City University London

Doctoral Thesis

Mortality: Modelling, Socio-Economic Differences and Basis Risk

Author:
Andrés M. Villegas Ramírez

Supervisor:
Dr. Steven Haberman

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

Faculty of Actuarial Science and Insurance
Cass Business School

August 2015

Declaration of Authorship

I, Andrés M. Villegas Ramírez, declare that this thesis titled, 'Mortality: Modelling, Socio-Economic Differences and Basis Risk' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only simple copies made for study purpose, subject to the normal conditions of acknowledgement.

Signed:

Date:

“Mortality rates are not just frivolous statistics, they are a reminder to cherish each and every moment with the people we love."

“Las tasas de mortalidad no son solo estadísticas frívolas, son también una advertencia para que apreciemos cada momento con las personas que queremos.”

CITY UNIVERSITY LONDON

Abstract

Faculty of Actuarial Science and Insurance
Cass Business School

Doctor of Philosophy

Mortality: Modelling, Socio-Economic Differences and Basis Risk

by Andrés M. Villegas Ramírez

During the last two centuries the developed world experienced a persistent increase in life expectancy. Although past trends suggests that life expectancy will continue to increase, there is considerable uncertainty surrounding the future evolution of mortality. In addition, past mortality improvements have not been shared equally across the population, resulting in a widening of socio-economic inequalities in mortality. The uncertainty and socio-economic variability of life expectancy pose a challenge for the design of pension systems and the management of longevity risk in pension funds and annuity portfolios.

This thesis is devoted to the investigation of the trends and financial implications of socio-economic differences in mortality. It comprises three parts. The first part introduces new modelling techniques for the quantification of socio-economic mortality differentials in aggregate and cause-specific mortality, which are applied in the study of the relationship between mortality and deprivation in the English population. The second part evaluates the suitability of several multipopulation stochastic mortality models for assessing basis risk in longevity hedges and provides guidelines on how to use these models in practical situations. Finally, the third part introduces new modelling tools which aim to permit a more effective and widespread use of stochastic mortality models.

Acknowledgements

The best outcome from this PhD has been starting a life with my loving and supporting wife Maria. I thank her for being my perfect complement and laughing with me in the ups and crying with me in the downs. I am fortunate to have her by my side and look forward to many more adventures together.

I would like to start by thanking my supervisor Steven Haberman for his intellectual generosity. The things I have learned from him go far beyond mortality modelling. He is a role model of what an academic should be like.

There are numerous people who have contributed to the development of this thesis and to my academic training during my time at Cass. I am indebted to Professors Vladimir Kaishev and Pietro Millossovich for their guidance, advice and support; working with them so closely (and intensively) on the research presented in Chapters 4 and 5 of this thesis has been a fantastic experience. I would like to thank Maddy Bajekal for the many insightful conversations on socio-economic differences in mortality and for showing me the value of interdisciplinary research. I would also like to thank Professors Ales Cerny, Andreas Tsanakas and Jens Nielsen for giving me the opportunity to work with them and sharing their vast academic experience with me. I am grateful to Les Mayhew, Jaap Spreeuw and David Smith for their very constructive comments on the MRes Dissertation and Transfer Document which preceded this thesis. I would like to thank Ana Debón for her willingness to discuss ideas and for her feedback on several chapters of this thesis. I wish to thank Andrew, James, Mario, Sveinn and Steven for having made me feel welcome during the time I spent working with them at Hymans Robertson. I am grateful to Chloe, Anna, Adina for all the administrative support over these years. Malla and Abdul deserve an especial mention, I thank them for the administrative support but most importantly for their friendliness and genuine interest in the best for the PhD students at Cass. I thank Santander Bank, the Institute and Faculty of Actuaries and the Life and Longevity Market Association for the funding they provided in the last two years of my studies.

I would like to thank my PhD colleagues and friends Natalia, Juraj, Hazel, Munir and Kaustav for having made this PhD such an enjoyable experience. I am especially indebted to my friend, colleague and co-author Andrew Hunt; this PhD would not have

been the same without having him at the other side of the office whether to bounce ideas or just to listen to my occasional rants about Colombian politics. I cannot forget my friend Juan Fernando, he became like a brother to me and I thank him for letting me see the world from a different perspective.

My deepest gratitude to my professional mentors and friends Luis Gutierrez and Armando Zarruk. They have sown the seeds for my passion for actuarial science and without their encouragement this journey would have never started. I would also like to thank my former supervisors Professors Juan Dario Restrepo, Jaime Londoño, Andrés Medaglia and Luis Fernando Zuluaga for sharing their experience and playing a key role in my formation as a researcher. I am grateful to Angela, Karol, Luisa and Camila, my ex-colleagues from Suramericana, keeping in touch with them has made me feel closer to my country in spite of the distance.

These years in London have given me the opportunity to meet very good friends or reencountered with old ones. I thank Valeria and Andrea for being our family while in the UK. I would also like to thank Juan David and Erica for the many wonderful nights we spent together. I am grateful for having met Humberto and Carolina, knowing another couple experiencing the same things we were experiencing made things much easier and I hope this friendship is for life. I thank my friends at Olympia Handball Club with whom we have had some of the best times of our lives. I am really grateful to Sebas and Paula, having them around for the first year of the PhD made the transition much easier. I thank my friend Valentina for her random Skype calls which cheered me up on some grey London days.

I cannot thank enough to my brother and colleague Juan Guillermo for always being there for his little brother and have the right advice whether personal or academic.

Finally but most importantly, I want to thank my parents for their love and unconditional support over all these years. Their example has made me who I am and a PhD in actuarial science is the natural culmination of a childhood spent between a university and an insurance company, thanks to a dad passionate for academia and a mum who devoted her life to the insurance industry.

Contents

List of Figures

List of Tables

To the loved ones who have already left us...
Para los seres queridos que ya partieron...

1

Introduction

During the last two centuries the developed world has experienced a persistent increase in life expectancy. For instance, data from the Human Mortality Database (2014) shows that life expectancy at birth in England and Wales passed from 41.59 years in 1841 to 81.04 years in 2011.

This increase in life expectancy, though a sign of social progress, poses a challenge to governments, private pension plans and life insurers as it can have a potential impact on pension and health costs. However, this challenge does not arise as a consequence of the increase in life expectancy by itself. If mortality improvements were fully anticipated, they could be properly managed. Rather, the real challenge emerges as a result of the uncertainty surrounding the increases in life expectancy. The magnitude of this uncertainty is illustrated by the ongoing debate among demographers on the extent to which past increases in life expectancy will continue into the future. Whereas some demographers (Oeppen and Vaupel, 2002), based on recent mortality trends, argue that no natural limit can be set to life expectancy, others (Olshansky et al., 2005) suggest that life expectancy might level off or even decline.

The uncertainty of future mortality trends, i.e. longevity risk, can result in important economic implications for individuals, pension plans and annuity providers. On the one hand, individuals face the risk of outliving their savings and having to reduce their living standards at old ages. On the other hand, pension plans and annuity providers may incur significant financial losses if they consistently underestimate the survival probabilities of their pensioners and policyholders. These potential consequences of longevity risk are

particularly acute in current times when the baby boom generation is starting to retire and real interest rates are at low levels.

Besides the uncertainty in life expectancy, there exists a considerable variability in the way mortality improvements are spread over the population. For instance, in England and Wales workers in different occupational classes not only have different life expectancies, but have also experienced significantly different mortality improvements. Whilst 65 year old professional males have gained 4.6 years of life expectancy increasing from a life expectancy of 14.0 in 1972-1976 to a life expectancy of 18.6 in 2002-06, unskilled manual workers have only gained 2.9 years increasing from 11.6 in 1972-1976 to 14.5 in 2002-06 (Johnson, 2011).

These socio-economic differences in mortality result in additional challenges for the design of pension systems and the management of longevity risk in pension funds and annuity portfolios. For instance, the equity and solidarity of national pensions systems can be considerably undermined by socio-economic mortality differentials as they may induce an undesirable redistribution of wealth from less affluent socio-economic groups with shorter life expectancy to more affluent socio-economic groups with above average longevity (Brown, 2002; Liebman, 2002; Whitehouse and Zaidi, 2008; Mazzaferro et al., 2012). Moreover, the inadequate allowance of mortality heterogeneity in the pricing and reserving of annuities can lead to annuities only being attractive to a limited portion of the population (Knox and Tomil, 1997) or to an inadequate funding of annuity and pension obligations (Meyricke and Sherris, 2013). Further, the flourishing of the emerging market of standardised longevity securities requires a good understanding of socio-economic mortality differentials as they are in most situations the main determinant of the basis risk associated with index-based longevity hedges (Coughlan et al., 2011).

The successful addressing of the challenges derived from longevity risk requires the best possible understanding of the past and possible future evolution of mortality improvements and of how these improvements have been shared across the society. To contribute to this understanding, this PhD thesis is devoted to the investigation of the trends and financial implications of socio-economic differences in mortality. Specifically, it introduces modelling techniques for the assessment of the magnitude of mortality differentials and for the examination of their possible future evolution.

The technical backbone of this thesis is the rapidly growing actuarial and demographic literature looking at the simultaneous modelling of mortality in two (or more) related populations (e.g. Li and Lee (2005); Jarner and Kryger (2011); Cairns et al. (2011a); Dowd et al. (2011); Li et al. (2015)). These multiple population stochastic mortality models provide a flexible tool which allows the assessment of historical mortality differentials as well as the forecasting of future mortality trends while taking into account the uncertainty around these trends. As such, this thesis also makes some contributions to the emerging area of multiple population mortality modelling and to the more mature field of single population stochastic mortality modelling.

This thesis comprises three parts: the first part is devoted to the modelling and forecasting of socio-economic differences in mortality with particular reference to the English population; the second part presents a rigorous comparison of two-population models for the assessment of basis risk in longevity hedges; and the third part presents some general tools for mortality modelling which were by-products of the research carried out in the first two parts. The document is composed of five self-contained chapters stemming from five research papers. These papers were developed in collaboration with co-authors who are mentioned at the beginning of each chapter. Being self-contained, each chapter has its own introduction, notation, conclusions and references. In addition the last chapter presents global conclusions and future research perspectives. A brief description of the contributions of each chapter follows.

Chapter 2: On the Modelling and Forecasting of Socio-economic Mortality Differentials: An Application to Deprivation and Mortality in England

This chapter has been published in the North American Actuarial Journal in 2014, Volume 18, issue 1, pages 168-193

In this chapter we discuss the suitability for the modelling and forecasting of socio-economic differences in all-cause mortality of several multiple population extensions of the Lee-Carter model (Lee and Carter, 1992), including a newly introduced relative model based on the modelling of the mortality in socio-economic subpopulations alongside the mortality of a reference population. We then use these modelling techniques

to analyse the extent of mortality differentials across deprivation subgroups in England and to evaluate the impact of these differentials on the valuation of annuities.

Chapter 3: Modelling Mortality by Cause of Death and Socio-economic Stratification: Trends and Projections for England, 1981-2030

This chapter is a research collaboration with Madhavi Bajekal from University College London, who, besides providing the data, has assisted with the epidemiological background for the modelling assumptions and for the interpretation of the research findings.

After looking at all-cause mortality differentials in Chapter 2, in this chapter we turn our attention to socio-economic difference in cause-specific mortality and introduce new modelling techniques that enable the modelling and projection of mortality trends by cause of death and socio-economic stratification.

One of the key challenges for the modelling of mortality by cause of death is posed by the frequent changes in the classification of causes of death which, if not taken into account appropriately, can compromise the validity of any analysis of trends in cause-specific mortality. To deal with this issue, in this chapter we first extend the widely-used model of Lee and Carter (1992) to enable the consideration of coding changes in cause-specific mortality data. Then, we embed this model into a multiple population setting to allow for the consideration of socio-economic differences in mortality. Finally, in order to complement the analysis of all-cause mortality differentials in England performed in Chapter 2 and to shed light on the drivers behind these differentials, we apply the newly introduced modelling approach to the examination of the relationship between deprivation and mortality for the leading causes of death in England.

Chapter 4: A Comparative Study of Two-population Models for the Assessment of Basis Risk in Longevity Hedges

The majority of this chapter has been carried out within the framework of a joint research project between Cass Business School and Hymans Robertson LLP. This project was commissioned by the Longevity Basis Risk Working Group of the Institute & Faculty of Actuaries and the Life & Longevity Markets Association with the objective of developing a practical methodology for assessing basis risk for longevity transactions. Therefore,

part of this chapter draws from the technical report documenting this research. The paper underpinning this chapter is joint work with Steve Haberman, Vladimir Kaishev and Pietro Millossovich. We have all participated in the development of the research ideas in the paper, while I have been in charge of their computational implementation and of the majority of the drafting of the document.

In Chapters 2 and 3 we found that mortality inequalities are persistent and seem to be widening, which has implications for the managing of longevity risk in pension funds and annuity portfolios. In particular, the basis risk emerging from the mismatch between the socio-economic characteristics of annuitants and those of the national population is one of the key obstacles preventing many pension schemes and insurers from considering index based longevity transactions as a legitimate alternative for the management of their longevity risk (LLMA, 2012). Two-population stochastic mortality models offer an alternative to overcome this obstacle as they allow market participants to compare and project the mortality experience for the reference and target populations and thus assess the amount of demographic basis risk involved in an index-based longevity hedge. Aiming to contribute to the development of the market of standardised longevity transactions, in this chapter we systematically evaluate the suitability of several multi-population stochastic mortality models for assessing basis risks and provide guidelines on how to use these models in practical situations.

Chapter 5: StMoMo: An R Package for Stochastic Mortality Modelling

This chapter is joint work with Vladimir Kaishev and Pietro Millossovich. We have all participated in the development of the research ideas in the chapter, while I have been in charge of their computational implementation and of the majority of the drafting of the document.

Due to the significant number of two-population mortality models that have recently been proposed in the literature, the rigorous comparison of models carried out in Chapter 4 was a major empirical exercise. To facilitate this exercise, we developed general **R** code which permitted the implementation of many of the models under a unifying modelling framework. Since we believe that the wider actuarial science community could benefit from these developments, we have turned part of this code into a publicly available **R** package, whose usage is documented in this chapter. More specifically, in this chapter

we define the family of Generalised Age-Period-Cohort (GAPC) stochastic mortality models and introduce the **R** package **StMoMo**, which exploits the unifying framework of the GAPC family to provide computational tools for implementing many of the single-population stochastic mortality models proposed to date.

Chapter 6: Robustness and Convergence in the Lee-Carter Model with Cohort Effects

The research presented in this chapter has been carried out in collaboration with Andrew Hunt, a fellow PhD Student at Cass Business School. Both of us have contributed equally to the development of the ideas and to the writing of the associated paper. This chapter has been published in Insurance: Mathematics and Economics in 2015, Volume 64, pages 186-202

The model for assessing socio-economic mortality differentials introduced in Chapter 2 has as one of its integral parts the Lee-Carter model with cohort effects proposed by Renshaw and Haberman (2006). During the implementation of this latter model we encountered numerous robustness and convergence issues which prompted the development of this chapter. The existence of robustness and convergence problems with the Lee-Carter model with cohort effects is not new and several author have conjectured that such problems are the result of an unresolved identifiability issue (Cairns et al., 2009, 2011b; Haberman and Renshaw, 2011; Currie, 2014). In this chapter, after investigating systemically the robustness of cohort extensions of the Lee-Carter model and the convergence of the algorithms used to fit it to data, we demonstrate the existence of such an identifiability issue and propose an additional approximate identifiability constraint which solves many of the problems found.

Bibliography

Brown, J. R., 2002. Differential mortality and the value of individual account retirement annuities. In: Feldstein, M., Liebman, J. B. (Eds.), The Distributional Aspects of Social Security and Social Security Reform. University of Chicago Press.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., 2011a. Bayesian stochastic mortality modelling for two populations. ASTIN Bulletin 41, 29–59.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2011b. Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics 48 (3), 355–367.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1–35.

Coughlan, G. D., Khalaf-Allah, M., Ye, Y., Kumar, S., Cairns, A. J., Blake, D., Dowd, K., 2011. Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. North American Actuarial Journal 15 (2), 150–176.

Currie, I. D., 2014. On fitting generalized linear and non-linear models of mortality. Scandinavian Actuarial Journal.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., Khalaf-allah, M., 2011. A gravity model of mortality rates for two related populations. North American Actuarial Journal 15 (2), 334–356.

Haberman, S., Renshaw, A., 2011. A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics 48 (1), 35–55.

Human Mortality Database, 2014. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL www.mortality.org

Jarner, S. F., Kryger, E. M., 2011. Modelling adult mortality in small populations: The Saint Model. ASTIN Bulletin 41 (2), 377–418.

Johnson, B., 2011. Deriving trends in life expectancy by the National Statistics Socio-economic Classification using the ONS Longitudinal Study. Health Statistics Quarterly (49), 9–52.

Knox, D., Tomil, A., 1997. An analysis of pensioner mortality by pre-retirement income. Working Paper Series. Centre for Actuarial Studies, University of Melbourne.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Li, J. S.-H., Zhou, R., Hardy, M. R., 2015. A step-by-step guide to building two-population stochastic mortality models. Insurance: Mathematics and Economics.

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

Liebman, J., 2002. Redistribution in the current US social security system. In: Feldstein, M., Liebman, J. (Eds.), The Distributional Aspects of Social Security and Social Security Reform. University of Chicago Press.

LLMA, 2012. Basis risk in longevity hedging: parallels with the past. Institutional Investor Journals 2012 (1), 39–45.

Mazzaferro, C., Morciano, M., Savegnago, M., 2012. Differential mortality and redistribution in the Italian notional defined contribution system. Journal of Pension Economics and Finance, 500–530.

Meyricke, R., Sherris, M., 2013. The determinants of mortality heterogeneity and implications for pricing annuities. Insurance: Mathematics and Economics 53 (2), 379–387.

Oeppen, J., Vaupel, J. W., 2002. Broken limits to life expectancy. Science 296 (5570), 1029–1031.

Olshansky, S., Passaro, D., Hershow, R., Layden, J., Carnes, B., Brody, J., Hayflick, L., Butler, R., Allison, D., Ludwig, D., 2005. A potential decline in life expectancy in the United States in the 21st century. New England Journal of Medicine 352 (11), 1138–1145.

Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556–570.

Whitehouse, E. R., Zaidi, A., 2008. Socio-economic differences in mortality: Implications for pensions policy. OECD Social, Employment and Migration Working Papers, No. 71, OECD Publishing, Paris.

Part I

Socio-Economic Differences in Mortality

2

On the Modelling and Forecasting of Socio-economic Mortality Differentials: An Application to Deprivation and Mortality in England

This chapter has been published in the North American Actuarial Journal. The full reference follows:

– Villegas, A. M., Haberman, S., 2014. On the modeling and forecasting of socio-economic mortality differentials: an application to deprivation and mortality in England. North American Actuarial Journal 18 (1), 168–193.

Previous versions of this chapter were presented at the following conferences:

– June 2012. Demographic Analysis and Research International Conference, Chania, GREECE. "A model for the forecasting of socio-economic mortality differentials".

– September 2012. Longevity 8, Waterloo, CANADA. "A model for the forecasting of socio-economic mortality differentials".

– January 2013. Perspectives on Actuarial Risks in Talks of Young Researchers, Ascona, SWITZERLAND. "On the modelling and forecasting of socio-economic mortality differentials: an application to deprivation and mortality in England".

– March 2013. Socio-demographics: Exploring the Future and Defining the Questions, Institute and Faculty of Actuaries, London, UK. "On the modelling and forecasting of socio-economic mortality differentials: an application to deprivation and mortality in England". Video available at http://openchannel.multichanneltv.com/the-actuarial-profession/socioeconomic-inequalities/platform.php?noCPD

On the Modelling and Forecasting of Socio-economic Mortality Differentials: An Application to Deprivation and Mortality in England

Andrés M. Villegas, Steven Haberman

Cass Business School, City University London, United Kingdom

Abstract

In any country, mortality rates and indices such as life expectancy usually differ across subpopulations, for example, defined by gender, geographic area or socio-economic variables (e.g. occupation, level of education, income). These differentials, and in particular those related to socio-economic circumstances, pose important challenges for the design of public policies for tackling social inequalities, as well as for the design of pension systems and the management of longevity risk in pension funds and annuity portfolios. We discuss the suitability for the modelling and forecasting of socio-economic differences in mortality of several multiple population extensions of the Lee-Carter model, including a newly introduced relative model based on the modelling of the mortality in socio-economic subpopulations alongside the mortality of a reference population. Using England mortality data for socio-economic subpopulations defined using a deprivation index, we show that this new relative model exhibits the best results in terms of goodness of fit and ex-post forecasting performance. We then use this model to derive projections of deprivation specific mortality rates and life expectancies at pensioner ages and analyse the impact of socio-economic differences in mortality on the valuation of annuities.

Keywords: Mortality modelling; multipopulation models; socio-economic circumstances; annuity pricing

2.1 Introduction

In any country, mortality rates and indices such as life expectancy usually differ across subpopulations, for example, defined by gender, geographic area or socio-economic variables. In particular, there is a well established inverse relationship between socio-economic circumstances – whether measured by educational attainment, occupation, income or area deprivation – and mortality, with higher socio-economic subgroups having lower mortality rates and, in most cases, also experiencing faster mortality improvements than lower socio-economic subgroups (see, e.g., Shkolnikov et al. (2006); Johnson (2011); Tarkiainen et al. (2012); Raleigh and Kiri (1997)).

These socio-economic differences in mortality not only pose significant challenges for the design of public policies for tackling social inequalities, but also for the design of pension systems and the management of longevity risk in pension funds and annuity portfolios. On the one hand, differential mortality can have important consequences on the redistribution properties of both defined benefit and defined contribution pension schemes (Liebman, 2002; Brown, 2002), for example, undermining the equity and solidarity of national pensions systems by inducing an undesirable transfer of wealth away from lower socio-economic groups with shorter life expectancy to higher socio-economic groups with above average longevity. On the other hand, the ignorance of mortality heterogeneity when valuing pension liabilities or pricing annuities could result in an inadequate funding of annuity and pension obligations. Furthermore, the successful development of a market of standardised longevity securities requires a good understanding of socio-economic mortality differentials as they are in most situations the main determinant of the basis risk associated with index-based longevity hedges (Coughlan et al., 2011).

In view of this, there is a need for methods that help us assess the magnitude of socio-economic mortality differentials within a population and that enable us to examine their possible future evolution. In general, to measure and project mortality differentials we require a modelling approach that permits the simultaneous modelling of mortality in a group of subpopulations. Moreover, in the specific context of socio-economic subpopulations, such a model should ideally capture both mortality level differentials and mortality improvement differentials among the subpopulations, that is, differentials in the average level of mortality and differentials in the pace of mortality change.

Some authors have suggested the application of generalised linear models (Madrigal et al., 2011) and survival models (Richards, 2008) in the quantification of socio-economic mortality differentials. These statistical methods have specifically been proposed for the assessment of baseline (level) mortality differentials, ignoring, in part due to the lack of appropriate data, the differences in improvements by socio-economic characteristics and the modelling of their possible future evolution. Provided that data requirements are met, a modelling alternative that does allow the consideration of both level and trend differentials in mortality, as well as the projection of their future evolution, is offered by the numerous stochastic mortality models that have recently been proposed for the simultaneous modelling and forecasting of mortality in a group of populations. Unfortunately, most of these multipopulation models have not been designed with the aim of assessing socio-economic differences in mortality, but with the purpose of comparing the mortality evolution of a group of countries or of genders or regions within a country. As a result, some of these models may lack some of the desirable features of an approach for the modelling and forecasting of mortality in a group of socio-economic subpopulations, which include: transparency for the disentangling of level and improvement differentials in mortality; consistency of subpopulation-specific mortality forecasts with national mortality forecasts; ability to produce adequate interval forecasts of mortality differentials; and ability to produce mortality rates forecast that preserve the inverse relationship between socio-economic circumstances and mortality.

In this chapter, we discuss the suitability for the modelling and forecasting of socio-economic differences in mortality of several multiple population extensions of the Lee-Carter model (Lee and Carter, 1992), including a newly introduced relative model based on the modelling of the mortality in socio-economic subpopulations alongside the mortality of a reference population. This latter model has especially been designed taking into consideration the characteristics of socio-economic subpopulations as well as some of the typical issues of mortality data disaggregated by socio-economic circumstances. Although previous studies have investigated the use of multipopulation mortality models in the assessment of mortality differentials among two countries (Li and Hardy, 2011) and regions within a country (Debón et al., 2011), this is the first study to apply such models to socio-economic mortality differentials in particular. More specifically, we use multipopulation mortality models to analysis the extent of mortality differentials across deprivation subgroups in England.

The reminder of this chapter is organised as follows. Section 2.2 describes several alternative methods for the measurement and projection of mortality differentials. Particularly, Section 2.2.4 introduces our proposed new relative modelling approach. In Section 2.3, we apply these models in the examination of the relationship between deprivation and mortality in the English population, emphasising on its implications for the valuation of life annuities. Finally, Section 2.4 concludes with a discussion of our main findings.

2.2 Modelling socio-economic mortality differentials

In this section, we describe several multipopulation extensions of the Lee-Carter model (Lee and Carter, 1992), and discuss their suitability for the modelling of socio-economic mortality differentials. All these models propose a parametric representation of the central death rate ${}_n\mu_{xtg}$ in year t for people age $[x, x+n)$ in subpopulation g, based on a cross-classified mortality experience containing the observed number of deaths ${}_nd_{xtg}$ at ages $[x, x+n)$, $x \in \mathcal{X} := \{x_1, \ldots, x_k\}$, in year t, $t \in \mathcal{T} := \{t_1, \ldots, t_n\}$, for subpopulation g, $g \in \mathcal{G} := \{g_1, \ldots, g_m\}$, with matching exposures ${}_ne_{xtg}$.

2.2.1 Independent modelling

The simplest approach for modelling mortality in a set of subpopulations would be to use independent unrelated Lee-Carter models for each subpopulation. Thus, we could model mortality in each subpopulation $g \in \mathcal{G}$ using the specification

$$\log {}_n\mu_{xtg} = \alpha_{xg} + \beta_{xg}\kappa_{tg}, \tag{2.1}$$

where α_{xg} captures the general age-specific mortality pattern for subpopulation g, κ_{tg} is a time varying mortality index representing the overall level of mortality in year t for subpopulation g, and β_{xg} measures the age-specific response to changes in the general level κ_{tg}. The parameters of this model are identifiable only up to a transformation as for any constants c_1 and $c_2 \neq 0$, if we replace α_{xg} by $\alpha_{xg} + c_1\beta_{xg}$, β_{xg} by $\frac{1}{c_2}\beta_{xg}$, and κ_{tg} by $c_2(\kappa_{tg} - c_1)$, equation (2.1) will produce the same log death rates. Therefore, in

order to ensure the identifiability of the model, the constraints

$$\sum_{x\in\mathcal{X}} \beta_{xg} = 1, \quad \sum_{t\in\mathcal{T}} \kappa_{tg} = 0$$

are imposed.

In this model mortality forecasts are obtained by modelling and forecasting the period indexes using independent univariate ARIMA processes. In most applications, each period index k_{tg}, $g \in \mathcal{G}$, can be modelled using a random walk with drift (ARIMA(0,1,0)):

$$\kappa_{tg} = d_g + \kappa_{t-1,g} + \xi_{tg}, \tag{2.2}$$

where d_g is the drift term and ξ_{tg} is a normally distributed error term with zero mean and variance σ_g^2. We note that under the independent modelling approach the error terms for each subpopulation ξ_{tg}, $g \in \mathcal{G}$, are assumed to be independent of one another.

The independent modelling approach is straightforward to implement. However, it has several shortcomings. First, it assumes no interdependence among the mortality of the subpopulations, a very unrealistic assumption for socio-economic subpopulations within a country, which are likely to follow similar mortality trends. Second, although mortality level differentials could be assessed by comparing the α_{xg} terms and trend differentials by comparing the β_{xg} and κ_{tg} terms, this is not a straightforward task.

The assumption of complete independence among the subpopulations can be relaxed by, instead of using independent random walks to model the time indexes, using a multivariate random walk with drift so that

$$\boldsymbol{\kappa}_t = \mathbf{d} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{2.3}$$

where $\boldsymbol{\kappa}_t := (\kappa_{t,1}, \kappa_{t,2} \ldots, \kappa_{t,m})'$, $\mathbf{d}$ is a vector of drift parameters and $\boldsymbol{\Sigma}$ is the variance-covariance matrix. Since the diagonal terms of the variance-covariance matrix $\boldsymbol{\Sigma}$ coincide with the variances σ_g^2, $g \in \mathcal{G}$, of ξ_{tg} in (2.2), the central and interval projections of mortality rates for each subpopulation derived from the multivariate random walk and the independent random walk with drifts are identical. However, the multivariate random walk yields more precise interval projections of mortality differentials as the

non-diagonal terms of the variance-covariance matrix $\mathbf{\Sigma}$ capture the dependence among the subpopulations.

A further enhancement of the independent approach would be the use of co-integration methods to model $\boldsymbol{\kappa}_t$ as suggested by Carter and Lee (1992) and described in Li and Hardy (2011) and in Yang and Wang (2013). However, although theoretically sound, mortality data disaggregated by socio-economic variables are typically available for short periods of time, hampering the application of the required econometric techniques.

2.2.2 The joint κ-model

A second alternative for modelling mortality differentials is the joint-κ model proposed by Carter and Lee (1992). This extension of the Lee-Carter model assumes that all the subpopulations are driven by a single period index. Formally, the model can be expressed as

$$\log {}_n\mu_{xtg} = \alpha_{xg} + \beta_{xg}\kappa_t,$$

where α_{xg} describes the average age profile for subpopulation g, κ_t is a period index driving the mortality trend for all subpopulations, and β_{xg} represents the age-subpopulation-specific pattern of mortality change.

In order to facilitate the assessment of mortality differentials, it is convenient to re-parametrised the model as[1]:

$$\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + (\beta_x + \beta_{xg})\kappa_t \tag{2.4}$$

with the following constraints

$$\sum_{x\in\mathcal{X}} \beta_x = 1, \quad \sum_{t\in\mathcal{T}} \kappa_t = 0$$
$$\sum_{g\in\mathcal{G}} \alpha_{xg} = 0, \quad \sum_{g\in\mathcal{G}} \beta_{xg} = 0 \quad \text{for all } x \in \mathcal{X}$$

[1] Delwarde et al. (2006) have considered a similar parametrisation in the simultaneous modelling of mortality in five developed countries.

to ensure the identifiability of the model. This parametrisation allows the measurement of subpopulation mortality as deviations from the general mortality pattern of the population. Thus, parameters α_x and α_{xg} capture mortality level differentials, whereas parameters β_x and β_{xg} capture mortality improvement differentials. Specifically, the term $\exp(\alpha_{xg})$ quantifies the average percentage deviation of subpopulation g from the the general level of mortality in the population $\exp(\alpha_x)$. That is, if $\exp(\alpha_{xg}) > 1$ then at ages $[x, x+n)$ mortality in subpopulation g is higher than in the total population, and if $\exp(\alpha_{xg}) < 1$ then mortality is lower. Similarly, if $\beta_{xg} > 0$ then at ages $[x, x+n)$ mortality in subpopulation g is improving at a faster pace than in the total population and if $\beta_{xg} < 0$ then mortality is improving at a slower rate.

Whilst the independent modelling approach includes a period index for each subpopulation, the joint-κ has a single period index. Hence, mortality forecast for all the subpopulations can be derived by modelling this period index using a univariate random walk with drift

$$\kappa_t = d + \kappa_{t-1} + \xi_t$$

Besides the transparency provided by parametrisation (2.4) for the identification of both level and trend differentials in mortality, there are other compelling statistical and demographics reasons that suggest that the joint-κ may be a very attractive approach. Statistically, a single time index is a very parsimonious way of linking the mortality of multiple populations. Demographically, a single mortality driver may impose greater consistency among the subpopulations, ruling out the possibility that the mortality of the subpopulations evolves in completely different ways. Nevertheless, a single mortality driver also implies that the improvement rates of the subpopulations will be perfectly correlated, resulting in extremely narrow interval forecasts of mortality differentials.

There are two restricted versions of the joint-κ model which are worth exploring as alternatives for modelling mortality differentials: the "common factor" model introduced by Li and Lee (2005) and the stratified or additive Lee-Carter model considered in Butt and Haberman (2009) and in Debón et al. (2011). The common factor model, obtained

by setting $\beta_{xg} = 0$, can be expressed as

$$\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + \beta_x \kappa_t$$

and the stratified Lee-Carter, which also sets $\alpha_{xg} = \alpha_g$, is given by

$$\log {}_n\mu_{xtg} = \alpha_x + \alpha_g + \beta_x \kappa_t \tag{2.5}$$

These two variants of the joint-κ model, although very parsimonious, may be too stringent for some applications. Both models assume the same mortality improvements for all subpopulations, which implies that improvement differentials in mortality are assumed to be non-existent and makes it impossible to draw any conclusion as to whether relative mortality differentials are increasing or decreasing. In addition, the stratified Lee-Carter model ignores any variation of mortality differentials with age, failing to capture the commonly observed decrease in socio-economic differentials in mortality with rising age (Hoffmann, 2005).

2.2.3 The three-way Lee-Carter

The three-way extension of the Lee-Carter model proposed by Russolillo et al. (2011) provides an additional alternative for modelling mortality differentials. This variant of the Lee-Carter model adds a subpopulation parameter that deals with trend differences in mortality. Specifically, the model is given by

$$\log {}_n\mu_{xtg} = \alpha_{xg} + \beta_x \lambda_g \kappa_t,$$

where α_{xg} measures the age-subpopulation-specific pattern of mortality, β_x and κ_t have the same interpretation as in the Lee-Carter model, and λ_g is a new term capturing the variability in the improvement rates of the subpopulations. As with the joint-κ model, the interpretation of mortality differentials is made easier if we consider the re-parametrisation

$$\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + \beta_x \lambda_g \kappa_t,$$

with constraints

$$\sum_{g\in\mathcal{G}} \alpha_{xg} = 0 \quad \text{for all } x \in \mathcal{X}$$
$$\sum_{x\in\mathcal{X}} \beta_x = 1, \quad \sum_{t\in\mathcal{T}} \kappa_t = 0, \quad \frac{1}{m}\sum_{g\in\mathcal{G}} \lambda_g = 1$$

to ensure identifiability of the model. In this new parametrisation parameters α_x and α_{xg} capture level differentials in mortality and have the same interpretation as in the joint-κ model. Improvement differentials in mortality are captured by parameters λ_g, with $\lambda_g > 1$ ($\lambda_g < 1$) meaning that mortality in subpopulation g improves at a faster (slower) rate than mortality in the general population. As in the joint-κ model mortality forecasts are obtained by modelling the singe mortality driver κ_t using a univariate random walk with drift.

The joint-κ model and the three-way Lee-Carter share similar advantages and disadvantages. As the joint-κ model, the three-way Lee-Carter offers a parsimonious and transparent way of assessing mortality differentials, but, as a result of having a single mortality driver, it also suffers from very narrow interval forecasts of mortality differentials.

2.2.4 A relative modelling approach

Finally, in order to model socio-economic mortality differentials, we introduce a new relative modelling approach whereby subpopulation mortality is modelled relative to the mortality of a reference population, which would in most cases be the national population of the country from which the subpopulations come from. A relative approach offers several advantages when compared with the previously described approaches that rely exclusively on the mortality data of the socio-economic subpopulations. First, national mortality data are normally available for a longer period than mortality data disaggregated by socio-economic circumstances, permitting a more precise estimation of the long-run mortality trend. Moreover, modelling the subpopulations alongside the national population will ensure the consistency of the subpopulation-specific mortality forecasts with the national mortality forecasts. Second, as opposed to mortality data for socio-economic subpopulations which are normally available in an age-grouped format,

national mortality data are typically available for individual ages, facilitating the consideration of the effect of year of birth (cohort) in mortality which has been identified in some populations (see, e.g., Willets (2004) for the UK).

Thus, we assume that besides the subpopulation mortality experience $({}_nd_{xtg}, {}_ne_{xtg})$, there is available an additional experience containing the number of deaths d'_{xt} at age x, $x \in \mathcal{X}' =: \{x'_1, \ldots, x'_{k'}\}$, in year t, $t \in \mathcal{T}' := \{t'_1, \ldots, t'_{n'}\}$ for a reference population with corresponding exposures e'_{xt}, possibly covering a wider age range and a longer period of time than in the subpopulations' data, i.e., $x'_1 \leq x_1$, $x'_{k'} \geq x_k + n$, $t'_1 \leq t_1$, and $t'_{n'} \geq t_k$.

With this additional data at hand, we follow a modelling approach similar to that of Jarner and Kryger (2011) and model the subpopulations death rates, ${}_n\mu_{xtg}$, relative to the reference population death rates, μ'_{xt}. Specifically, we model the mortality rates of the reference population as

$$\log \mu'_{xt} = \alpha'_x + \beta'_x \kappa'_t + \gamma'_{t-x} \tag{2.6}$$

and the mortality rates of the subpopulations as

$$\log {}_n\mu_{xtg} = \log {}_n\bar{\mu}'_{xt} + \alpha_{xg} + \beta_x \kappa_{tg}, \tag{2.7}$$

where

$${}_n\bar{\mu}'_{xt} = \left(\prod_{i=0}^{n-1} \mu'_{x+i,t} \right)^{\frac{1}{n}} = \exp\left(\frac{1}{n} \sum_{i=0}^{n-1} \left(\alpha'_{x+i} + \beta'_{x+i} \kappa'_t + \gamma'_{t-x-i} \right) \right)$$

is the geometric average of the mortality rates in the reference population between age x and age $x + n - 1$. The identifiability of the model is ensured by imposing the following parameter constraints

$$\sum_{x\in\mathcal{X}'} \beta'_x = 1 \tag{2.8}$$

$$\sum_{t\in\mathcal{T}'} \kappa'_t = 0 \tag{2.9}$$

$$\sum_{\substack{z=t-x \\ t\in\mathcal{T}', x\in\mathcal{X}'}} \gamma'_z = 0 \tag{2.10}$$

$$\sum_{x\in\mathcal{X}} \beta_x = 1 \tag{2.11}$$

$$\sum_{t\in\mathcal{T}} \kappa_{tg} = 0 \quad \text{for all } g \in \mathcal{G} \tag{2.12}$$

The parametric structure defined by (2.6) was introduced by Renshaw and Haberman (2006) as a generalisation of the Lee-Carter model to allow for the consideration of cohort effects. In (2.6), α'_x captures the general age-specific mortality pattern in the reference population, κ'_t represents the overall time trend of mortality in the reference population, β'_x measures the age-specific response to changes in the general level of mortality, and γ'_{t-x} captures the cohort effect. This parametrisation assumes that cohorts effects are the same for all the subpopulations, which might be reasonable for socio-economic subpopulations within a country but might not be appropriate for the modelling of multiple populations of a different nature such as countries. We also note that for some applications the cohort effect might not be significant and one might consider the original Lee-Carter model, omitting γ'_{t-x} from the model.

Equation (2.7) models mortality in the subpopulations relative to mortality in the reference population. Within this parametric structure α_{xg} captures mortality level differentials, whilst β_x and κ_{tg} capture mortality improvement differentials. As in the joint-k and three-way Lee-Carter models, the term $\exp(\alpha_{xg})$ measures the average percentage deviation of subpopulation g from the pattern of mortality in the reference population. The subpopulation-specific time index κ_{tg} measures the deviations of mortality improvements in population g from the mortality improvements of the reference population. Therefore, a decreasing trend in κ_{tg} implies that mortality in subpopulation g is improving at a faster rate than in the reference population, while on the contrary, an increasing trend means that mortality is improving at a slower pace than in the reference population.

The age-modulating parameter β_x indicates the magnitude of mortality improvement differentials at each particular age. In principle, we could consider a model structure where the age-modulating parameter is subpopulation specific (i.e., β_{xg} as opposed to β_x) as in the augmented common factor model proposed by Li and Lee (2005). However, besides being more parsimonious, a subpopulation-independent specification of the age-modulating parameter is convenient in the forecasting of mortality in socio-economic subpopulations, where an ordering of mortality levels is natural (subpopulations with lower socio-economic conditions tend to have higher mortality than subpopulations with higher socio-economic conditions). Notice that if subpopulation g_1 has historically had lower mortality than subpopulation g_2 (i.e. $\alpha_{x,g_1} < \alpha_{x,g_2}$ for all $x \in \mathcal{X}$), then a sufficient condition for maintaining this ordering in the forecasted mortality rates (i.e. ${}_n\mu_{x,t_n+h,g_1} < {}_n\mu_{x,t_n+h,g_2}$, $h > 0$) is $\beta_x > 0$, and $(\kappa_{t_n+h,g_1} - \kappa_{t_n,g_1}) > (\kappa_{t_n+h,g_2} - \kappa_{t_n,g_2}), h > 0$. Hence, the problem of preserving the mortality ordering among the subpopulations reduces to modelling appropriately the multivariate time index κ_{tg}.

Mortality rate extrapolations for the subpopulations require time series forecasts of the multivariate period index $\boldsymbol{\kappa}_t := (\kappa_{t,1}, \kappa_{t,2} \ldots, \kappa_{t,m})'$ as well as projected values of the mortality rates in the general population, which in turn require times series forecasts of κ'_t and possibly of γ'_{t-x}, depending on the projected trajectories of interest. We employ a random walk with drift to model the reference population period index, and extrapolation of the cohort parameter is not required in the case study considered in this chapter. For the period index of the subpopulations, we consider a multivariate random walk with drift (see equation (2.3)) so that any potential dependence among the subpopulations is captured.

2.2.5 Fitting the models

To fit the models we consider that subpopulation death counts are independent Poisson responses ${}_nD_{xtg} \sim Poisson({}_ne_{xtg}\,{}_n\mu_{xtg})$ and derive parameter estimates by maximising the log-likelihood

$$\mathcal{L}({}_nd_{xtg}, {}_n\hat{d}_{xtg}) = \sum_{x\in\mathcal{X}}\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} \omega_{xtg}\left\{{}_nd_{xtg}\log {}_n\hat{d}_{xtg} - {}_n\hat{d}_{xtg} - \log {}_nd_{xtg}!\right\},$$

where ${}_n\hat{d}_{xtg}$ denotes the expected number of death predicted by the model and ω_{xtg} are 0-1 weights indicating empty or omitted data cells.

For the new relative model we suppose that the reference population death counts are also independent Poisson responses $D'_{xt} \sim Poisson(e'_{xt}\mu'_{xt})$ and estimate simultaneously the parameters of equations (2.6) and (2.7) by maximising the total model log-likelihood under the assumption of independence between D'_{xt} and ${}_nD_{xtg}$:

$$\begin{aligned}\mathcal{L}({}_nd_{xtg}, {}_n\hat{d}_{xtg}, d'_{xt}, \hat{d}'_{xt}) &= \mathcal{L}_{ref}(d'_{xt}, \hat{d}'_{xt}) + \mathcal{L}_{sub}({}_nd_{xtg}, {}_n\hat{d}_{xtg}) \\ &= \sum_{x\in\mathcal{X}'}\sum_{t\in\mathcal{T}'} \omega'_{xt}\left\{d'_{xt}\log\hat{d}'_{xt} - \hat{d}'_{xt} - \log d'_{xt}!\right\} \\ &+ \sum_{x\in\mathcal{X}}\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} \omega_{xtg}\left\{{}_nd_{xtg}\log{}_n\hat{d}_{xtg} - {}_n\hat{d}_{xtg} - \log{}_nd_{xtg}!\right\},\end{aligned} \tag{2.13}$$

where

$${}_n\hat{d}_{xtg} = {}_ne_{xtg}\exp\left(\frac{1}{n}\sum_{i=0}^{n-1}(\alpha'_{x+i} + \beta'_{x+i}\kappa'_t + \gamma'_{t-x-i}) + \alpha_{xg} + \beta_x\kappa_{tg}\right),$$

$$\hat{d}'_{xt} = e'_{xt}\exp\left(\alpha'_x + \beta'_x\kappa'_t + \gamma'_{t-x}\right),$$

and ω'_{xt}, ω_{xtg} are 0-1 weights that indicate empty or omitted data cells. In order to maximise the log-likelihood functions, we employ suitable straightforward extensions of the Newton-Raphson iterative procedure used by Brouhns et al. (2002) in the estimation of the Lee-Carter model.

It is well-known that as a result of the relationship cohort = period – age, age-period-cohort (APC) modelling is problematic (Renshaw and Haberman, 2006). Hence, the inclusion of a cohort effect in the modelling of the reference population of the relative model makes the fitting of this model complicated. Specifically, it has been reported that cohort-based extension of the Lee-Carter model have a slow rate of convergence and a lack of stability in the fitted parameters (Cairns et al., 2009). Cairns et al. (2009) suggest that these issues might be the result of a remaining identifiability problem, with the log-likelihood function being flat or approximately flat in certain dimensions. In fact, in Chapter 6 we show that if the period index κ'_t is approximately linear, as is the case in the mortality experience of most developed countries (Tuljapurkar et al., 2000), then

an approximately invariant parameter transformation arises. Therefore, to overcome the problems with the fitting of the relative model we use the results of Chapter 6 and add to constraints (2.8), (2.9), (2.10), (2.11) and (2.12) the constraint

$$\sum_{\substack{c=t-x \\ t\in\mathcal{T}',x\in\mathcal{X}'}} (c-\bar{c})\gamma'_c = 0, \tag{2.14}$$

where $\bar{c} = \frac{1}{k+n-1}\sum_{c=t-x,t\in\mathcal{T}',x\in\mathcal{X}'} c$, which combined with constraint (2.10) ensures that any linear trend in the cohort effect is eliminated, and improves the stability and convergence of the model.

2.3 Case study: Mortality by deprivation in England

In this section we employ the previously discussed models in the investigation of the relationship between socio-economic circumstances and mortality in England. To this end, we have derived a socio-economic classification of the English population using the Index of Multiple Deprivation 2007 (IMD 2007) which measures socio-economic circumstances at a small area level (Noble et al., 2007). The IMD 2007 is a composite index of deprivation comprising seven deprivation domains[2] and is calculated for each geographically defined Lower Layer Super Output Area (LSOA) in England. There are 32,482 LSOAs in England covering approximately 1,500 people each. In our analysis LSOAs are ranked and then grouped into deprivation quintiles based on their IMD 2007 score. For each deprivation quintile we have population and deaths estimates for the period 1981-2007, classified by sex and age groups 50-54,...,80-84. Throughout this chapter, we will refer to the deprivation quintiles as Q1, Q2, Q3, Q4, and Q5, with Q1 being the least deprived quintile of the population and Q5 the most deprived quintile. A particular feature of this dataset is that it addresses the issues of low volume, consistency and credibility that have been encountered in other similar datasets of mortality disaggregated by socio-economic circumstances for the English population. The mortality counts grouped by deprivation quintile were obtained from the ONS (Office for National

[2]The seven deprivations domains with their percentage participation in the index are: i) Income deprivation (22.5%), ii) Employment deprivation (22.5%), iii) Health deprivation and disability (13.5%), iv) Education, skills and training deprivation (13.5%), v) Barriers to housing and services (9.3%), vi) Crime (9.3%), and vii) Living Environment deprivation (13.5%).

Statistics, UK) and the corresponding populations were estimated. The details of the compilation of this dataset are described in Lu et al. (2014).

For the proposed new relative model, we consider deprivation-specific mortality relative to mortality in England and Wales. Thus, we use as reference population data the England and Wales mortality experience for calendar years 1961-2009 and individual ages $10, 11, \ldots, 99$, obtained from the Human Mortality Database (2012). Despite the fact that the available subpopulation data cover the age range 50-84 and the period 1981-2007, we use for the reference population a wider age range, 10-99, and a longer observation period, 1961-2009, so that more reliable estimates of the cohort effect and of the long-run mortality trend can be produced. We exclude ages below 10 years as infant mortality tends to exhibit a significantly different behaviour than mortality at older ages.

2.3.1 Comparison of the models

We first compare the models discussed in Section 2.2 in terms of their ability to fit and forecast the mortality differentials in the deprivation subpopulations. Thus, we have fitted the independent Lee-Carter model, the joint-κ model and its constrained variants the stratified Lee-Carter and common factor models, the three-way Lee-Carter model, and the new relative model using both the Lee-Carter model and the age-period-cohort model (2.6) for the England and Wales reference population. Due to space constraints, we limit our comparisons to the male population, but most of the conclusions of this section also apply to the female population. Figures 2.1 to 2.7 present the corresponding parameters estimates for the male population. These parameters plots reveal a clear association between deprivation and mortality in the English population. Specifically, α_g in the stratified Lee-Carter model and α_{xg} in all the other models show a marked inverse relationship between deprivation and mortality, with more deprived subpopulations having considerably higher mortality than less deprived ones. In addition, parameters β_{xg} of the joint-κ model, λ_g of the three-way Lee-Carter model, and κ_{tg} of the relative models indicate that less deprived subpopulations have experienced faster mortality improvements than most deprived ones. Although this last conclusion could

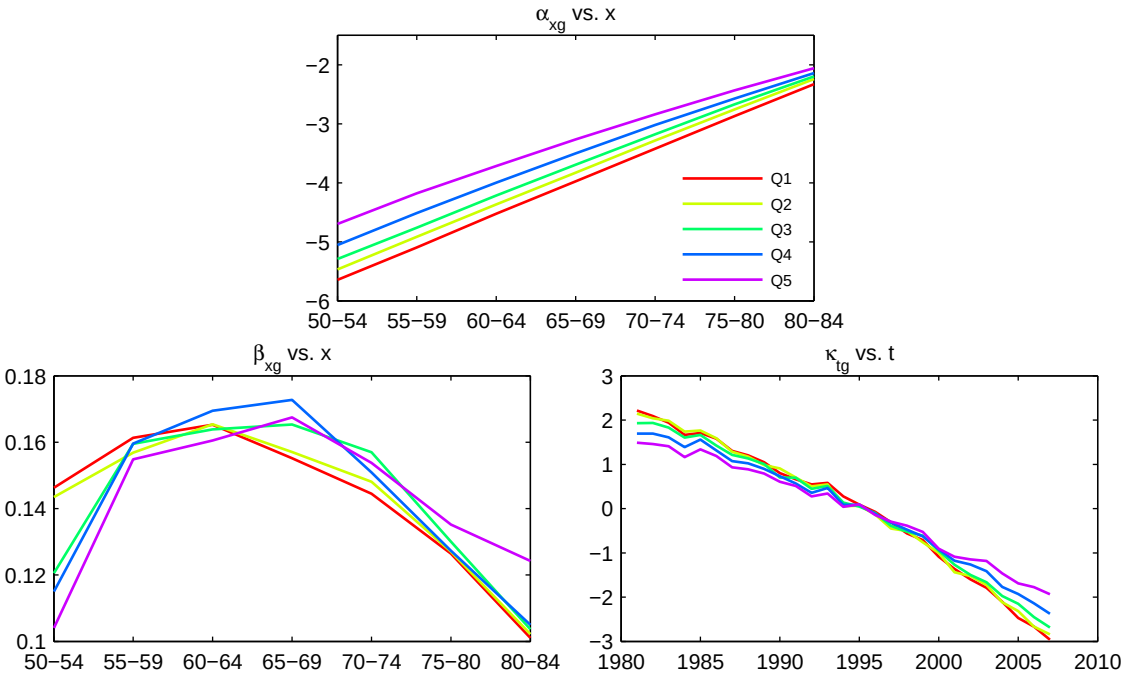


FIGURE 2.1: Parameter estimates of the independent Lee-Carter model for the male deprivation subpopulations.

also be drawn from the parameters of the independent Lee-Carter model, it is less obvious as it requires the analysis of the interactions between β_{xg} and κ_{tg}, underscoring one of the main disadvantages of this approach.

A first way of evaluating the goodness of fit of the models is inspecting the standardised deviance residuals of the fitted subpopulation model

$$r_{xtg} = \text{sign}({}_nd_{xtg} - {}_n\hat{d}_{xtg})\sqrt{\frac{\text{dev}(x,t,g)}{\hat{\phi}}}, \qquad \hat{\phi} = \frac{D({}_nd_{xtg}, {}_n\hat{d}_{xtg})}{N-\nu}$$

and, when applicable, the deviance residuals of the reference population model

$$r'_{xt} = \text{sign}(d'_{xt} - \hat{d}'_{xt})\sqrt{\frac{\text{dev}'(x,t)}{\hat{\phi}}}, \qquad \hat{\phi}' = \frac{D'(d'_{xt}, \hat{d}'_{xt})}{N'-\nu'},$$

where

$$D({}_nd_{xtg}, {}_n\hat{d}_{xtg}) = \sum_{x\in\mathcal{X}}\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\text{dev}(x,t,g) = \sum_{x\in\mathcal{X}}\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}2\omega_{xtg}\left\{{}_nd_{xtg}\log\frac{{}_nd_{xtg}}{{}_n\hat{d}_{xtg}} - ({}_nd_{xtg} - {}_n\hat{d}_{xtg})\right\},$$

$$D'(d'_{xt}, \hat{d}'_{xt}) = \sum_{x\in\mathcal{X}'}\sum_{t\in\mathcal{T}'}\text{dev}'(x,t) = \sum_{x\in\mathcal{X}'}\sum_{t\in\mathcal{T}'}2\omega'_{xt}\left\{d'_{xt}\log\frac{d'_{xt}}{\hat{d}'_{xt}} - (d'_{xt} - \hat{d}'_{xt})\right\},$$

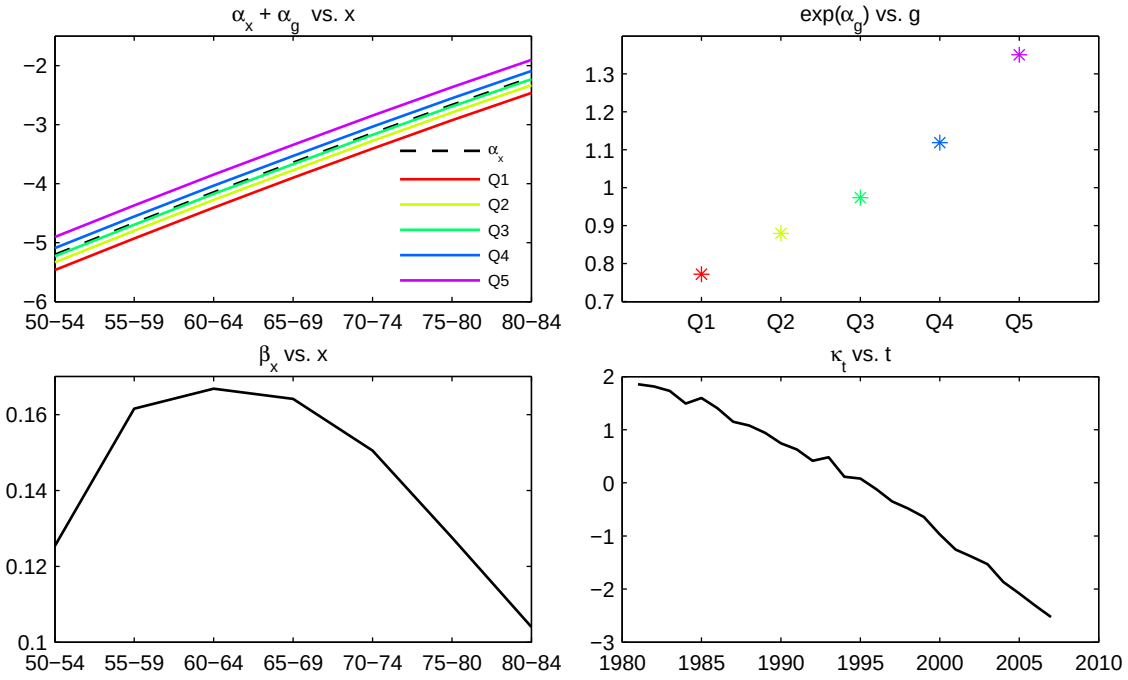


FIGURE 2.2: Parameter estimates of the stratified Lee-Carter model for the male deprivation subpopulations.

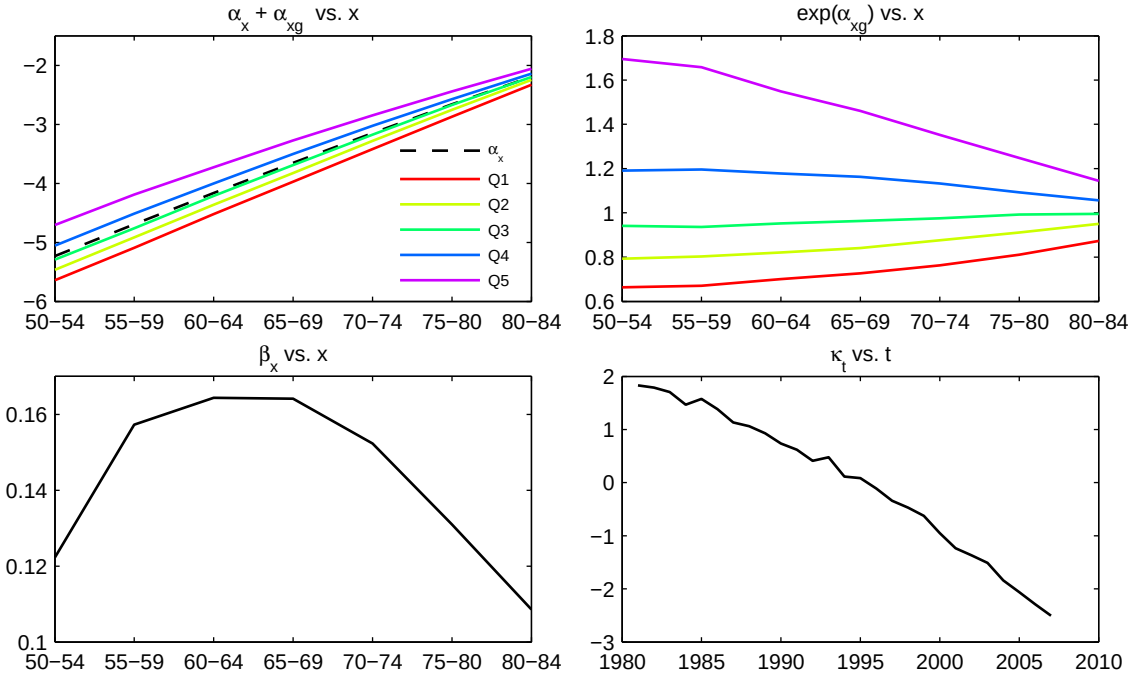


FIGURE 2.3: Parameter estimates of the common factor model for the male deprivation subpopulations.

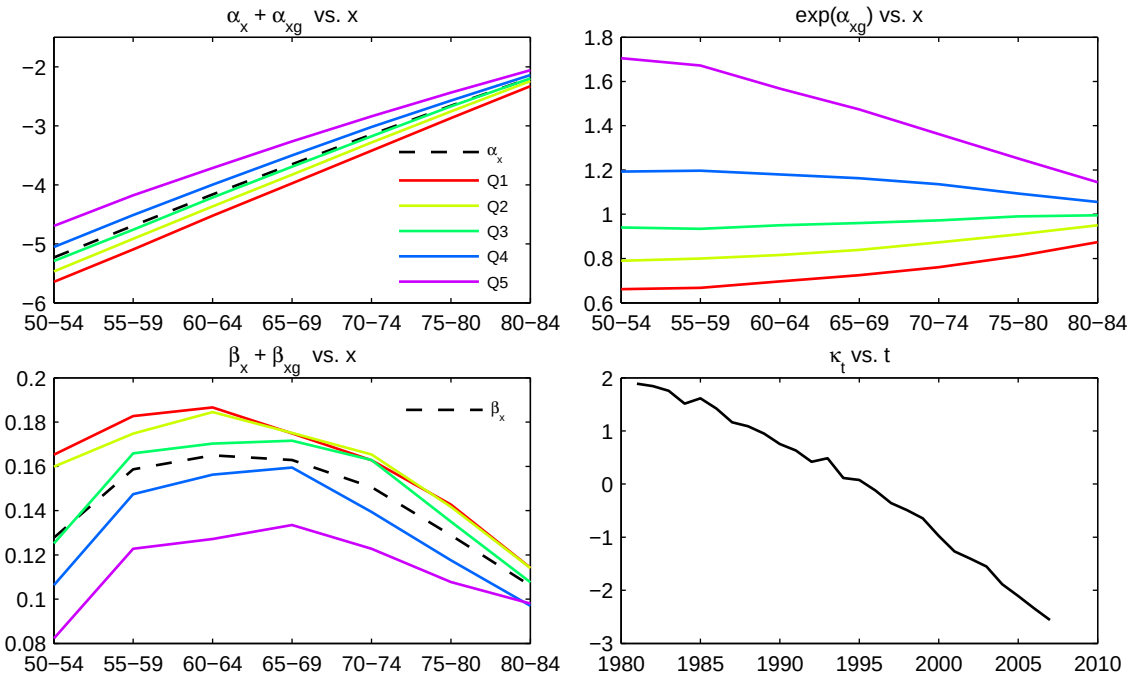


FIGURE 2.4: Parameter estimates of the joint-κ model for the male deprivation subpopulations.

$N = \sum_{xtg} \omega_{xtg}$, $N' = \sum_{xt} \omega'_{xt}$ are the number of observations in the subpopulation and reference population data, respectively, ν is the effective number of parameters in the subpopulation part of the model, and ν' is the effective number of parameters in the reference population part of the model[3]. Regular patterns in the residuals are a sign of the inability of the model to describe all of the phenomena appropriately.

Figures 2.9 plots the standardised deviance residuals of the models without a reference population, whereas Figure 2.10 plots the residuals of the relative models with a reference population. The residuals of the stratified Lee-Carter model show systematic patterns by age and calendar time, indicating that deprivation subpopulations do not satisfy the underlying model assumption that on a log scale the age-specific mortality profile of the subgroups defined by the extra variate (g) deviate parallelly from the population age-specific mortality profile. The common factor model does not exhibit any clear pattern by age and deprivation quintile, signifying that it captures satisfactorily mortality level differentials among the deprivation subpopulations. However, this model assumes that

[3]For example, the stratified Lee-Carter model (2.5) requires the estimation of k values of α_x, m values of α_g, k values of β_x and n values of κ_t, totalling $2k+m+n$ parameters, but since the parameters have to satisfy the constraints $\sum_x \beta_x = 1$, $\sum_t \kappa_t = 0$ and $\sum_g \alpha_g = 0$, the effective number of parameters is $\nu = 2k + m + n - 3$. For the relative model with age-period-cohort reference we take v' as the number of parameters in the reference population equation (2.6) less 4 to account for constraints (2.8)-(2.10), (2.14) and v as the number of parameters that appear exclusively in the subpopulation equation (2.7) less 2 for constraints (2.11) and (2.12).

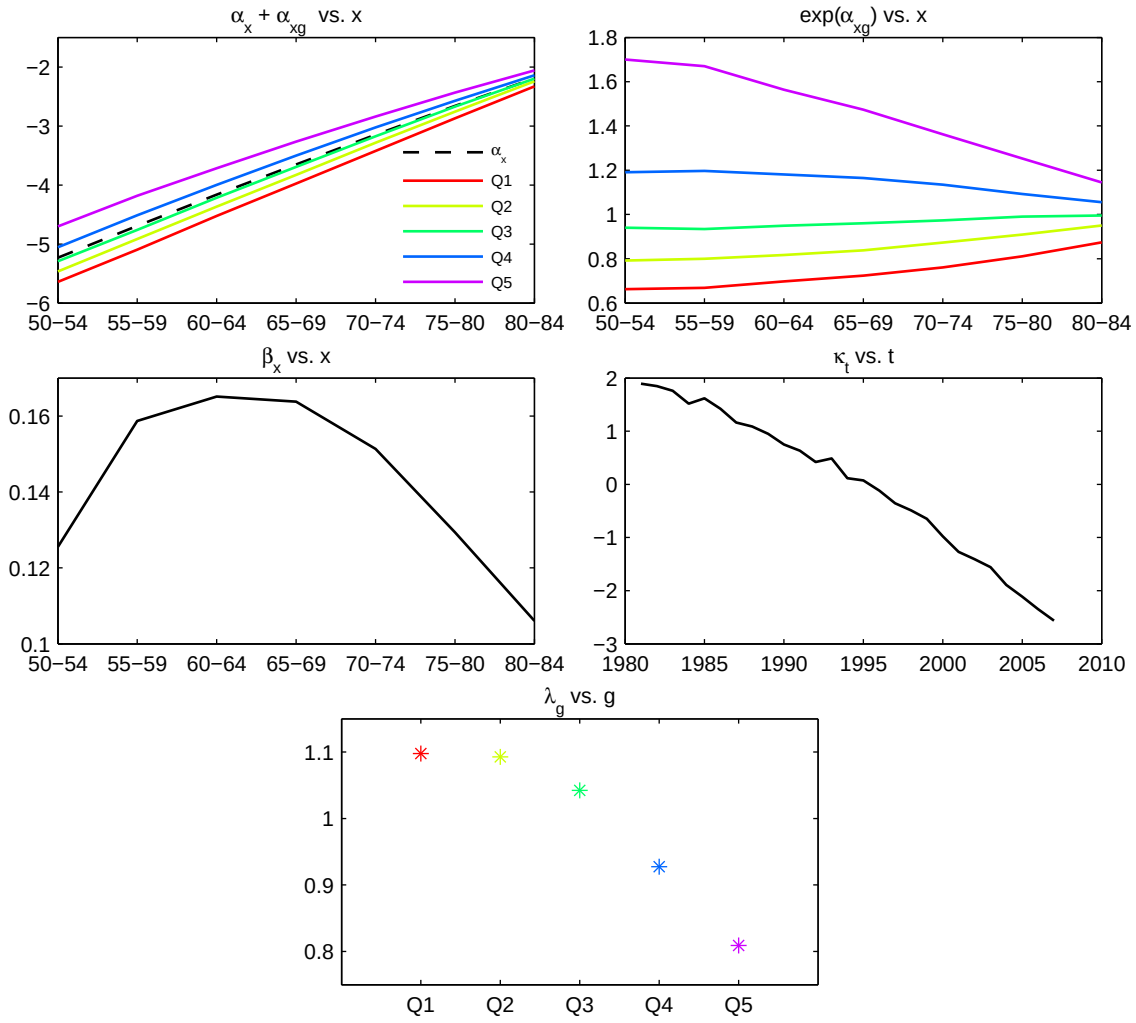


FIGURE 2.5: Parameter estimates of the three-way Lee-Carter model for the male deprivation subpopulations.

mortality level differentials are constant over time, failing to capture differences in the rate of mortality improvement as revealed by the diagonal patterns observed in the calendar year residual plots. The independent Lee-Carter model, the joint-κ model and the three-way Lee-Carter model do not exhibit any clear pattern in their residual plots, suggesting that all three models capture successfully both mortality level and mortality improvement differentials among the deprivation subgroups. In Figure 2.10 we observe a distinctive ripple effect in the year-of-birth residual plots under the Lee-Carter model for the reference population, indicating a failure to capture cohort effects. This contrasts with the absence of any major distinctive pattern in both the reference population and subpopulation residual plots under age-period-cohort modelling, indicating the success of this model in capturing the main effects of the reference population as well as in capturing mortality differentials in the deprivation subpopulations.

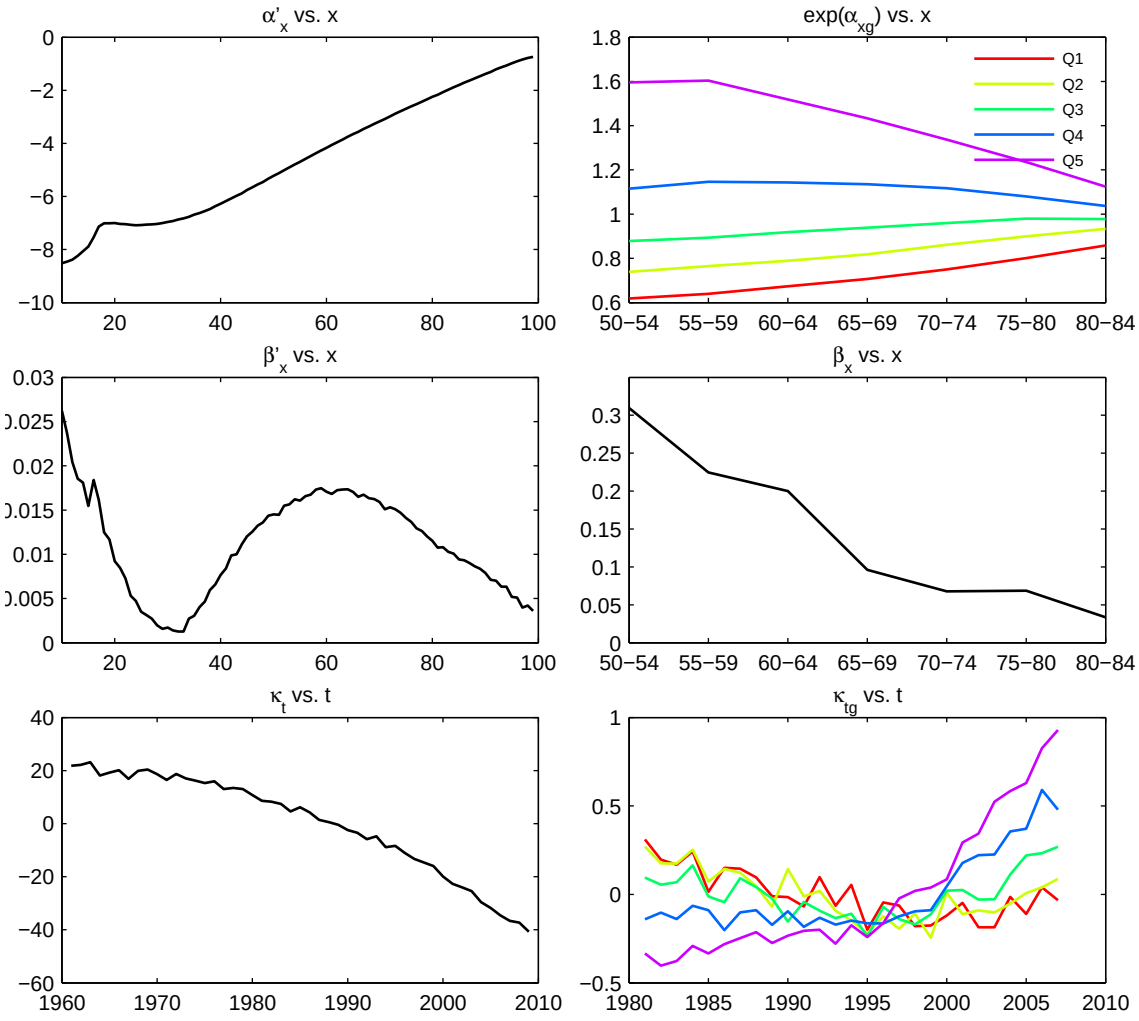


FIGURE 2.6: Parameter estimates of the relative model for the male deprivation sub-populations with Lee-Carter reference.

When assessing the goodness of fit of different models, it is natural that models with more parameters provide a better fit to the data. Therefore, we use information criteria to rule out the possibility that the better fit observed in a model is a result of overparametrisation. Information criteria modify the maximum likelihood criterion by including a penalty function so that overparametrised models are penalised. Here, we consider the Bayes Information Criterion (BIC) which has been proposed by Cairns et al. (2009) for the quantitative comparison of mortality projection models. This criterion is given by $BIC = \mathcal{L} - 0.5\nu \log N$, where $\mathcal{L}$ denotes the maximum log-likelihood of the model and ν is the effective number of parameters of the model. From this definition, it is clear that models with higher BIC values are preferable. In order to allow the comparison between models with and without a reference population, for models with a reference population we take the total model log-likelihood as being comprised of a

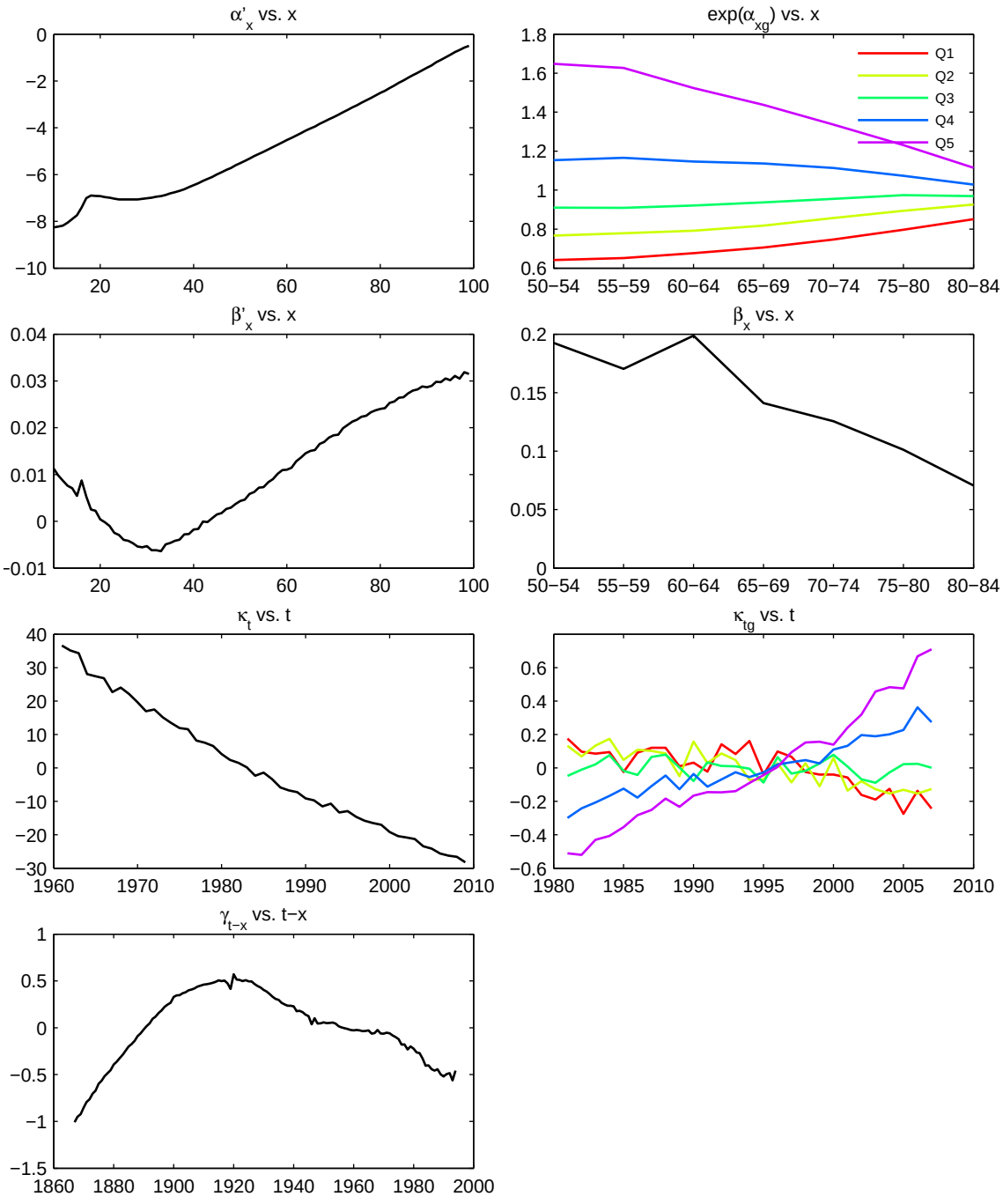


FIGURE 2.7: Parameter estimates of the relative model for the male deprivation sub-populations with age-period-cohort reference.

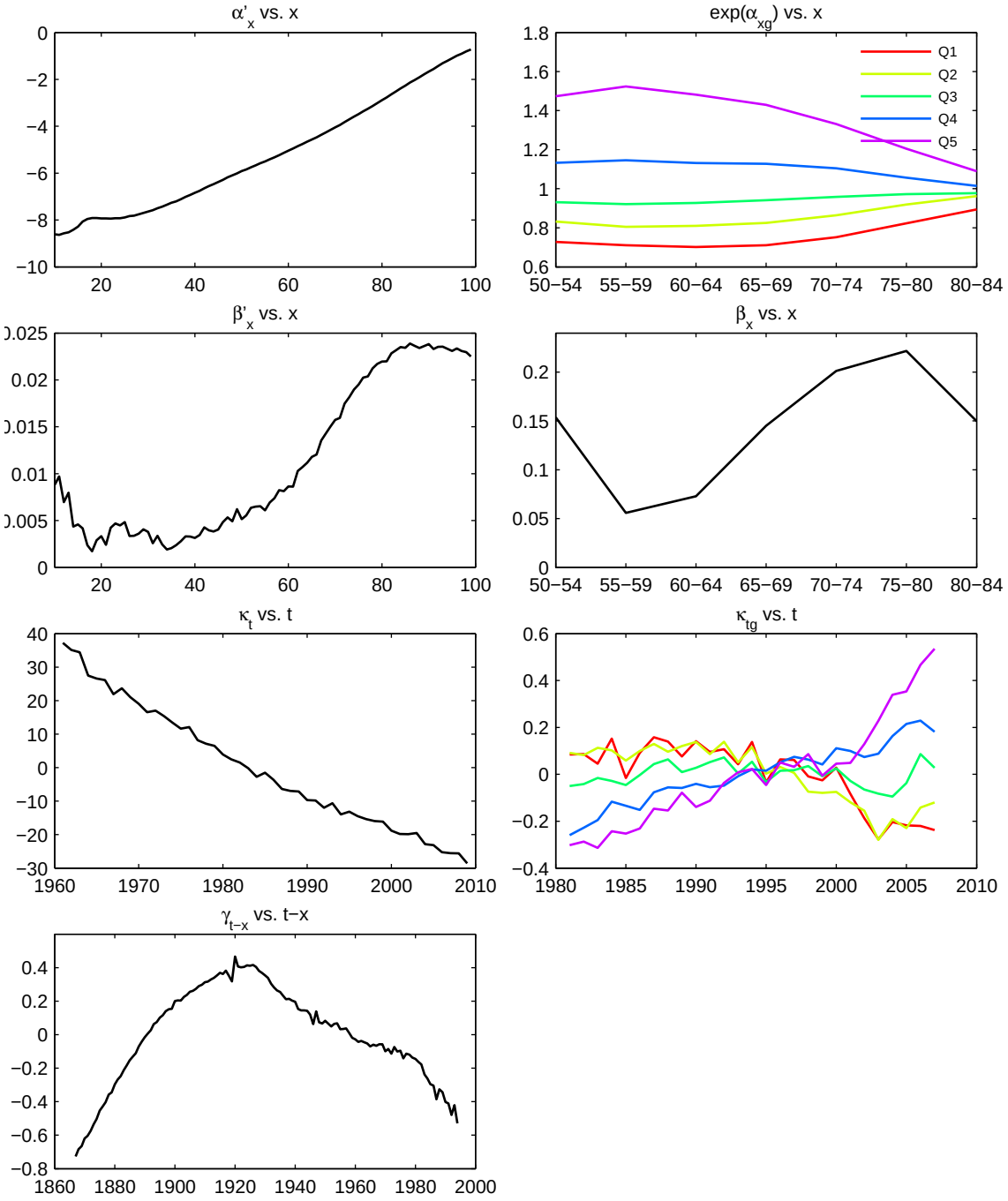
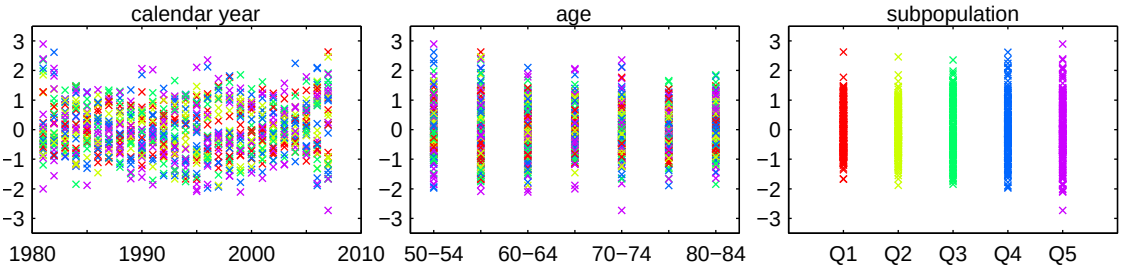
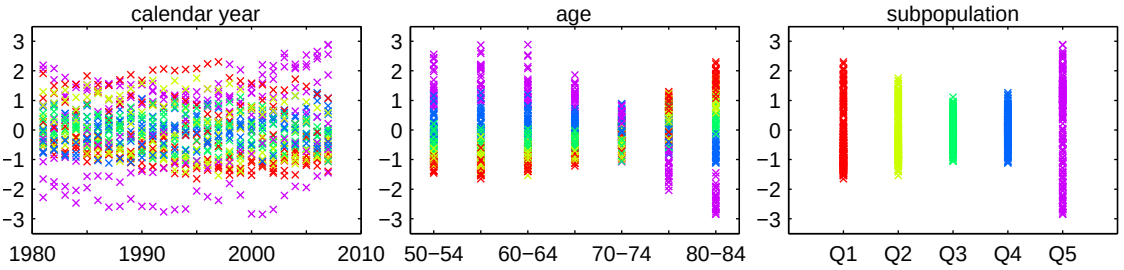


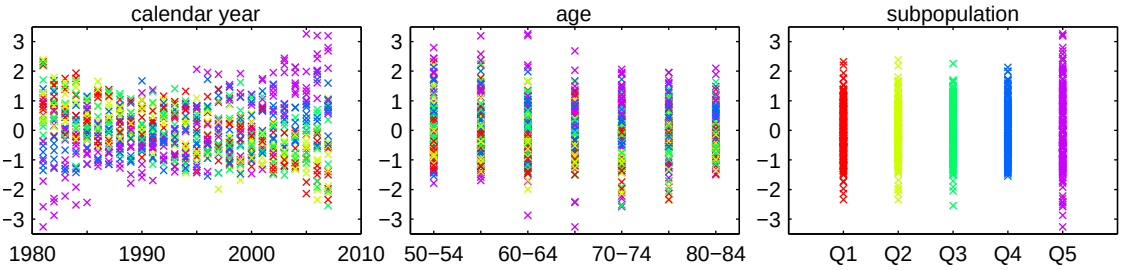
FIGURE 2.8: Parameter estimates of the relative model for the female deprivation subpopulations with age-period-cohort reference.



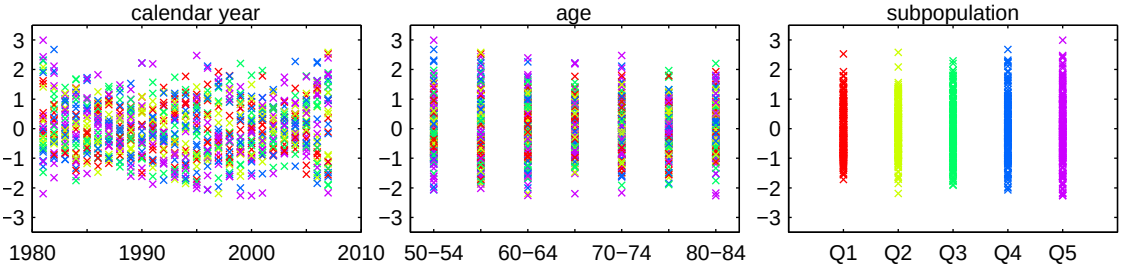
(A) Lee-Carter: $\log {}_n\mu_{xtg} = \alpha_{xg} + \beta_{xg}\kappa_{tg}$



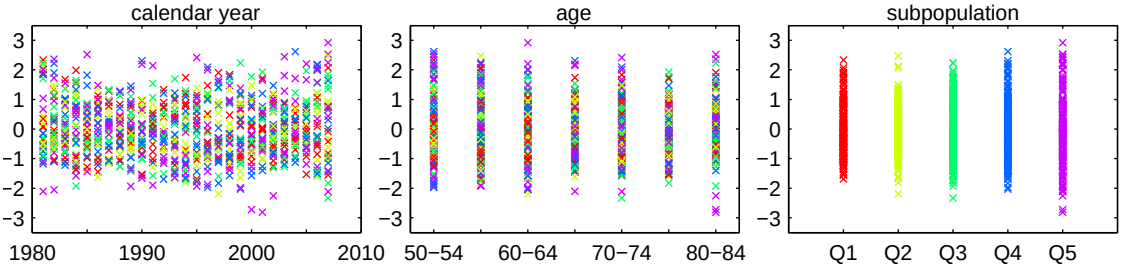
(B) Stratified Lee-Carter: $\log {}_n\mu_{xtg} = \alpha_x + \alpha_g + \beta_x\kappa_t$



(C) Common factor: $\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + \beta_x\kappa_t$

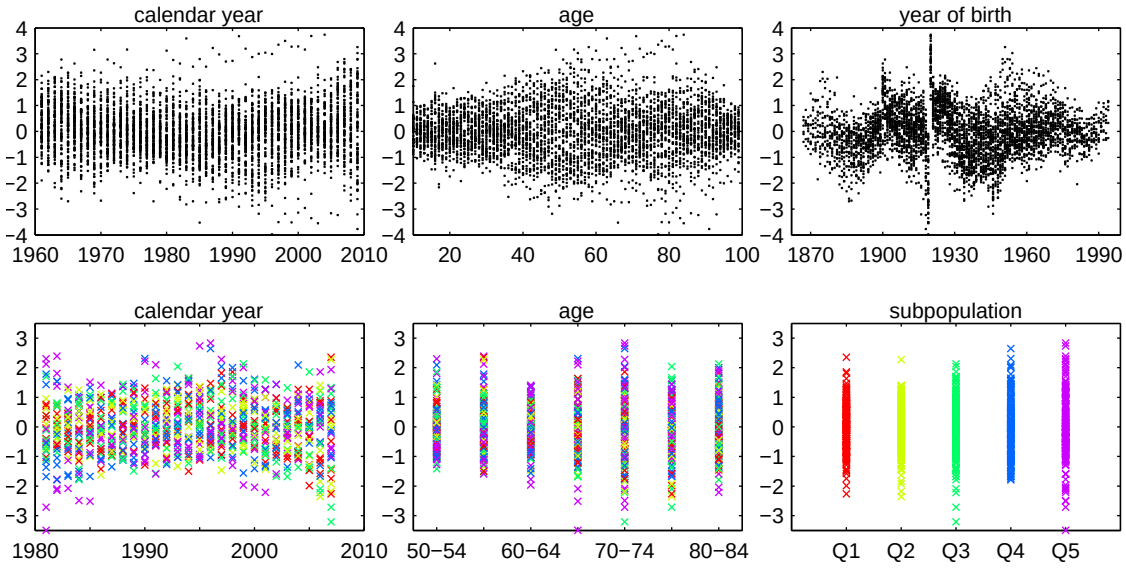


(D) Joint-κ: $\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + (\beta_x + \beta_{xg})\kappa_t$

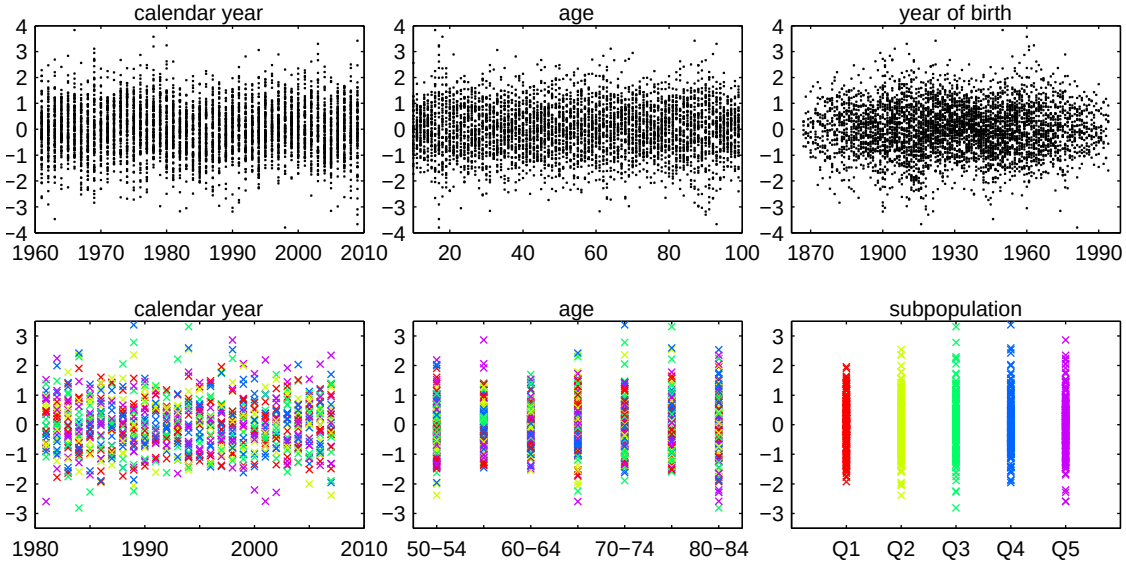


(E) Three way Lee-Carter: $\log {}_n\mu_{xtg} = \alpha_x + \alpha_{xg} + \beta_x\lambda_g\kappa_t$

FIGURE 2.9: Deviance residuals for the models without a reference population applied to the male deprivation subpopulations.



(A) Lee-Carter reference: $\log \mu'_{xt} = \alpha'_x + \beta'_x \kappa'_t$



(B) Age-period-cohort reference: $\log \mu'_{xt} = \alpha'_x + \beta'_x \kappa'_t + \gamma'_{t-x}$

FIGURE 2.10: Deviance residuals for the models with a reference population applied to the male deprivation subpopulations. The first row of each subfigure displays the residuals for the reference population and the second row of each subfigure displays the residuals for the deprivation subpopulations.

reference population part, $\mathcal{L}_{ref}$, and a subpopulation part, $\mathcal{L}_{sub}$, as in equation (2.13), and report thus the two corresponding BIC values $BIC_{ref} = \mathcal{L}_{ref} - 0.5\nu' \log N'$ and $BIC_{sub} = \mathcal{L}_{sub} - 0.5\nu \log N$.

Table 2.1 contains BIC values for each of the seven adjusted models, together with their respective rankings (in brackets). The small BIC values for the stratified Lee-Carter model and the common factor model confirm the inability of these models to capture adequately the differentials in mortality of the deprivation subpopulations. We observe

TABLE 2.1: Log-likelihood $\mathcal{L}$, effective number of parameters ν, and BIC values for the different models applied to the male deprivation subpopulations.

Model	Subpopulations			Reference population		
	ν	$\mathcal{L}$	BIC	ν'	$\mathcal{L}$	BIC
Independent LC	195	-6 448	-7 116(5)	-	-	-
Stratified LC	43	-22 770	-22 918(7)	-	-	-
Common factor	66	-7 827	-8 053(6)	-	-	-
Joint-κ	94	-6 538	-6 860(2)	-	-	-
Three way LC	71	-6 670	-6 913(4)	-	-	-
Relative (LC reference)	171	-6 310	-6 896(3)	227	-31 025	-31 997(2)
Relative (APC reference)	171	-5 630	-6 216(1)	353	-23 251	-24 731(1)

that the relative model with age-period-cohort reference, the only model capturing the cohort effect, is by far on top of the subpopulation BIC ranking. This model is followed by the joint-κ model, the relative model with Lee-Carter reference and the three-way Lee Carter, all of them with very close matching statistics. We also note that among models with a reference population, the age-period-cohort model performs much better than the Lee-Carter model, providing evidence of the existence of a cohort effect.

It is possible that a model has good in-sample-fit but produces poor ex-post forecast, that is, forecasts that deviate significantly from the realised outcomes. Therefore, in addition to the assessment of goodness of fit, it is necessary to evaluate the ex-post forecasting performance of the models using suitable backtesting methods. For that purpose, we apply an expanding horizon backtest as proposed by Dowd et al. (2010). This backtesting procedure entails the following steps:

(a) Select a metric of interest, that is, the forecasted variable which is the focus of the backtest. Since our interest is forecasting both individual subpopulation mortality rates and mortality rates differentials among the subpopulations, we base our backtest exercise on the subpopulation-specific probability of dying between ages 50 and 85, ${}_{35}q_{50,t,g}$, as well as on the ratio of these probabilities in the most and least deprived subpopulations, ${}_{35}q_{50,t,Q5}/{}_{35}q_{50,t,Q1}$, and in the fourth and second IMD quintiles of the population, ${}_{35}q_{50,t,Q4}/{}_{35}q_{50,t,Q2}$, with ${}_{35}q_{50,t,g}$ calculated on a period basis using the expression

$$ {}_{35}q_{50,t,g} = 1 - (1 - {}_5q_{50,t,g})\,(1 - {}_5q_{55,t,g}) \cdots (1 - {}_5q_{80,t,g}) $$

where

$$ {}_5q_{x,t,g} = \frac{5\,{}_5\mu_{x,t,g}}{1 + 2.5\,{}_5\mu_{x,t,g}}, \quad x = 50, 55, \ldots, 80. $$

(b) Select the historical "lookback" period used to estimate the parameters of the model and select the time horizon over which the forecasts will be made. We fit the models over the restricted period 1981-2000 and perform forecasts for the period 2001-2007.

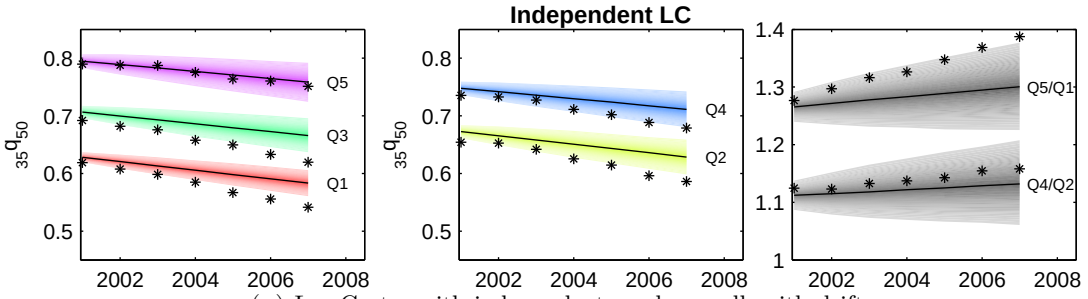
(c) Compare graphically the forecasts against realised outcomes of the metrics of interest. We plot 95% prediction intervals (fan charts) of the metrics of interest obtained by simulating 10 000 paths of the period indexes of each of the models with realised outcomes of the metrics superimposed. We note that these prediction intervals ignore any provision for parameter uncertainty.

We exclude the stratified Lee-Carter model, the common factor model, and the relative model with Lee-Carter reference from the backtesting exercise as these models show a poor performance in terms of in-sample-fit. For the independent Lee-Carter model, we include both the forecasted intervals using independent random walks with drift for each subpopulation's period index and using a multivariate random walk with drift. Figure 2.11 shows the comparison of the prediction intervals and the realised outcomes of the metrics of interest. From this figure we note the following:

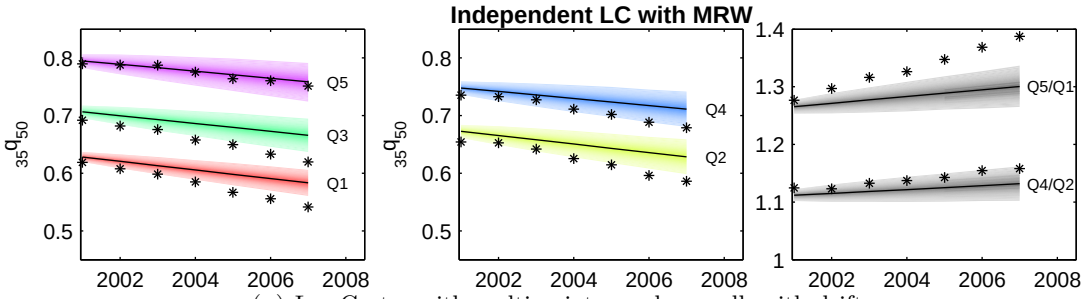
- The consistent underestimation of mortality decline by all the models with the exception of the relative model with an age-period-cohort reference, giving evidence that the possibility of capturing cohort effects results in substantial improvements in forecasting performance.

- The exact agreement between prediction intervals of subpopulation mortality rates of the two alternative forecasting methods for the independent Lee-Carter model, but the difference in the prediction intervals of the mortality differentials, with the independent random walk approach resulting in excessively wide prediction intervals.

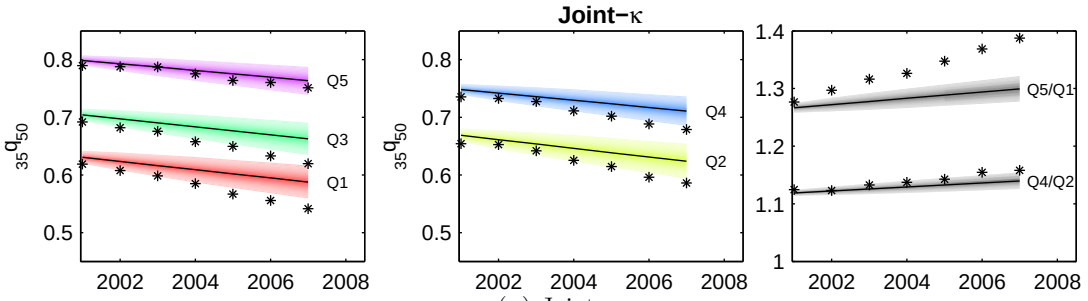
- The extremely narrow prediction intervals of mortality differentials produced by the joint-κ model and the three-way Lee-Carter model, which is a result of the perfect correlation of subpopulation mortality improvements induced by the use



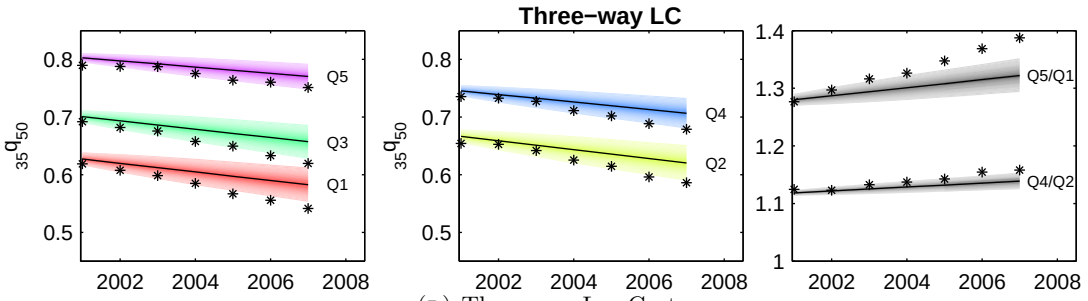
(A) Lee-Carter with independent random walk with drift



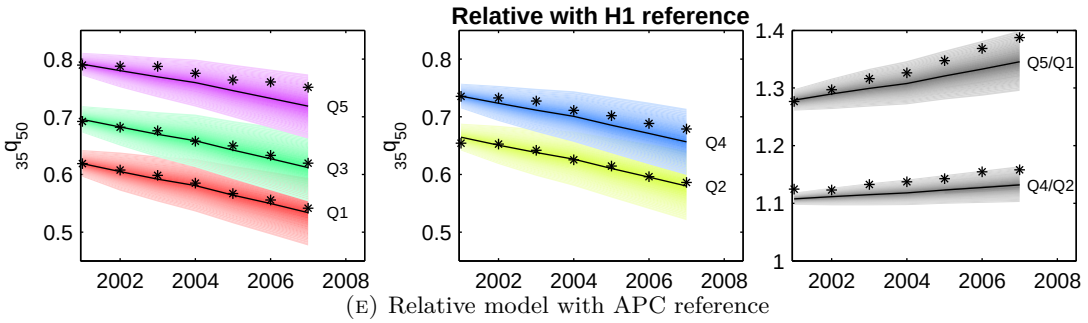
(B) Lee-Carter with multivariate random walk with drift



(C) Joint-κ



(D) Three-way Lee-Carter



(E) Relative model with APC reference

FIGURE 2.11: Predicted and realised outcomes of the backtesting metrics for the period 2001-2007. Left and central frames depict the values for ${}_{35}q_{50,t,g}$ and right frames depict the values for ${}_{35}q_{50,t,Q5}/{}_{35}q_{50,t,Q1}$ and ${}_{35}q_{50,t,Q4}/{}_{35}q_{50,t,Q2}$

of a single mortality driver for all the subpopulations. Therefore, although these two models show reasonable goodness of fit and provide a transparent framework for quantifying historical mortality differentials, they are likely to underestimate the uncertainty associated with the predictions of mortality differentials.

- The relatively poor performance of all the models in the forecasting of mortality differentials with a consistent underestimation of the widening of mortality differentials. Nevertheless, the relative model with an age-period-cohort reference shows the best performance, with central predictions that deviate the least from the realised outcomes and predictions intervals that match the observed variability.

In summary, the independent Lee-Carter model, the joint-κ model, the three-way Lee-Carter model and the relative model with age-period-cohort reference all succeed in fitting historical mortality differentials in the English male population, with the latter model outperforming the others according to the BIC criterion. In addition, the relative model with age-period-cohort reference outperforms the others in terms of ex-post evaluation of the model predictions. Equivalent analyses – not shown here – for the female population lead to similar conclusions. For these reasons, our subsequent discussions on the differentials in mortality across deprivations quintiles will be based on the relative model with age-period-cohort reference.

2.3.2 Historical mortality differentials: 1981-2007

We now use the new relative model with age-period-cohort reference to analyse the socio-economic differentials in mortality observed in the English male and female populations during the period 1981-2007. Figures 2.7 and 2.8 depict the corresponding parameter estimates for the male and female populations, respectively.

We first turn our attention to the parameter estimates of the England and Wales reference populations. The main age-effects plots (α'_x vs. x) for both sexes have the usual features of static life tables, with a more pronounced accident hump in males and lighter mortality in females. For both males and females, the period index plot (κ'_x vs. t) exhibits a steady decline in mortality with a mild curvature. For both genders, the plot of the age-modulating parameter β'_x reveals that the ages in the range 20-40 have

experienced the slowest mortality decline, with β'_x being even negative for the male experience. This result coincides with the pronounced decrease in the rate at which mortality improved in the last quarter of the past century reported by Renshaw and Haberman (2003) for both males and females in the age band 20-40. Finally, the plots of the cohort effect (γ'_{t-x} vs. $t-x$) for both sexes are similar and reveal the impact of the 1919 influenza pandemic and the rapid mortality improvement experienced by the generation born between 1925 and 1945 as reported by Willets (2004). It is worth mentioning that the cohort effects for the generations born after 1945 need to be interpreted with care as they only reflect the behaviour of mortality at young and middle ages, and it is not totally clear whether these effects will still hold at older ages.

The subpopulation parameters reveal very different subpopulation mortality rate levels as well as very different evolutions of these rates over time. In particular, the plots of $\exp(\alpha_{xg})$ show a steep increase in mortality rate levels with increasing deprivation. For instance, during the period 1981-2007 males age 50-54 in the least deprived quintile experienced 36% lower mortality than the reference population, which contrasts with the 65% higher mortality experienced by males age 50-54 in the most deprived quintile. Significant differences in the level of mortality are observed for both genders and all ages, though with varying magnitudes. Males show higher mortality level differentials than females. For instance, while for males age 60-64 the ratio between mortality in the most deprived and the least deprived quintile, $\exp(\alpha_{x,Q5})/\exp(\alpha_{x,Q1})$, is 2.25, for females this ratio reduces to 2.11. In addition, mortality level differentials narrow with age. At age 50-54 males and females in the most deprived quintile have respectively 2.57 and 2.03 higher mortality than persons of the same sex and age in the least deprived quintile. By contrast, at age 80-84 males and females in the most deprived quintile have respectively 1.31 and 1.22 higher mortality than persons of the same sex and age in the least deprived quintile. This decrease in socio-economic mortality differences with rising age is commonly reported in mortality research (Hoffmann, 2005).

The plots of κ_{tg} indicate that, in spite of the overall decrease in mortality levels, the more deprived quintiles have experienced slower mortality reductions than the less deprived ones, evidencing a widening of the mortality gap between the least and the most deprived areas of England. The sharp deceleration after 1997 in the pace of mortality improvement of the most deprived quintile of the female population indicated by $\kappa_{t,Q5}$ is particularly noticeable. For both males and females, it is also worth noticing the close alignment

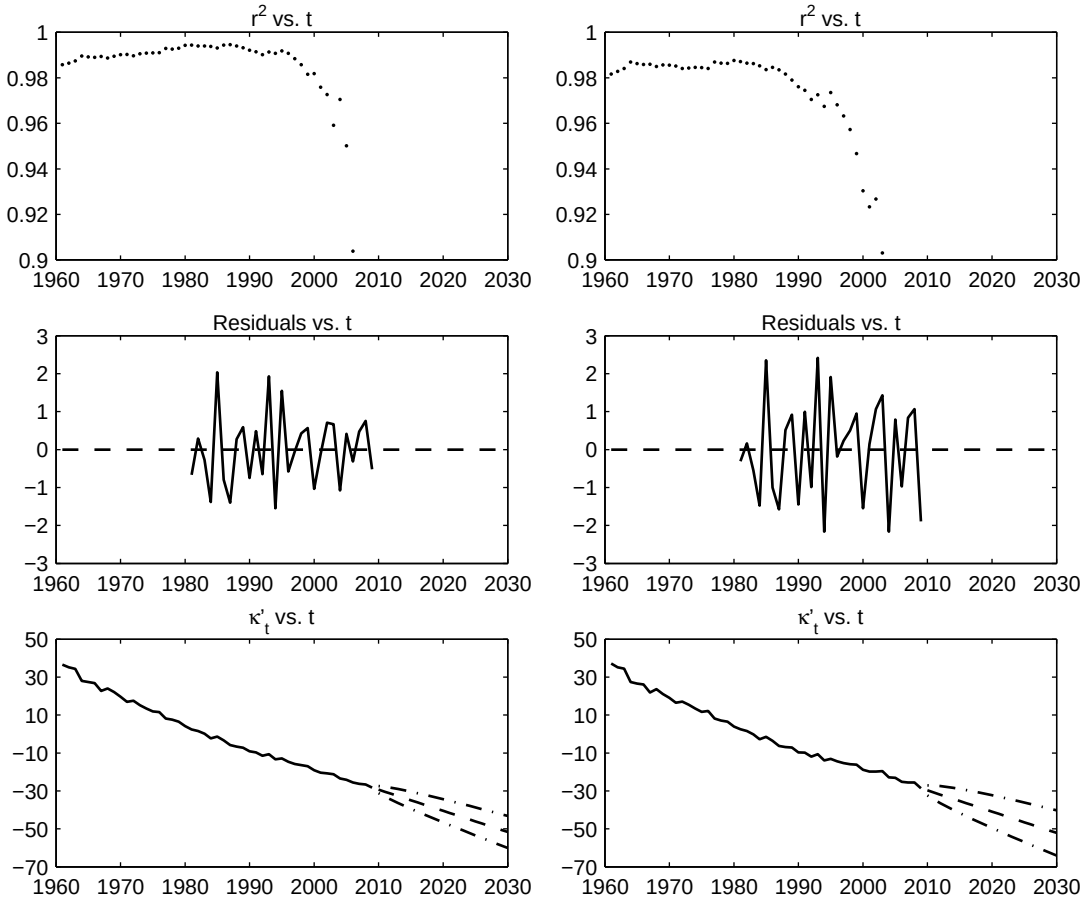
between the period indexes for the two least deprived quintiles, $\kappa_{t,Q1}$ and $\kappa_{t,Q2}$, which contrasts with the marked differences in the corresponding level parameters $\alpha_{x,Q1}$ and $\alpha_{x,Q2}$. This indicates that despite the fact that there are material level differences between these two subpopulations, improvement differences are negligible.

The plots of β_x show that in the male population the highest differentials in the rate of mortality improvement are observed at ages 50-54 and 60-64, whereas in the female population the greatest mortality improvement differentials are achieved at ages 70-74 and 75-79. We also notice that at ages 55-59 and 60-64 females from all deprivation quintiles have experienced more or less the same mortality reduction. For males, we see a decline in mortality improvement differentials with rising age similar to that observed in mortality level differentials. Nonetheless, this inverse relationship between age and mortality improvement differential does not hold for females. In fact, from age group 55-59 until age group 75-79 improvement differences increase with age.

2.3.3 Mortality differential projections

In order to project age-subpopulation-specific mortality rates and examine the possible future evolution of mortality differentials, we extrapolate the period indexes κ'_t and κ_{tg}. The presence of a mild curvature in κ'_t complicates its forecasting. Although second order ARIMA models should produce good time series fits, Haberman and Renshaw (2009) argue against the use of this approach as it tends to produce excessively wide prediction intervals. Therefore, we instead follow Haberman and Renshaw (2009) and curtail the time series at a perceived point of departure from linearity. To assist with this subjective task, we monitor the profile of the R^2 linear regression goodness-of-fit statistic, constructed backwards following the approach proposed by Denuit and Goderniaux (2005). For both males and females, the R^2 statistic attains a maximum around 1980: thus, we model κ'_t post 1980 using a random walk with drift. The corresponding residual and time series plots, together with R^2 profiles are presented in Figure 2.12.

Since κ_{tg} does not show any significant departure from linearity, we use a multivariate random walk with drift to model its dynamics. We recall from Section 2.2 that this approach permits the consideration of correlations in the mortality evolution of the different subpopulations which is a necessary feature for the adequate estimation of prediction intervals of mortality differentials.



(A) Male population (B) Female population

FIGURE 2.12: Period index time series for the England and Wales reference population (modelled post 1980). First row: R^2 goodness of fit statistics. Second row: random walk with drift time series residuals. Third row: time series with predictions and with 95% prediction intervals.

We use the forecasted values of κ'_t and κ_{tg} together with the estimated α'_x, β'_x, γ'_{t-x}, α_{xg}, and β_x to obtain forecasts of subpopulation specific mortality rates for ages 50-84 up to year 2030. We note that forecasts for this period and age range do not require the extrapolation of the cohort component.

These projected mortality rates are then employed to compute: the ratio between the mortality rates of the most and least deprived quintiles ${}_5\mu_{x,t,Q_5}/{}_5\mu_{x,t,Q_1}$, the difference between the mortality rates of the most and least deprived quintiles ${}_5\mu_{x,t,Q_5} - {}_5\mu_{x,t,Q_1}$, subpopulation specific period life expectancies

$$\mathring{e}_{xtg} = 0.5 + \sum_{h=1}^{110-x} {}_hp_{xtg}, \tag{2.15}$$

and the life expectancy gap between the most and least deprived quintiles $\mathring{e}_{x,t,Q_1} - \mathring{e}_{x,t,Q_5}$, where ${}_hp_{xtg} = p_{xtg}p_{x+1,t,g}\cdots p_{x+h-1,t,g}$ and $p_{xtg} = \exp(-\mu_{xtg})$.

The implementation of the period life expectancy formula (2.15) requires subpopulation mortality rates, μ_{xtg}, for individual ages $x, x+1, \ldots, 110$, which are not available due to the age-grouped format of our subpopulation dataset and its truncation at age 84. To tackle this problem, we set $\mu'_{110,t} = 0.7$ and extrapolate the England and Wales mortality rates along the age axis up to age 110 using the topping-out by age technique proposed by Haberman and Renshaw (2009). This technique, which is a variant of the widely used demographic method introduced by Coale and Kisker (1990), uses the quadratic differencing formula

$$\begin{aligned} u_j &= \log q'_{99+j,t+j} = a + bj + cj(j+1), \\ q'_{99+j,2009+j} &= 1 - \exp(-\mu'_{99+j,2009+j}), \quad j = -1, 0, \ldots, 11, \end{aligned} \tag{2.16}$$

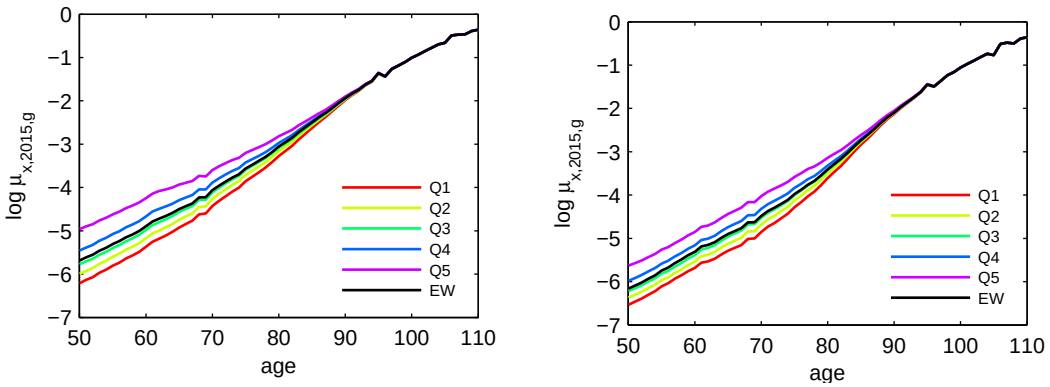
which requires $u_{-1} = \log q'_{98,t-1}$, $u_0 = \log q'_{99,t}$ and $u_{11} = \log q'_{110,t+11}$.

These extrapolated mortality rates are then used as a reference to expand the abridged subpopulation mortality rates under the assumptions that ${}_5\mu_{xtg}/{}_5\bar{\mu}'_{xt} \approx \mu_{x+2,t,g}/\mu'_{x+2,t}$ and that mortality differences vanish at age 100. More specifically, for ages $x \leq 82$ we interpolate log-linearly the mortality ratio $\mu_{x,t,g}/\mu'_{x,t}$ after setting $\mu_{x+2,t,g}/\mu'_{x+2,t} = {}_5\mu_{xtg}/{}_5\bar{\mu}'_{xt}$, $x \in \{50, 55, \ldots, 80\}$; for ages $82 < x \leq 99$, which are outside the observable age range, we modify (2.16) and use the differencing formula

$$u_j = \log \left| \log \frac{\mu_{82+j,t,g}}{\mu'_{82+j,t}} \right| = a + bj + cj(j+1), \quad j = -1, 0, \ldots, 17,$$

with $u_{-1} = \log \left| \log \mu_{81,t,g} \big/ \mu'_{81,t} \right|$, $u_0 = \log \left| \log \mu_{82,t,g} \big/ \mu'_{82,t} \right|$ and $u_{17} = \log 0.01$; and for ages $100 \leq x \leq 110$ we set $\mu_{xtg}/\mu'_{xt} = 1$. By way of illustration, Figure 2.13 depicts the set of mortality rates used to compute period life expectancies in 2015.

Figures 2.14 and 2.15 depict the projected evolution of age-deprivation-specific mortality rates. A key feature of note is that these projected mortality rates exhibit a coherent behaviour, in the sense that they are consistent with the England and Wales forecasted rates and that there are no cross-overs between the subpopulations. This is a major



(A) Male population (B) Female population

FIGURE 2.13: Projected period mortality rate profiles for year 2015.

characteristic of mortality projections produced using the relative modelling approach proposed in this chapter. The forecasted fast decline of mortality rates at ages 75-79 and 80-84, especially in the female population, is also particularly noteworthy. This feature is a reflection of the faster mortality decline experienced by those born between 1925-1945 captured by the cohort parameters of the reference population.

Figures 2.14 and 2.15 suggest a tendency to convergence in the mortality rates of the deprivation subpopulations. To investigate further this feature, we display in Figure 2.16 the relative mortality difference between the most and least deprived quintiles, ${}_5\mu_{x,t,Q_5}/{}_5\mu_{x,t,Q_1}$, as well as the corresponding absolute mortality difference, ${}_5\mu_{x,t,Q_5} - {}_5\mu_{x,t,Q_1}$. In relative terms, we observe a widening of the mortality differences between the extreme deprivation quintiles, reflecting the slowest mortality improvements experienced by the most deprived subpopulation. By contrast, in absolute terms, with the exception of some old ages in the female population, a general narrowing in mortality differences is observed. Therefore, when drawing conclusions on the widening or narrowing of mortality differences, the reference measure needs to be clearly stated.

Figure 2.17 presents a closer look at mortality differentials projections at age 65-69: central predictions are accompanied with 95% prediction intervals (fan charts) obtained by simulating 10 000 paths of the random walk model of κ'_t and of the multivariate random walk model of κ_{tg}. For both sexes a widening of mortality relative differentials is projected, although with a steeper trend and slightly wider prediction intervals for males. Whereas for men the mortality rate at age 65-69 in 2030 of the most deprived quintile is forecasted to be between 2.5 and 3.2 times the mortality rate of the least

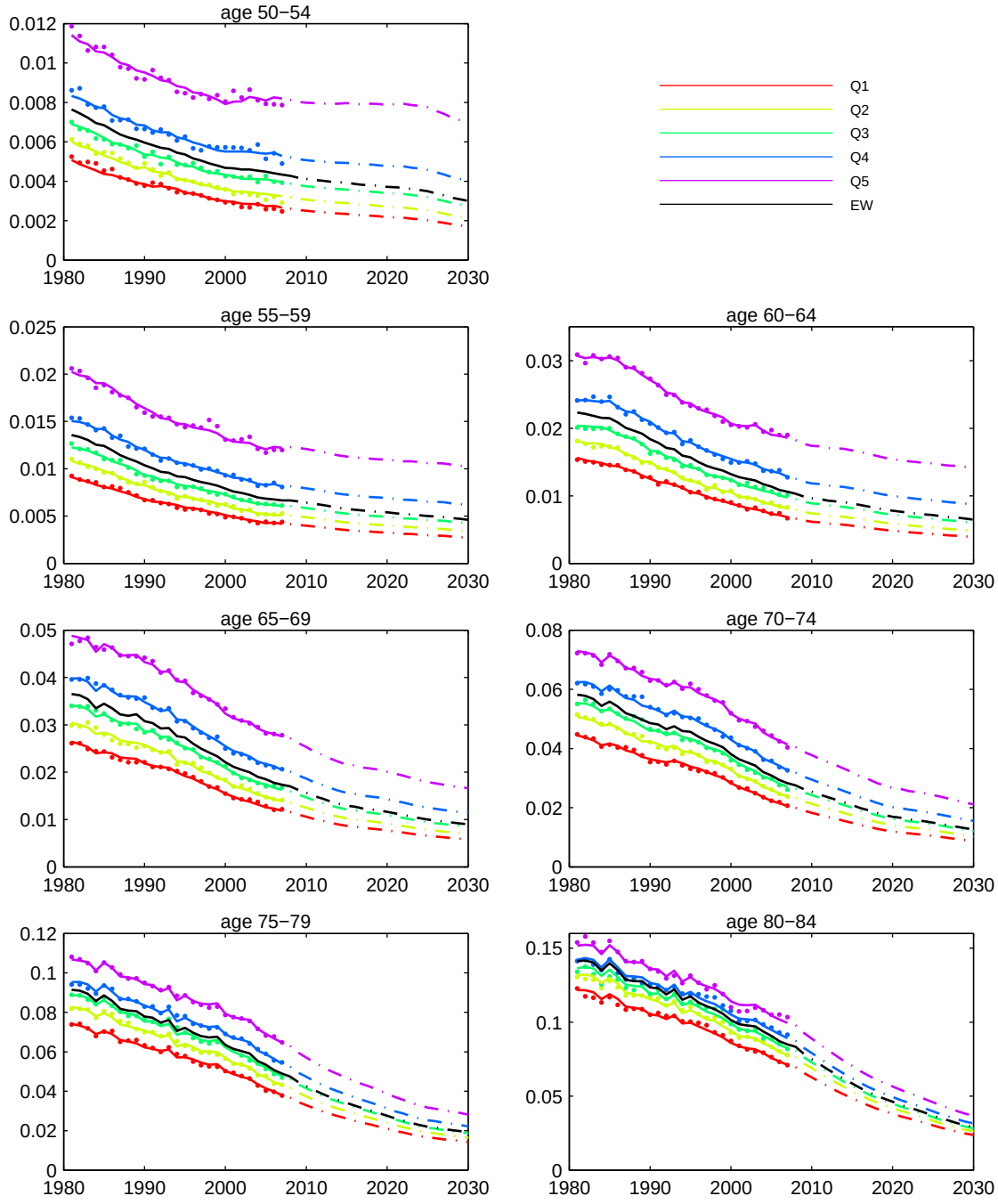


FIGURE 2.14: Time series of fitted and forecasted mortality rates ${}_5\mu_{xtg}$ for the male deprivation subpopulations. The lines labelled "EW" correspond to the fitted and projected values of ${}_5\bar{\mu}'_{xt}$ for the England and Wales reference population.

deprived quintile, for women the ratio between the mortality rate at age 65-69 of the the most deprived quintile and the least deprived quintile is projected to range between 2.3 and 2.9.

Forecasted period life expectancies at age 65 together with the life expectancy gap between the most and least deprived quintiles are displayed in Figure 2.18. We note that

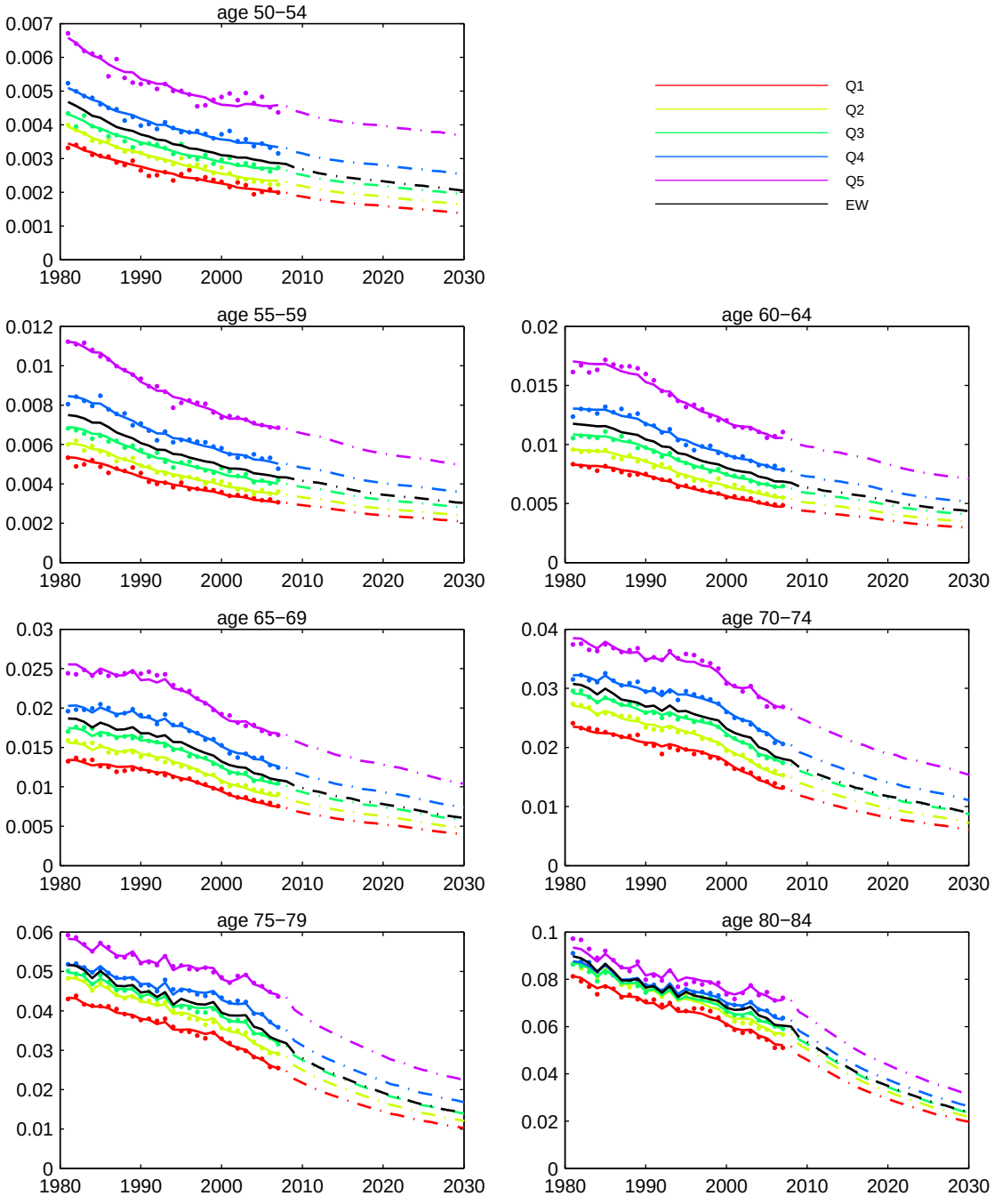
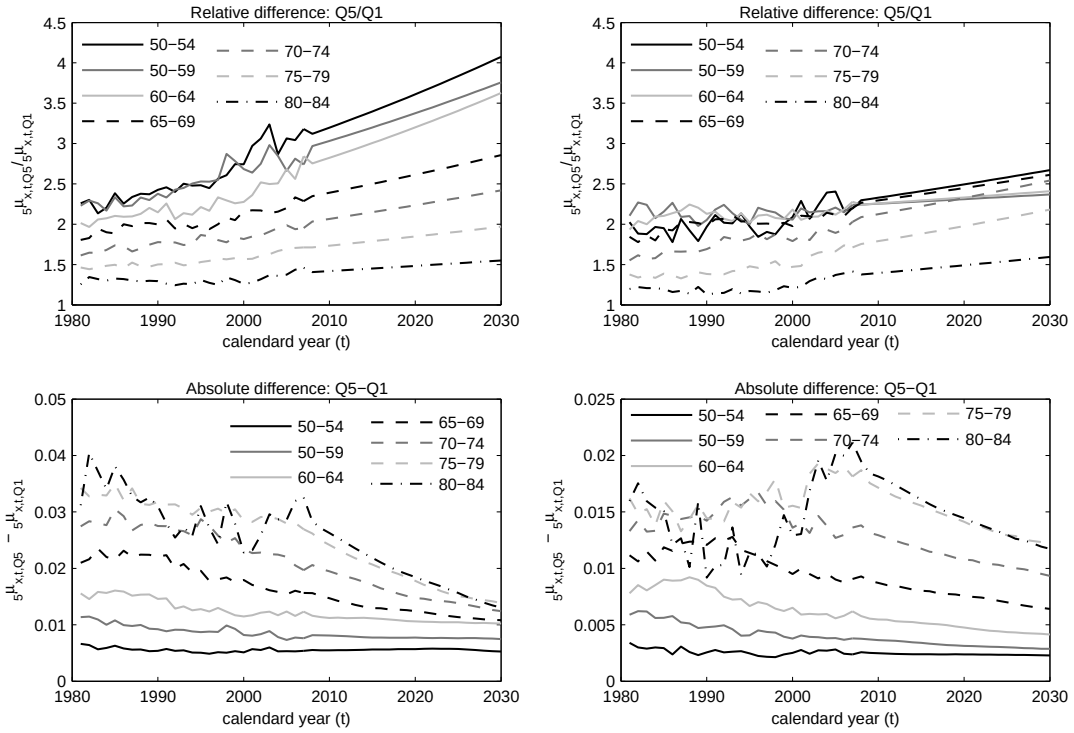


FIGURE 2.15: Time series of fitted and forecasted mortality rates ${}_5\mu_{xtg}$ for the female deprivation subpopulations.
The lines labelled "EW" correspond to the fitted and projected values of ${}_5\bar{\mu}'_{xt}$ for the England and Wales reference population.

despite the overall increase in life expectancy, the most deprived quintile has gained fewer years in life expectancy than the least deprived quintile. Specifically, the life expectancy gap between the most and the least deprived quintile increased from 2.9 years for males and 2.3 for females in 1981 to 3.8 and 3.2 years in 2007, respectively. For males, the life



(A) Male population (B) Female population

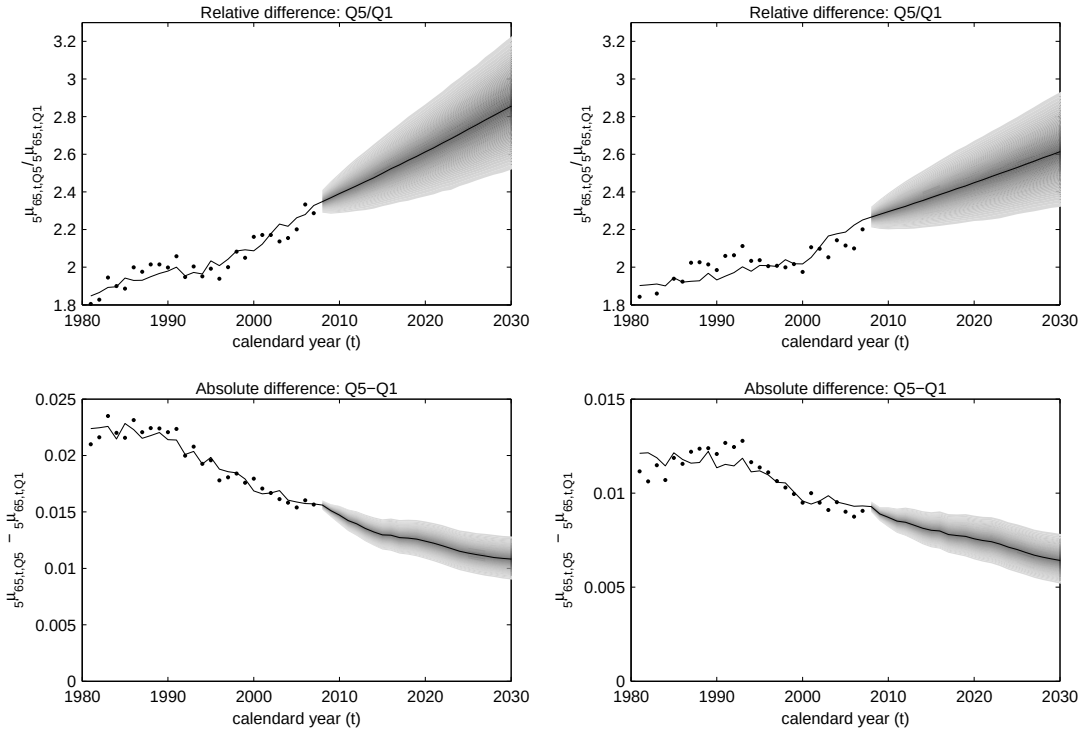
FIGURE 2.16: Time series of mortality rate differentials between the most and least deprived quintiles of the population. Values prior to 2007 are observed differentials and values post 2007 are projected differentials.

expectancy gap is projected to reach a value of 4 years in 2030 with a 95% confidence interval of 3.5-4.5 years. For females, the life expectancy gap is forecasted to remain practically unchanged at 3.1 years, but with a considerably wide prediction interval of 2.6-3.9 years in 2030.

2.3.4 Implications for life annuities

In order to assess the financial implications of socio-economic mortality differentials, in this section we investigate the variability of annuities values across deprivation sub-groups.

A typical practice in the valuation of annuities is to vary the baseline mortality assumptions according to socio-economic characteristics, but to assume the same future mortality improvements for all individuals, regardless of their socio-economic characteristics (Madrigal et al., 2011; Lu et al., 2014). Therefore, to examine the extent to which these assumptions are reasonable, we consider two alternative mortality scenarios. In



(A) Male population (B) Female population

FIGURE 2.17: Central projections and 95% prediction intervals of mortality rate differences at age 65-69.

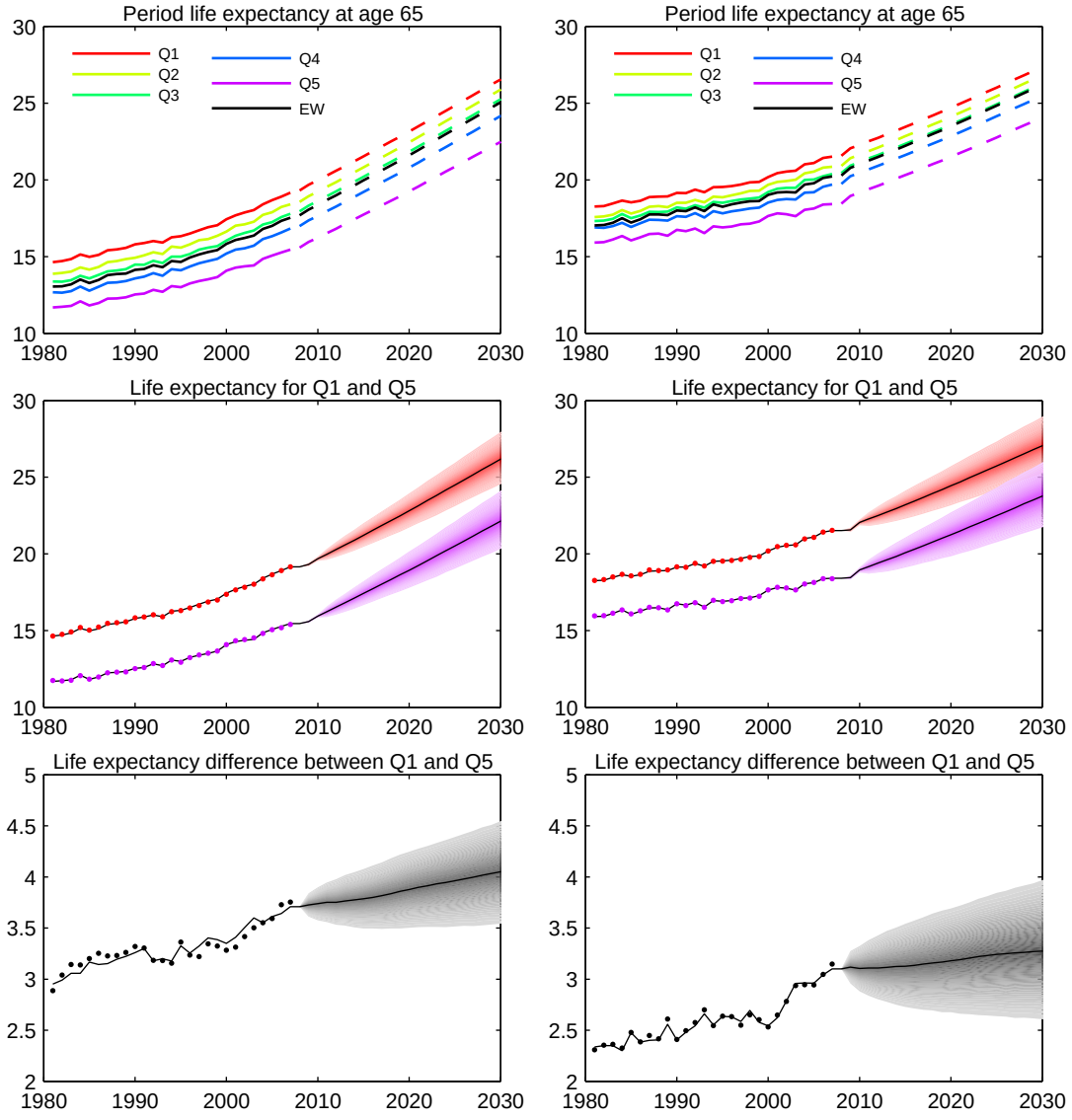
a first scenario we suppose that there are both level and improvement differences in mortality and compute projected mortality rates using the full forecasting model:

$$ {}_n\mu_{x,2007+j,g} = {}_n\bar{\mu}'_{x,2007+j} \exp\left(\alpha_{xg} + \beta_x \kappa_{2007+j,g}\right), \quad j > 0. $$

In a second scenario we assume that improvement differentials are non-existent and take mortality differences as being fixed at their 2007 level. Thus, we project mortality rates using the expression

$$ {}_n\mu_{x,2007+j,g} = {}_n\bar{\mu}'_{x,2007+j} \exp\left(\alpha_{xg} + \beta_x \kappa_{2007,g}\right), \quad j > 0, $$

which assumes that the mortality improvements of all the subpopulations follow the same behaviour of the improvements of the England and Wales reference population. Table 2.2 presents for these two mortality scenarios the percentage deviation of the annuity rate of each deprivation quintile with respect to the comparable England and Wales annuity rate. The values presented correspond to level immediate annuities for individuals age



(A) Male population (B) Female population

FIGURE 2.18: Historical and projected period life expectancies and life expectancy gap at age 65. The lines labelled "EW" correspond to period life expectancies for the England and Wales reference population. Central projections of the life expectancy in Q1 and Q5 (middle panels) and of the life expectancy gap $\mathring{e}_{x,t,Q_1} - \mathring{e}_{x,t,Q_5}$ (bottom panels) are accompanied by 95% prediction intervals.

$50, 55, \ldots, 80$ in 2007, computed under the cohort trajectory[4] and assuming an interest rate of 4%. Referring to Table 2.2 we note the following:

[4] As opposed to a period approach, a cohort approach makes full allowance for the future evolution of mortality rates. Thus, the annuity rate for an individual of subpopulation g age x in year t is computed using mortality rates $\mu_{xtg}, \mu_{x+1,t+1,g}, \mu_{x+2,t+2,g} \cdots, \mu_{110,t+110-x,g}$

TABLE 2.2: Percentage increase/decrease in annuity rates relative to the England and Wales annuity rate

Males

		Scenario 1: Level and improvement differences					Scenario 2: Only level differences				
Age	E&W	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
50	18.71	4.5%	2.7%	0.8%	-2.7%	-8.8%	4.1%	2.5%	0.8%	-2.1%	-7.2%
55	17.23	5.3%	3.1%	0.9%	-3.1%	-9.7%	5.0%	2.9%	1.0%	-2.5%	-8.2%
60	15.55	6.1%	3.5%	1.0%	-3.4%	-10.3%	5.8%	3.3%	1.0%	-2.8%	-8.9%
65	13.53	6.9%	3.9%	1.1%	-3.5%	-10.4%	6.6%	3.7%	1.1%	-3.0%	-9.2%
70	11.37	7.4%	3.9%	1.1%	-3.4%	-10.0%	7.1%	3.8%	1.1%	-3.0%	-9.1%
75	8.93	7.7%	3.9%	1.1%	-3.0%	-9.0%	7.4%	3.8%	1.1%	-2.7%	-8.4%
80	6.58	7.0%	3.4%	1.3%	-2.0%	-6.8%	6.9%	3.4%	1.3%	-1.8%	-6.4%

Females

		Scenario 1: Level and improvement differences					Scenario 2: Only level differences				
Age	E&W	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
50	19.23	2.9%	1.7%	0.4%	-1.7%	-5.4%	2.6%	1.5%	0.5%	-1.2%	-4.2%
55	17.99	3.4%	1.9%	0.4%	-1.9%	-6.1%	3.1%	1.7%	0.5%	-1.4%	-4.9%
60	16.28	4.2%	2.3%	0.6%	-2.3%	-7.1%	3.9%	2.1%	0.6%	-1.7%	-5.8%
65	14.59	4.6%	2.5%	0.5%	-2.4%	-7.6%	4.2%	2.2%	0.5%	-1.9%	-6.4%
70	12.44	5.1%	2.5%	0.5%	-2.4%	-7.8%	4.7%	2.3%	0.6%	-1.9%	-6.8%
75	9.99	5.0%	2.2%	0.5%	-2.2%	-7.3%	4.8%	2.0%	0.5%	-1.8%	-6.5%
80	7.42	4.4%	1.6%	0.5%	-1.5%	-5.7%	4.2%	1.5%	0.4%	-1.2%	-5.1%

Note: The column "E&W" presents the annuity rate for a person in England and Wales age x in 2007 at a 4% interest rate.

- As expected, annuity values decrease significantly as the level of deprivation increases. For example, under the assumption of both level and improvement differentials in mortality, the annuity rate at age 65 for males in the most deprived quintile is 10.4% less than the corresponding rate for the England and Wales population. This contrasts with a 6.9% excess for males age 65 in the least deprived quintile compared to the England and Wales population.

- Although mortality differentials decrease with rising age, the impact of these differentials on annuity values does not decrease significantly with age, and, in some cases, can even increase. For instance, at age 80 the annuity rate for males in the least deprived quintile is 7.0% higher than the corresponding annuity rate for the whole of the England and Wales population, which is an even greater difference than the 4.5% observed at age 50.

- The impact of improvement differentials on the valuation of annuities is in general of second order when compared to that of level differentials. In fact, for most ages and deprivation subgroups the difference between annuity rates computed under the two alternative mortality scenarios is less than 0.7%. This difference is only

significant for the youngest ages of the most deprived quintile, reaching a maximum of 1.6% at age 50 in the male population. These results suggest that assuming the absence of improvement differentials in mortality is in principle reasonable for the valuation of annuities. In contrast, the correct estimation of initial socio-economic differentials in mortality is critical in the pricing and reserving of annuities and pensions.

- The variability of annuity rates by socio-economic characteristics can be more significant than the variability of annuity rates by gender. For instance, at age 65 the percentage difference between the annuity rate for females and males in England and Wales is $14.59/13.53-1 = 7.8\%$, whereas the percentage difference between the annuity rate for the two extreme deprivations quintiles is $6.9\%-(-10.4\%) = 17.3\%$ and $4.6\% - (-7.6\%) = 12.2\%$, for males and females, respectively. Although in this simple comparison the significance of socio-economic circumstances is exaggerated by the fact that deprivation quintile is a multilevel factor while gender is a binary factor, this result is still interesting in view of the recent European Court of Justice ruling that insurance premiums and benefits after 21 December 2012 should be gender neutral (Court of Justice of the European Union, 2011), and, hence, socio-economic characteristics might be a candidate for substituting part of the role of gender in the pricing of annuities.

2.4 Discussion

In this chapter we have examined a number of alternatives for the modelling and forecasting of socio-economic differentials in mortality, including several existing multipopulation extensions of the Lee-Carter model and a newly proposed relative model. An application to deprivation subpopulations in England showed that in the presence of both level and improvement differentials, the new relative model exhibits the best results in terms of goodness of fit and forecasting performance. A key feature for the success of this model is the possibility of relying on the wider mortality experience of a reference population which enables the consideration of cohort effects and a more reliable estimation of the long-run mortality trend.

We note, however, that in other applications with different aims and using datasets with other characteristics, a different model might be appropriate. For instance, if the focus of the study is on the assessment of historical mortality differentials rather than on their forecasting, the joint-κ model and the three-way Lee-Carter model deserve serious consideration. Given the commonly observed widening of relative mortality differentials and the decrease in socio-economic mortality differentials with rising age, models that are not able to capture any of these features, such as the common factor model and the stratified Lee-Carter, are likely to be too rigid for most applications. In spite of that, these models may still be applied successfully in situations where there are data limitations and more parsimonious models are preferred, or in the study of mortality differentials in populations of a different nature. An example of the latter is the application of the stratified Lee-Carter in the modelling of mortality in Spanish regions discussed in Debón et al. (2011).

A noticeable advantage of the Lee-Carter based methods discussed in this chapter is that they offer a simple methodology for projecting subpopulation-specific mortality rates. However, since these methods are based on pure extrapolation of past trends, they assume that socio-economic mortality improvement differentials will remain constant in the future, ignoring the fact that policy interventions for reducing health inequalities will likely change the future mortality gradient. Nevertheless, many of the factors behind health inequalities change very slowly and, hence, the assumption that relative mortality differentials will not narrow seems a reasonable starting point for forecasting (Wanless et al., 2012).

A simplifying assumption of the relative model introduced in this chapter is that cohort effects are the same between socio-economic subpopulations. However, there is some evidence that cohort effects may vary across socio-economic subgroups. For instance, it has been reported that the cohort effect for assured lives, who are more likely to belong to higher socio-economic subgroups, is centred in a slightly earlier generation than seen in the general population of England and Wales (CMI Bureau Mortality Sub-Committee, 2002; Willets, 2004). Consequently, the development of models that allow for socio-economic variations in cohort effects deserves further investigation.

In the second part of this chapter we have applied the new proposed relative model to

analyse the extent of mortality differentials across deprivation subgroups in England during 1981 through 2007. This analysis reveals a clear association between area deprivation and mortality rates, with people living in more deprived areas having higher mortality rates than those living in less deprived areas. The mortality differentials found in this study are substantial. In fact, at some ages the mortality rates of the most deprived quintile can be more than twice the mortality rates of the least deprived quintile. In addition, our analysis indicates a widening of the relative mortality gap between more and less deprived areas of England, mainly as a result of the slower mortality improvements experienced by the lowest socio-economic subgroups.

It has been shown that socio-economic differences in mortality have a significant impact on the pricing of life annuities. Moreover, despite the fact that socio-economic differences in mortality tend to decrease with age, it was found that their financial impact is still very significant at old ages. These results become more relevant in light of the European Court of Justice ban on the use of gender as an underwriting variable. With gender removed from the list of admissible rating factors, socio-economic related rating factors gain relative importance in the modelling of longevity risk.

Finally, we recognise some data issues that might distort our results on the association between mortality and deprivation. First, our quintile groups are defined using the IMD 2007, implicitly assuming that the relative ranking by deprivation of small areas in England remained unchanged over the 1981-2007 period. Second, it is plausible that healthier people will tend to move from more deprived areas to less deprived ones and vice versa, resulting in a potential bias toward higher mortality inequalities. However, previous studies have shown that the majority of small areas in England have stayed in the same deprivation quintile over our period of study (Norman, 2009; Lu et al., 2014), and that selective migration has a negligible impact on the analysis of trends in mortality inequalities (Norman et al., 2005). A detailed discussion of these two issues with particular reference to the dataset used in this chapter is provided by Lu et al. (2014, Appendix A).

Acknowledgements

We are grateful for comments received at the Longevity 8 Conference in Waterloo, Canada in 2012; at the Demographic Analysis and Research International Conference in Chania, Greece in 2012; and at a presentation for the Longevity Science Advisory Panel, London. We thank Madhavi Bajekal from University College London for providing the deprivation-specific mortality data and for the fruitful discussions on socio-economic differentials in mortality trends. We would also like to thank Paul Norman from University of Leeds who compiled the deprivation-specific population denominators; and Ana Debón and Andrew Hunt for the comments received on an earlier version of this chapter.

Bibliography

Brouhns, N., Denuit, M., Vermunt, J., 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics 31 (3), 373–393.

Brown, J. R., 2002. Differential mortality and the value of individual account retirement annuities. In: Feldstein, M., Liebman, J. B. (Eds.), The Distributional Aspects of Social Security and Social Security Reform. University of Chicago Press.

Butt, Z., 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, Cass Business School.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1–35.

Carter, L. R., Lee, R. D., 1992. Modeling and forecasting US sex differentials in mortality. International Journal of Forecasting 8 (3), 393–411.

CMI Bureau Mortality Sub-Committee, 2002. Working Paper 1: An interim basis for adjusting the '92 Series' mortality projections for cohort effects. CMI, London.

Coale, A., Kisker, E., 1990. Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. Asian and Pacific Population Forum 4, 1–31.

Coughlan, G. D., Khalaf-Allah, M., Ye, Y., Kumar, S., Cairns, A. J., Blake, D., Dowd, K., 2011. Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. North American Actuarial Journal 15 (2), 150–176.

Court of Justice of the European Union, 2011. Taking the gender of the insured individual into account as a risk factor in insurance contracts constitutes discrimination. Press Release, No 12/11.

Debón, A., Montes, F., Martínez-Ruiz, F., 2011. Statistical methods to compare mortality for a group with non-divergent populations: an application to Spanish regions. European Actuarial Journal 1 (2), 291–308.

Delwarde, A., Denuit, M., Guillén, M., Vidiella-i Anguera, A., 2006. Application of the Poisson log-bilinear projection model to the G5 mortality experience. Belgian Actuarial Bulletin 6 (1), 54–68.

Denuit, M., Goderniaux, A.-C., 2005. Closing and projecting life tables using log-linear models. Bulletin of the Swiss Association of Actuaries (1), 29–48.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2010. Backtesting stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts. North American Actuarial Journal 14 (3), 281–298.

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

Hoffmann, R., 2005. Do socioeconomic mortality differences decrease with rising age? Demographic Research 13, 35–62.

Human Mortality Database, 2012. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL www.mortality.org

Jarner, S. F., Kryger, E. M., 2011. Modelling adult mortality in small populations: The Saint Model. ASTIN Bulletin 41 (2), 377–418.

Johnson, B., 2011. Deriving trends in life expectancy by the National Statistics Socio-economic Classification using the ONS Longitudinal Study. Health Statistics Quarterly (49), 9–52.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Li, J. S.-H., Hardy, M. R., 2011. Measuring basis risk in longevity hedges. North American Actuarial Journal 15 (2), 177–200.

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

Liebman, J., 2002. Redistribution in the current US social security system. In: Feldstein, M., Liebman, J. (Eds.), The Distributional Aspects of Social Security and Social Security Reform. University of Chicago Press.

Lu, J. L. C., Wong, W., Bajekal, M., 2014. Mortality improvement by socio-economic circumstances in England (1982 to 2006). British Actuarial Journal 19 (1), 1–35.

Madrigal, A., Matthews, F., Patel, D., Gaches, A., Baxter, S., 2011. What longevity predictors should be allowed for when valuing pension scheme liabilities. British Actuarial Journal 16 (1), 1–38.

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Exley, S., Barnes, H., Dibben, C., 2007. The English indices of deprivation 2007. Department of Communities and Local Government, London.

Norman, P., 2009. Identifying change over time in small area socio-economic deprivation. Applied Spatial Analysis and Policy 3 (2-3), 107–138.

Norman, P., Boyle, P., Rees, P., 2005. Selective migration, health and deprivation: a longitudinal analysis. Social Science and Medicine 60 (12), 2755–71.

Raleigh, V. S., Kiri, V. a., 1997. Life expectancy in England: variations and trends by gender, health authority, and level of deprivation. Journal of Epidemiology and Community Health 51 (6), 649–58.

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

Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556–570.

Richards, S., 2008. Applying survival models to pensioner mortality data. British Actuarial Journal 14 (2), 257–303.

Russolillo, M., Giordano, G., Haberman, S., 2011. Extending the Lee-Carter model: a three-way decomposition. Scandinavian Actuarial Journal (2), 96–117.

Shkolnikov, V. M., Andreev, E. M., Jasilionis, D., Leinsalu, M., Antonova, O. I., McKee, M., Oct. 2006. The changing relation between education and life expectancy in central and eastern Europe in the 1990s. Journal of Epidemiology and Community Health 60 (10), 875–81.

Tarkiainen, L., Martikainen, P., Laaksonen, M., Valkonen, T., 2012. Trends in life expectancy by income from 1988 to 2007: decomposition by age and cause of death. Journal of Epidemiology and Community Health 66 (7), 573–578.

Tuljapurkar, S., Li, N., Boe, C., 2000. A universal pattern of mortality decline in the G7 countries. Nature 405 (6788), 789–792.

Wanless, D., Pattison, J., McPherson, K., Haberman, S., Blakemore, C., Wong, W., Lu, J., Jan. 2012. Life expectancy past and future variations by socio-economic group in England and Wales. Longevity Science Advisory Panel.
URL www.longevitypanel.co.uk

Willets, R., 2004. The cohort effect: Insights and explanations. British Actuarial Journal 10 (4), 833–877.

Yang, S. S., Wang, C.-W., 2013. Pricing and securitization of multi-country longevity risk with mortality dependence. Insurance: Mathematics and Economics 52 (2), 157–169.

3

Modelling Mortality by Cause of Death and Socio-economic Stratification: Trends and Projections for England, 1981-2030

Previous versions of this chapter were presented at the following conferences:

- July 2013. 17th International Congress on Insurance: Mathematics and Economics, Copenhagen, DENMARK. "Modelling mortality by cause of death and socio-economic stratification: an analysis of mortality differentials in England".
- July 2013. R in Insurance, London, UK. "Mortality modelling in R: an analysis of mortality trends by cause of death and socio-economic circumstances in England".
- August 2013. 48th Actuarial Research Conference, Philadelphia, US. "Modelling mortality by cause of death and socio-economic stratification: an analysis of mortality differentials in England".
- June 2014. First International Congress on Actuarial Science and Quantitative Finance, Bogotá, COLOMBIA. "Modelling mortality by cause of death and socio-economic stratification: an analysis of mortality differentials in England".

– September 2014. International Mortality and Longevity Symposium, Birmingham, UK. "Modelling Mortality by Cause of Death and Socio-economic Stratification: an Analysis of Mortality Differentials in England".

There have been some plagiarism issues with the Lee-Carter model with cause of death coding adjustments introduced in Section 3.2.1 of this chapter. Specifically, some people with whom we confidentially shared some of our preliminary results published the model as if it were of their authorship in the conference proceedings:

– Sibillo, M., D'Amato, V., Passannante, V., 2014. The prediction of mortality by causes of death in Critical Illness. In: Srivastava, H.M., Bohner,M., Avramidi, I. G., Schechter, M. , Adelman, M (Eds,), Mathematical Methods in Engineering and Economics, Praga 2-4 April 2014. ISBN:9781618042309

We have an official written communication from the authors of the above manuscript confirming our priority over the developed methods. This is available upon request.

Modelling Mortality by Cause of Death and Socio-economic Stratification: Trends and Projections for England, 1981-2030

Andrés M. Villegas[a], Madhavi Bajekal[b], Steven Haberman[a]

[a] *Cass Business School, City University London, United Kingdom*

[b] *Department of Applied Health Research, University College London, United Kingdom*

Abstract

It is well known that higher socio-economic groups – whether defined by educational attainment, occupation, income or area deprivation – have lower mortality rates and longer lives than lower socio-economic groups. In many cases, affluent subpopulations also experience faster rates of improvement in mortality. Socio-economic differentials in mortality and longevity pose important challenges when designing public policies for tackling social inequalities; and for managing longevity risk in pension funds and annuity portfolios. The successful addressing of these social and financial challenges requires the best possible understanding of the drivers of socio-economic mortality differentials. A key step in achieving this understanding is to investigate how mortality trends for leading causes of death differ between socio-economic groups. Accordingly, the main purpose of this chapter is to introduce new modelling techniques that enable the modelling and projection of mortality trends by cause of death and socio-economic stratification. We first introduce an extension the Lee-Carter model to allow for the consideration of coding changes in cause-specific mortality data. We then embed this model into a multiple population setting to allow for the quantification of socio-economic differences in cause-specific mortality. Using England mortality data for socio-economic subpopulations defined using the Index of Multiple Deprivation (IMD 2007), we show that this modelling approach can be satisfactorily employed both in the assessment of the magnitude of historical mortality differentials for the main causes of death and in the projection of the possible future evolution of these differentials.

Keywords: Mortality modelling; multipopulation models; socio-economic circumstances; causes of death; Lee-Carter model

3.1 Introduction

It is well known that higher socio-economic groups – whether defined by educational attainment, occupation, income or area deprivation – have lower mortality rates and longer lives than lower socio-economic groups. For instance, in England and Wales workers in different occupational classes not only have different life expectancies, but have also experienced significantly different mortality improvements. Whilst 65 years old professional males have gained 4.6 years of life expectancy increasing from a life expectancy of 14.0 in 1972-1976 to a life expectancy of 18.6 in 2002-06, unskilled manual workers have only gained 2.9 years increasing from 11.6 in 1972-1976 to 14.5 in 2002-06 (Johnson, 2011).

These socio-economic differences in life expectancy are widely considered to be unacceptable in a fair society and policy makers need to take action to reduce the social gradient in health (The Marmot Review, 2010). Furthermore, actuaries need to recognise mortality heterogeneity as not doing so could result, for instance, in an inadequate funding of annuity and pension obligations (Meyricke and Sherris, 2013; Villegas and Haberman, 2014). The design of effective policies for tackling social inequalities in health and the successful management of the financial implications of longevity risk therefore require the best possible understanding of the drivers of socio-economic mortality differentials. A key step in achieving this understanding is to investigate how mortality trends for leading causes of death differ between socio-economic groups.

Moreover, the utility and importance of looking at the twin issues of mortality by cause of death and by socio-economic stratification has been widely recognised by the actuarial profession. For instance, in a survey on mortality by cause of death and socio-economic stratification prepared for the International Actuarial Association, Ridsdale and Gallop (2010) identify that these would be of interest to actuaries:

- aiming to understand mortality trends and its drivers with the objective of pricing the mortality elements of insurance, annuity and pensions products, or of reserving effectively for liabilities;
- involved in the underwriting and in the setting of premium rates for "substandard lives"; or

- looking to inform or test the credibility of their assumptions for mortality projections.

Richards (2009) also acknowledges the close links between socio-economic circumstances and causes of death and hence argues that to avoid misleading results *"cause-of-death statistics need to be broken down by socio-economic group or deprivation index before being used for forecasting purposes for actuarial work"*.

However, in spite of its many benefits and uses, the analysis of cause-specific mortality trends faces several theoretical and practical challenges; much of them summarised in GAD (2001), Continuous Mortality Investigation (2004) and Richards (2009). The main theoretical problem when considering mortality by cause of death is how to capture the correlations between causes: causes are not always independent and may be linked in poorly understood manners. For instance, the same risk factors can affect several causes as is the case with smoking and numerous forms of cancer and heart diseases. In addition, even if two causes of death do not share risk factors, reductions in the relative importance of one cause can lead to further mortality improvements on the other one as medical research efforts shift from one cause to the other.

In addition, the validity of the analysis and projection of cause-specific mortality trends can be affected by changes in the diagnosis and in the classification of causes of death. The World Health Organisation (WHO) has reviewed the International Classification of Diseases (ICD) regularly in 20^{th} century to reflect new diseases, changes in medical terminology and the development of medical knowledge (Moriyama et al., 2011). Further, national statistical offices often make adjustments when adopting locally the ICD coding system (see e.g. Rooney and Devis (1996) for the case of England and Wales). These data production changes can cause substantial discontinuities in the mortality trends of some causes of death. This is for instance the case of respiratory diseases in England and Wales which, as illustrated in Figure 3.1, have seen major trend-breaks due to the numerous changes in the cause of death classification system.

As summarised by Richards (2009) *"when doing projections by cause of death we must not only take great care with socio-economic differentials, and also worry about projecting correlated time series, but we must also take note of uncertainty surrounding the classification of cause of death itself"*.

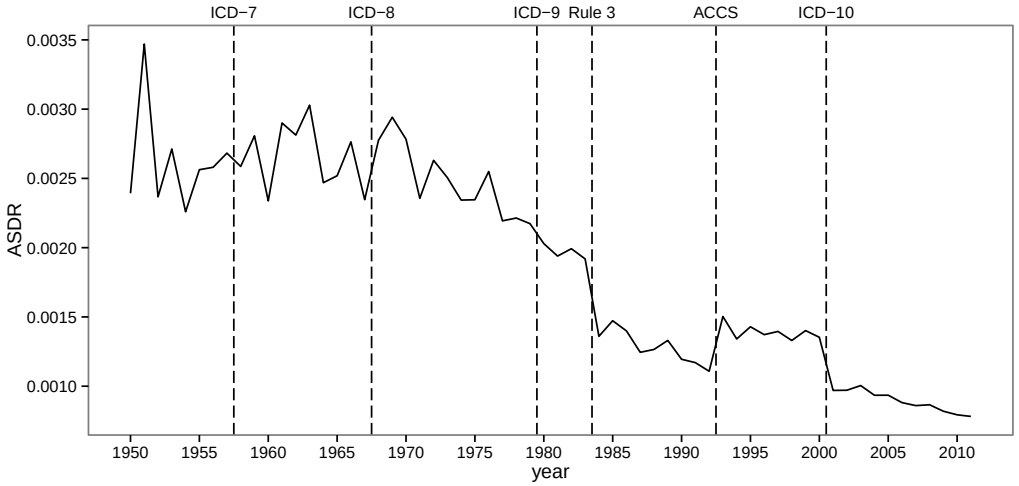


FIGURE 3.1: Age-standardised mortality rate for respiratory diseases for men aged 25-84 in England and Wales for the period 1950-2011.

Accordingly, the main purpose of this chapter is to introduce new modelling techniques that enable the modelling and projection of mortality trends by cause of death and socio-economic stratification, paying particular attention to the possible problems induced by changes in the cause of death coding system. A second purpose is to provide an analysis of the socio-economic differentials for the leading causes of death in England. This study is complementary to Chapter 2 where the modelling of socio-economic differences in all-cause mortality was addressed and which provided an investigation of all-cause mortality differentials in England across deprivation subgroups.

The chapter is organised as follows. In Section 3.2 we describe the proposed modelling approach for the quantification and projection of socio-economic mortality differentials by cause of death. Specifically, in Section 3.2.1 we introduce an extension of the widely-use model of Lee and Carter (1992) to control for coding changes in cause-specific mortality data. Then, in Section 3.2.2, we embed this model in to a multiple population setting to allow for the consideration of socio-economic differences in mortality. In Section 3.3, we apply the proposed modelling approach in the examination of the relationship between deprivation and mortality for the leading causes of death in England. Finally, in Section 3.4, we discuss our main findings and discuss possible limitations and extensions of our work.

3.2 Modelling and forecasting mortality by cause of death and socio-economic stratification

In this section we describe a modelling approach to quantify cause-specific mortality differentials in population subgroups induced by an additional covariate. Although our main interest is on population subgroups defined by socio-economic factors (e.g. occupation, level of education, social class), these subgroups could also be associated with geographical factors (e.g. regions within a country) or other factors such as marital status, gender, ethnicity and so on. The proposed approach is based on a newly introduced extension of the Lee-Carter model (Lee and Carter, 1992) to deal with data production changes in cause of death data, combined with the multipopulation Three-way Lee-Carter model proposed by Russolillo et al. (2011).

3.2.1 Lee-Carter model with cause of death coding adjustments

Assume that the death counts from a given cause of death are independent Poisson responses

$$D_{xt} \sim Poisson(e_{xt}\mu_{xt})$$

so that the force of mortality from the given cause at exact age x in time t, μ_{xt}, is given by $\mu_{xt} = \mathbb{E}(D_{xt})/e_{xt}$, with e_{xt} denoting the population at age x and time t. Also assume that mortality data are available for consecutive years $t_0, t_1, \ldots, t_n$ and let $\mathcal{S} = \{s_1, s_2, \ldots, s_h\}$ be the times at which cause of the coding changes occur with the convention that $s_0 = t_0$. In order to account for data production changes we extend the widely-used model of Lee and Carter (1992) to assume that the force of mortality from certain cause is modelled by:

$$\log \mu_{xt} = \alpha_x + \beta_x \kappa_t + \sum_{i=1}^{h} \delta_x^{(i)} f^{(i)}(t) \tag{3.1}$$

where, as in the standard Lee-Carter model, α_x captures the general age-specific mortality pattern, κ_t represents the overall level of mortality at time t and β_x measures the age-specific responses to changes in the general level κ_t. In addition, for $i = 1, \ldots, h$, $f^{(i)}(t) = \mathcal{I}_{s_{i-1} \leq t < s_i}$ denotes the indicator function taking value 1 if $t \in [s_{i-1}, s_i)$ and $\delta_x^{(i)}$

are new parameters which measure the magnitude of coding changes at age x due to the change in coding system which occurred at time s_i. This model controls for data production changes by assuming that there are different age-patterns $\alpha_x + \delta_x^{(i)}$ for each subperiod $[s_{i-1}, s_i)$, $i = 1, \ldots, h$, where a particular cause of death coding system is in force.

In a similar manner to the standard Lee-Carter model, the model defined by Equation (3.1) is only identifiable up to a set of transformations. Specifically, given constants b_1, $b_2 \neq 0$ and a_i, $i = 1, \ldots, h$, we can transform the parameters in Equation (3.1) in the following ways

$$\{\tilde{\alpha}_x, \tilde{\kappa}_t\} = \{\alpha_x + b_1\beta_x, \kappa_t - b_1\} \tag{3.2}$$

$$\left\{\tilde{\beta}_x, \tilde{\kappa}_t\right\} = \left\{\frac{1}{b_2}\beta_x, b_2\kappa_t\right\} \tag{3.3}$$

$$\left\{\tilde{\delta}_x^{(i)}, \tilde{\kappa}_t\right\} = \left\{\delta_x^{(i)} + a_i\beta_x, \kappa_t - a_i f^{(i)}(t)\right\}, \quad i = 1, \ldots, h, \tag{3.4}$$

leaving the fitted mortality rates in Equation (3.1) unchanged. Transformations (3.2) and (3.3) correspond to the transformations that can be applied to the standard Lee-Carter model, whilst the new family of transformations defined by (3.4) is induced by the new parameters $\delta_x^{(i)}$. This latter set of transformations mean that we are free to shift mortality levels from $\delta_x^{(i)}$ to κ_t.

In order to ensure the identifiability of the model, some constraints need to be imposed. To deal with transformations (3.2) and (3.3) we impose the constraints

$$\begin{aligned} \sum_x \beta_x &= 1 \\ \kappa_{t_n} &= 0 \end{aligned} \tag{3.5}$$

Choosing constraint (3.5) as opposed to the more common $\sum_t \kappa_t = 0$ gives the interpretation that α_x represents the fitted mortality rates in the last year of observation, t_n. This, coupled with the chosen definition of the set of coding changes times, $\mathcal{S}$, ensures that the coding system in force in the period $[s_h, t_n]$ (i.e. the most recent one) is the reference for the fitted and projected mortality rates.

It is desirable that the underlying coding-adjusted "jump-free" mortality trend from the cause of death of interest is captured by κ_t and that parameters $\delta_x^{(i)}$ capture the

discontinuities in mortality induced by the changes in the cause of death coding system. This goal can be accomplished by carefully choosing constants a_i, $i = 1, \ldots, h$, in the family of transformations defined by (3.4). To do so, we use a similar approach to the one introduced by Rey et al. (2011) for the removal of discontinuities in age-specific mortality rates by cause of death. Specifically, we set the constants a_i, $i = 1, \ldots, h$, by fitting the model

$$\kappa_t = g(t) + \sum_{i=1}^{h} a_i f^{(i)}(t) + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma) \qquad \text{i.i.d.}, \tag{3.6}$$

where $g(t)$ is a continuous function and ϵ_t is a Gaussian error term. Model (3.6) decomposes the time trend κ_t into:

i. a smooth function $g(t)$ representing the underlying "jump-free" mortality trend;

ii. the jumps in mortality $\sum_{i=1}^{h} a_i f^{(i)}(t)$ induced by data production changes; and

iii. the noise ϵ_t around the "jump-free" trend.

In order to estimate the smooth function $g(t)$ we follow Rey et al. (2011) and use a thin plate penalised regression spline with the smoothness parameter derived automatically using generalised cross-validation. As noted by Rey et al. (2011), this can be easily accomplished using the **MGCV R** package (Wood, 2015).

Finally, given constants a_i, $i = 1, \ldots, h$, estimated from model (3.6) and using the family of transformations defined by Equation (3.4), we can recover the "jump-free" trend using the expression

$$\tilde{\kappa}_t = g(t) + \epsilon_t = \kappa_t - \sum_{i=1}^{h} a_i f^{(i)}(t), \tag{3.7}$$

with revised data production adjustments

$$\tilde{\delta}_x^{(i)} = \delta_x^{(i)} + a_i \beta_x \quad i = 1, \ldots, h.$$

The procedure defined by Equations (3.6) and (3.7) is schematised in Figure 3.2.

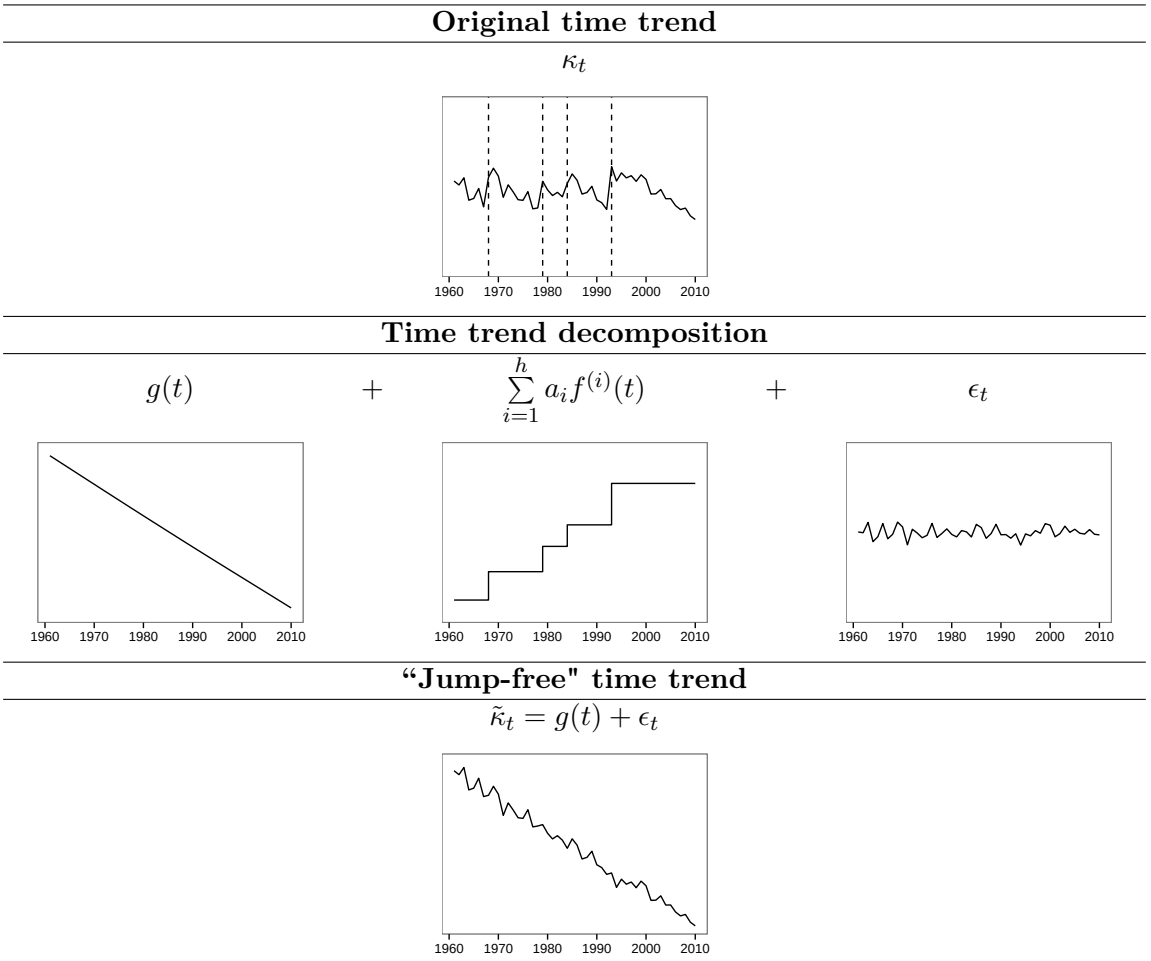


FIGURE 3.2: Schematic representation of the decomposition of the mortality trend κ_t,

3.2.2 Modelling differentials: The Three-way Lee-Carter model with coding adjustments

In order to allow for the quantification of socio-economic differences in cause-specific mortality, we now embed under a multipopulation setting the Lee-Carter model with coding adjustments introduced in the previous section. To this end, we use the Three-way Lee-Carter model proposed by Russolillo et al. (2011) to deal with mortality data disaggregated according to a third criterion besides age and time. Specifically, we assume that the central death rate for a given cause of death for age x at time t in subpopulation g, denoted μ_{xtg}, is given by

$$\log \mu_{xtg} = \alpha_x + \alpha_{xg} + \beta_x \lambda_g \kappa_t + \sum_{i=1}^{h} \delta_x^{(i)} f^{(i)}(t), \tag{3.8}$$

where α_x captures the general age-specific mortality pattern, α_{xg} captures age-subpopulation-specific deviations from the general age pattern α_x, κ_t represent the overall level of mortality at time, β_x measures the age-specific responses to changes in the general level κ_t, λ_g captures trend mortality differences across subpopulations, and $\delta_x^{(i)}$ and $f^{(i)}(t)$ have the same interpretation as in model (3.1). We note that in model (3.8) it is assumed that cause of death coding changes have a uniform impact across subpopulations. Section 3.2.4 elaborates further on the interpretation of the parameters of the model.

3.2.3 Parameter estimation

In order to estimate the parameters of the Three-way Lee-Carter model with coding adjustments, we resort to a two-stage estimation strategy using data from a reference population. More specifically, in the first stage the subpopulation independent parameters α_x, β_x, κ_t and $\delta_x^{(i)}$ are estimated using the reference population data. Then, in a second stage and conditional on the estimated values of α_x, β_x, κ_t and $\delta_x^{(i)}$, the remaining subpopulation dependent parameters α_{xg} and λ_g are estimated. Thus, our estimation procedure entails the following two steps:

i. Fit the Lee-Carter model with code adjustments defined by Equation (3.1) to the reference population to obtain parameter estimates $\hat{\alpha}_x$, $\hat{\beta}_x$, $\hat{\kappa}_t$ and $\hat{\delta}_x^{(i)}$.

ii. Given parameter estimates $\hat{\alpha}_x$, $\hat{\beta}_x$, $\hat{\kappa}_t$ and $\hat{\delta}_x^{(i)}$, obtain parameters estimates $\hat{\alpha}_{xg}$ and $\hat{\lambda}_g$ by fitting the generalised linear model

$$D_{xtg} \sim Poisson(e_{xtg}\mu_{xtg})$$
$$\log \mu_{xtg} = \hat{\alpha}_x + \alpha_{xg} + \hat{\beta}_x \lambda_g \hat{\kappa}_t + \sum_{i=1}^{h} \hat{\delta}_x^{(i)} f^{(i)}(t),$$

where, D_{xtg} and e_{xtg} are, respectively, the number of deaths for the cause of death of interest and the population at age x and time t for subpopulation g.

Such a two-stage estimation strategy has several advantages when the subpopulations represent different socio-economic strata of a national population. These advantages include:

- National mortality data are normally available for a longer period than mortality data disaggregated by socio-economic stratification, permitting a more precise estimation of the long-run mortality trend of the different causes of death.
- Modelling the subpopulations alongside the national population will ensure the consistency of the subpopulation-specific mortality forecasts with the national mortality forecasts.

3.2.4 Mortality level differential and mortality trend differentials

In general, we can distinguish between two type of mortality differentials among subgroups: mortality level differentials and mortality trend differentials. The former refers to differentials in the average level of mortality, whereas the latter refers to differentials in the pace of mortality change. In the modelling framework defined by Equation (3.8) parameters α_x and α_{xg} relate to mortality level differentials, whilst parameters β_x, λ_g and κ_t determine mortality trend differentials.

The term $\exp(\alpha_x)$ measures the general level of mortality in the reference population and the term $\exp(\alpha_{xg})$ quantifies the percentage deviations of each subgroup from this reference pattern of mortality. That is, if $\exp(\alpha_{xg}) > 1$ then at age x subgroup g has a higher mortality than the reference population and if $\exp(\alpha_{xg}) < 1$ then at age x subgroup g has a lower mortality than the reference population.

To understand how our modelling approach assesses trend differentials in mortality we can compare for each age x the rate of change over t of the log-mortality rates in the reference population,

$$\frac{d \log \mu_{xt}}{dt} = \beta_x \frac{d \kappa_t}{dt},$$

with the rate of change over t of the log-mortality rates in each subpopulation,

$$\frac{d \log \mu_{xtg}}{dt} = \beta_x \lambda_g \frac{d \kappa_t}{dt} = \lambda_g \frac{d \log \mu_{xt}}{dt}$$

From this last equation it is clear that λ_g can be interpreted as the proportionate deviation of mortality change in each subgroup relative to the mortality change in the reference population. That is, if mortality is improving in the reference population (i.e. $d \log \mu_{xt}/dt < 0$) then $\lambda_g > 1$ means that for subpopulation g mortality improves at

a faster pace than it does in the reference population, whereas $\lambda_g < 1$ means that it decreases at a slower pace. Similarly, if mortality from a specific cause in the reference population is worsening (i.e. $d \log \mu_{xt}/dt > 0$) then $\lambda_g > 1$ means that for subpopulation g mortality is worsening at a faster pace than in the reference and $\lambda_g < 1$ means that it is worsening at a slower pace.

3.2.5 Forecasting

Under the proposed model the time evolution of mortality and of mortality differentials is represented by the univariate time index κ_t. Therefore, forecasts of age-subgroup-specific mortality rates by cause of death can be obtained by modelling and forecasting κ_t using autoregressive moving average (ARIMA) models. In general, we assume that κ_t follows an ARIMA(p, q, d) with drift:

$$\Delta^d \kappa_t = \theta_0 + \phi_1 \Delta^d \kappa_{t-1} + \cdots + \phi_p \Delta^d \kappa_{t-p} + \xi_t + \theta_1 \xi_{t-1} + \cdots + \theta_q \xi_{t-q} \tag{3.9}$$

where Δ is the difference operator; θ_0 is the drift parameter, $\phi_1, \ldots, \phi_p$ are the autoregressive coefficients with $\phi_p \neq 0$; $\theta_1, \ldots, \theta_q$ are the moving average coefficients with $\theta_q \neq 0$; and ξ_t is a Gaussian white noise process with variance σ_ξ.

Model (3.9) can then be used to obtain projected values of the time index $\dot{\kappa}_{t_n+r}$, $r = 1, 2, \ldots$, which are then inserted into equation (3.8), omitting the coding adjustment terms, to provide forecasted age-group-specific mortality rates corrected for cause of the coding adjustments:

$$\dot{\mu}_{x,t_n+r,g} = \exp\left(\hat{\alpha}_x + \hat{\alpha}_{xg} + \hat{\beta}_x \hat{\lambda}_g \dot{\kappa}_{t_n+r}\right)$$

3.3 Case study: Mortality by cause of death and deprivation in England

In this section we present an empirical application that illustrates the previous discussion on the modelling of mortality by cause of death and socio-economic stratification.

Specifically, we examine the relationship between deprivation and mortality for the main causes of death in England for the period 1981-2007.

3.3.1 Data

In this study we use the English Index of Multiple Deprivation 2007 (IMD 2007) to measure socio-economic conditions at a small area level (Noble et al., 2007). The IMD 2007 is a composite index of deprivation made up of seven domains with an associated weight[1]. The IMD 2007 is based on small geographies called Lower Layer Super Output Areas (LSOAs). There are 32,482 LSOAs in England with an average population of approximately 1,500 people. In our analysis LSOAs are grouped into deprivation quintiles based on their IMD 2007 score, with Q1 denoting the least deprived quintile and Q5 the most deprived quintile.

Our data comprises death counts for seven major causes of death for the English population disaggregated by IMD 2007 quintiles for the period 1981 to 2007 and age groups, 50-54, ..., 80-84, 85+. We consider six groups of causes of death derived from the International Classification of Diseases (ICD) plus a group "Other" including the remaining deaths. The resulting seven group of causes together with a summary of common causes included in each group are presented in Table 3.1. In this study, we will only consider main causes so that problems with a too detailed disaggregation are avoided. To date, the six selected causes represent around 90% of the total number of deaths for people aged 50+, with Neoplasms, Circulatory and Respiratory diseases constituting around 80% of the deaths.

As for the reference population, we use matching cause of death data for the England and Wales population. These data, which is publicly available at the ONS website[2], covers ages 50-54, ..., 80-84, 85+ and the period 1961-2010. Over this 50 year period cause of death coding in England and Wales has undergone five major changes which are summarised in Table 3.2.

[1]The seven deprivations domains with their percentage participation in the index are: i) Income deprivation (22.5%), ii) Employment deprivation (22.5%), iii) Health deprivation and disability (13.5%), iv) Education, skills and training deprivation (13.5%), v) Barriers to housing and services (9.3%), vi) Crime (9.3%), and vii) Living Environment deprivation (13.5%).

[2]See http://www.ons.gov.uk/ons/rel/subnational-health1/the-21st-century-mortality-files/deaths-dataset--2001-2013/index.html and http://www.ons.gov.uk/ons/rel/subnational-health1/the-20th-century-mortality-files/20th-century-deaths/index.html

Table 3.1: Common causes of death under each ICD Chapter group

ICD Chapter	Deaths in 2010 (age 50+)	Common causes	ICD-9 code	ICD-10 code
Circulatory	30%	Ischemic heart disease, Cerebrovascular disease	390-459	I00-I99
Neoplasms (Cancer)	38%	Lung (trachea,bronchus), Colon (rectum, anus) , Breast, Prostate	140-239	C00-D48
Respiratory	12%	Influenza and pneumonia, Chronic obstructive respiratory disease (bronchitis, emphysema, asthma, other chronic)	460-519	J00-J99
Digestive	5%	Peptic ulcer, Cirrhosis and chronic liver disease, Hernia	520-579	K00-K93
Mental and Behavioural	2%	Dementia, Substance abuse	290-319	F00-F99
External	2%	Non-transport accidents, Intentional self-harm, Motor vehicle accidents	E800-E999	V01-Y89
Other	10%	Diabetes mellitus, Renal failure, Parkinson's, Alzheimer's disease	-	-

TABLE 3.2: Changes in cause of death coding implemented in England and Wales over the period 1961-2010

Change	ICD-8	ICD-9	Rule 3	ACCS	ICD-10
Year	1968	1979	1984	1993	2001

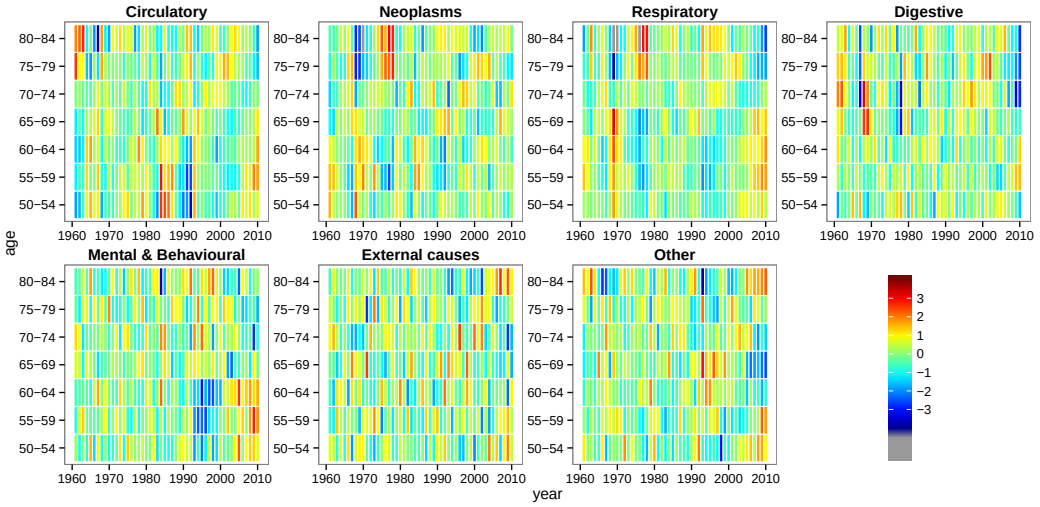
In order to examine the magnitude and trends of socio-economic differentials in mortality for the main causes of death, for each cause we have fitted the The Three-way Lee-Carter model with coding adjustments introduced in Section 3.2.2 to England mortality data for period 1981-2007 with IMD 2007 quintile as covariate and England and Wales data for the period 1961-2010 as reference population data. We have excluded age category 85+ from the analysis as it groups many ages into a single category, having a distorting effect on the results. In addition, we have assumed that mortality trends need to be adjusted for cause of death coding changes in years $s_1 = 1968$, $s_2 = 1979$, $s_3 = 1984$ and $s_4 = 1993$, corresponding to the introduction of ICD-8, ICD-9, the changes in the application of Rule 3 and the introduction of the automated cause coding system (ACCS), respectively. Finally, prior to the fitting of the models, death counts for the period 1993-2000 have been adjusted using a suitable bridge coding algorithm to account for changes in cause of death coding between ICD-9 and ICD-10 (Cook et al., 2002).

Figures 3.7 to 3.13 present for the seven groups of causes of death the estimated parameters for the England Wales reference population, while Figures 3.14 and 3.15 depict the mortality differential parameters for the deprivation subpopulations. We defer, however, the analysis of these fitted parameters until Sections 3.3.4 and 3.3.5 and concentrate first on the quality of the model fits and of the correction for coding adjustments performed by the models.

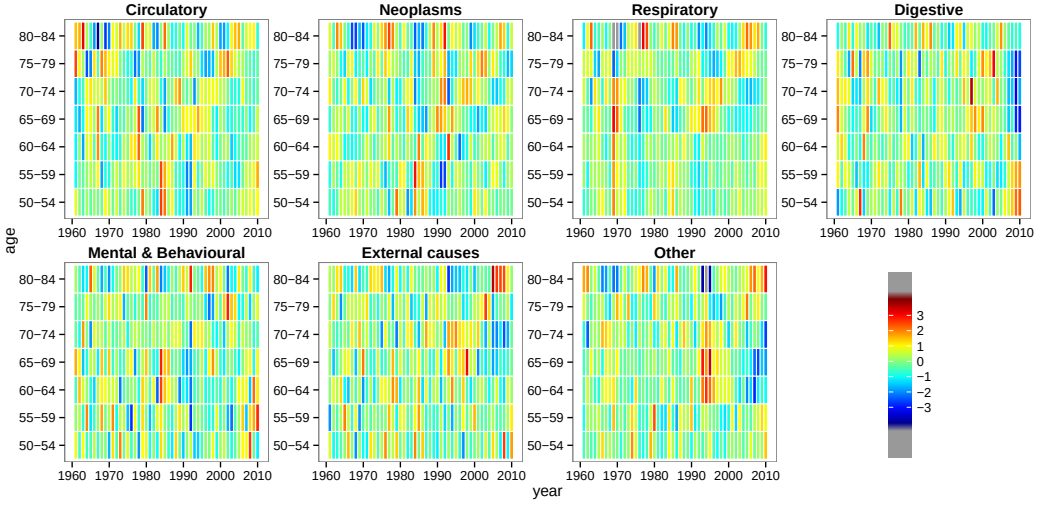
3.3.2 Goodness-of-fit

The goodness of fit of the models can be assessed by inspecting the standardised deviance residuals of the fitted models: regular patterns in the residuals are a sign of the inability of the model to describe all of the phenomena appropriately.

Figure 3.3 plots for each of the seven causes death the standardised residuals of the Lee-Carter model with coding adjustments defined by Equation (3.1) fitted to the male and female England and Wales reference populations. Similarly, Figures 3.4 and 3.5



(A) Men



(B) Women

FIGURE 3.3: Deviance residuals for the Lee-Carter model with coding adjustments fitted to the main causes of deaths in England and Wales.

depict, for men and women respectively, the corresponding deviance residuals for the deprivation subpopulations. From these plots we note the the following:

- There are distinctive diagonal patterns for Circulatory diseases, Neoplasms and, to a lesser extent, Respiratory diseases, suggesting the presence of a cohort effect for these causes of death. We also note that the diagonal clustering of residuals are stronger for the female population with the patterns observed for Neoplasms in the two most deprived quintiles of England being particularly strong.

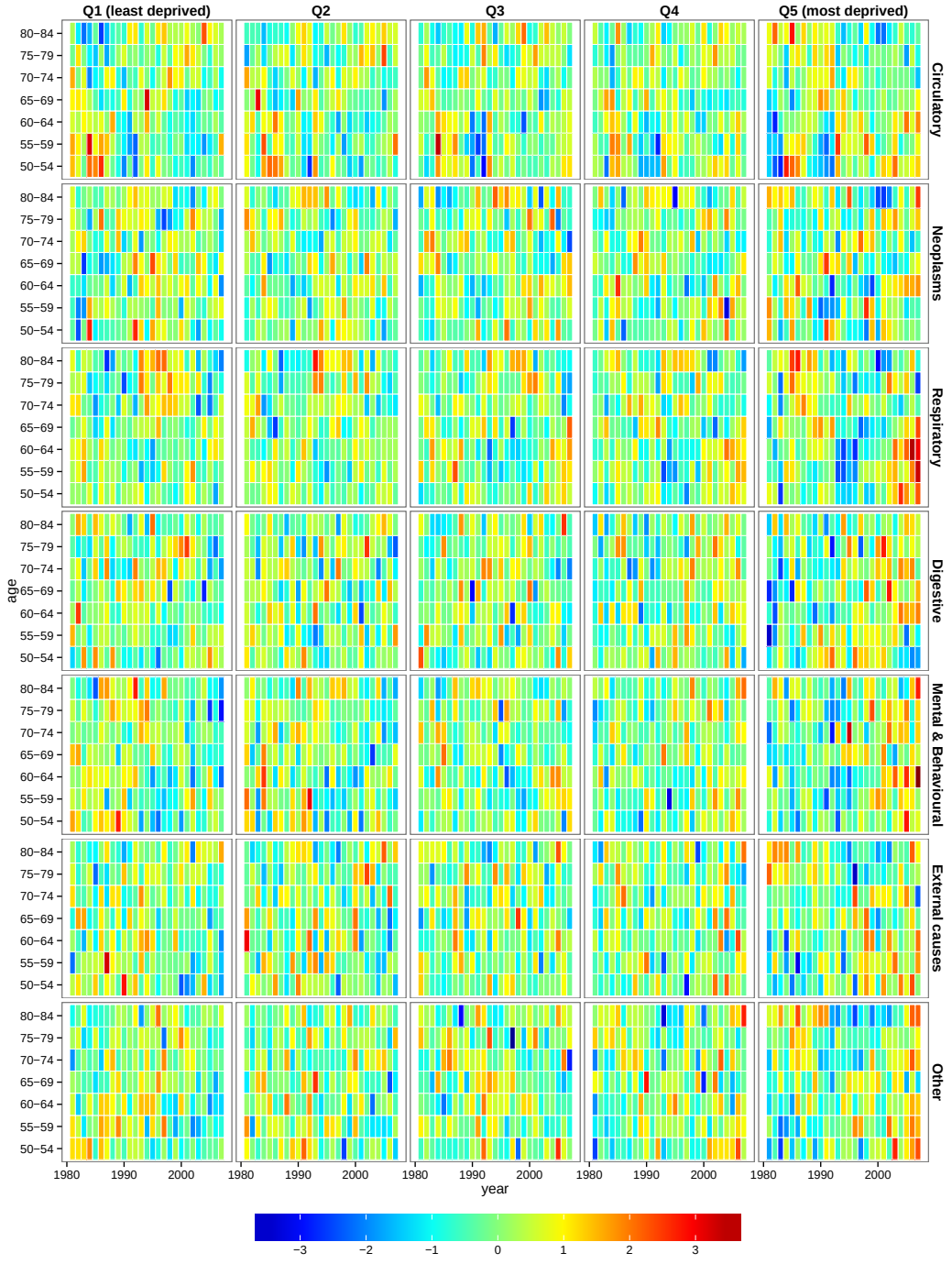


FIGURE 3.4: Deviance residuals for the Three-way Lee-Carter model with coding adjustments fitted to the main causes of deaths in the male deprivation subpopulations of England.

- In spite of some localised concentration of large residual (e.g. for Digestive causes in England and Wales for the most recent years), the residuals for Digestive diseases, Mental and Behavioural diseases, External and "Other" causes show a reasonably random pattern, indicating a satisfactory fit to the data.

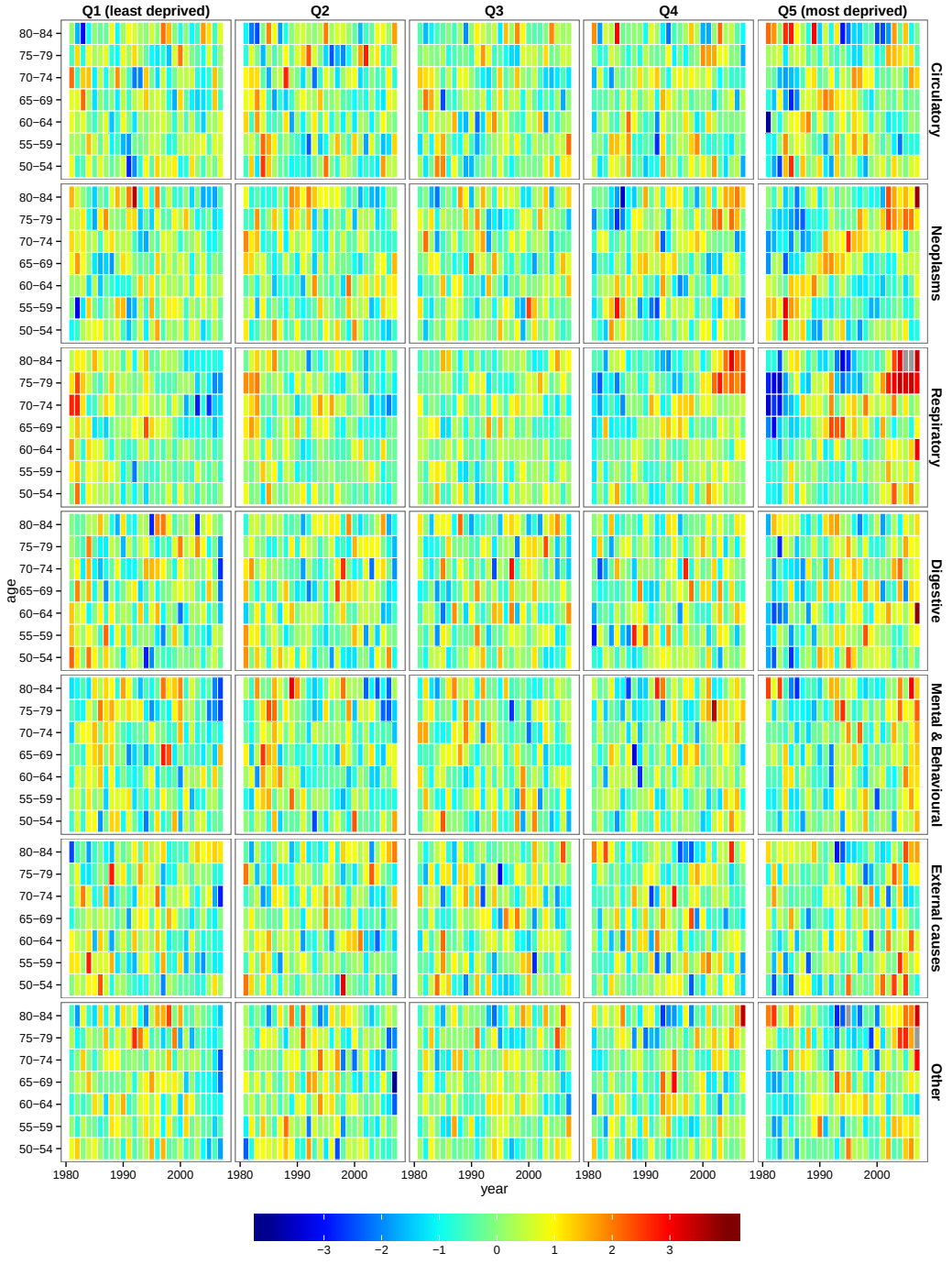


FIGURE 3.5: Deviance residuals for the Three-way Lee-Carter model with coding adjustments fitted to the main causes of deaths in the female deprivation subpopulations of England.

- There are generally more colourful residuals at older ages which suggest variable dispersion of mortality with age.

- There are notably higher residuals for the most deprived quintile of England. This may suggest the presence of different underlying risk factors driving causes of death

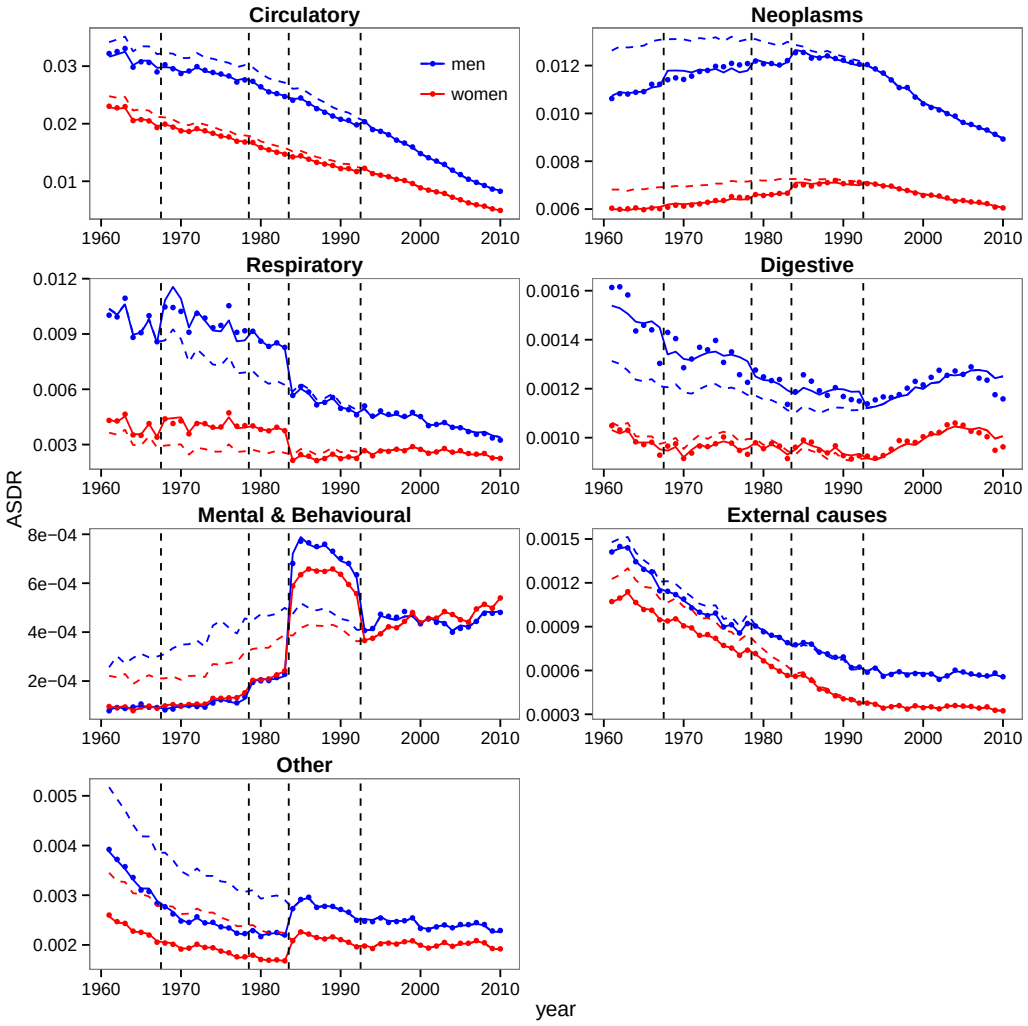


FIGURE 3.6: Age standardised death rates (ASDR) for the age range 50-84 for the main causes of death in England and Wales reference population. Dots show observed ASDRs, solid lines fitted ASDRs and dashed lines coding-adjusted ASDRs. Vertical dashed lines indicate coding-adjustment years.

in this subpopulation.

Overall, we see that the fit of the Lee-Carter model with coding adjustments is reasonable, particularly after 1981 which is the period of main interest for the analysis of mortality differentials.

3.3.3 Evaluation of coding adjustments

In order to examine the performance of the fitted models with regards to the quality of the coding adjustments, we plot in Figure 3.6 observed, fitted and code-adjusted cause-specific age standardised death rates (ASDR) in England and Wales for the age range

50-84. These ASDRs have been computed using the expression

$$ASDR_t = \frac{\sum_x \omega_x \mu_{xt}}{\sum_x \omega_x}$$

where ω_x represents the 1976 European Standard Population (Ahmad et al., 2001) and μ_{xt} corresponds to

- d_{xt}/e_{xt}, in the case of the observed ASDRs, with d_{xt} being the observed number of deaths from a given cause;
- $\exp\left(\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t + \sum_{i=1}^{4} \hat{\delta}_x^{(i)} f^{(i)}(t)\right)$ in the case of the fitted ASDRs; and
- $\exp\left(\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t\right)$ in the case of the code-adjusted ASDRs.

Referring to Figure 3.6 we highlight the following important features:

- The reassuringly close alignment of the fitted and observed ASDRs for all causes and both genders, with only minor deviations for Digestive diseases in the most recent and most distant years.
- The satisfactory adjustment for the marked trend breaks induced by the broadening in the application of Rule 3 of the ICD-9 introduced in England after 1984. The change in the application of Rule 3 meant that some "terminal conditions" were not coded as the underlying cause of death if some other "major condition" was reported in the death certificate (Rooney and Devis, 1996; Brock et al., 2006; Goldacre et al., 2003). One of the main effects of this was a sudden fall in the numbers of deaths from Respiratory diseases with Mental and Behavioural conditions (among other diseases) being the other side of the change in the application of Rule 3 and experiencing a sudden increase in the numbers of deaths (Rooney and Devis, 1996; Griffiths and Rooney, 2006).
- The noticeable improvement in consistency observed in the adjusted trends for Mental and Behavioural diseases. However, the nature of the Lee-Carter model with code adjustments means that it can only correct for sustained jumps and cannot remove gradual changes in coding as those that seem to have affected Mental and Behavioural diseases in the period 1984-1993. In fact, Rooney and

Devis (1996) and Griffiths and Rooney (2006) report that for this period there was a gradual drift in certification from terms such as senile dementia, coded under Mental and Behavioural causes, to more specific conditions such as Alzheimer's disease, coded under diseases of the nervous system which belong to the "Other" group of causes used in this chapter (see Table 3.1).

- The possible "false-positive" trend adjustment performed by the model in 1994 for Digestive diseases in men, where an apparent genuine change in trend may have been mistakenly confused with a change in coding.

Overall, the proposed Lee-Carter model with cause of death coding adjustments performs very well in recovering the underlying trend of mortality for the different causes of death. This will be crucial for the satisfactory forecasting of these trends as performed in Section 3.3.6.

3.3.4 Historical cause-specific mortality patterns: England and Wales 1961-2010

Having seen that the Three-way Lee-Carter model with coding adjustments shows a reasonable goodness-of-fit and performs a satisfactory adjustment for cause of death coding changes, we now use this model to analyse the historical pattern of cause-specific mortality for the period 1961-2010 in the England and Wales population and to analyse the differentials in mortality across deprivation quintiles in England for the period 1981-2007.

Figures 3.7 to 3.13 depict for the seven groups of causes of death the parameter estimates of the Lee-Carter model with coding adjustments fitted to mortality data for men and women in the England and Wales reference population. From these figures we note the following important features:

- The approximately linear increase with age in log death rates captured by parameters α_x for the three main causes Neoplasms, Circulatory and Respiratory diseases. Due to the importance of these causes in the total number of deaths, their age patterns determine the exponential growth of mortality with rising age which characterises all-cause mortality.

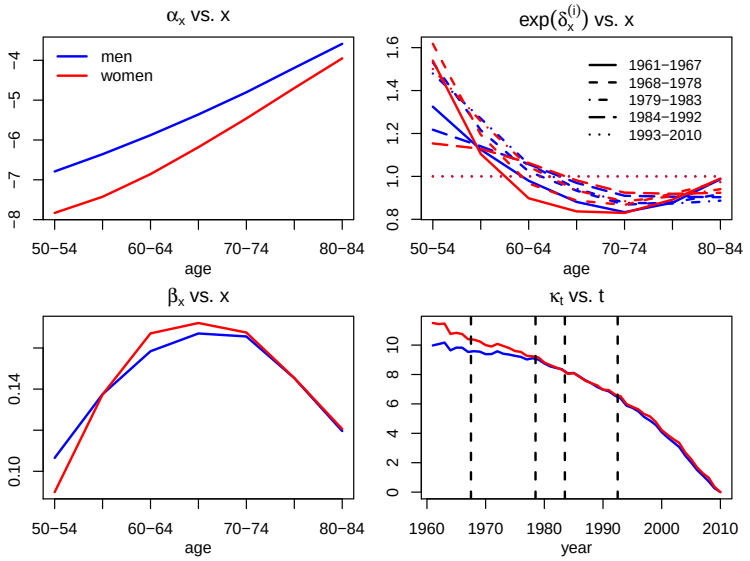


FIGURE 3.7: Fitted parameters for the Lee-Carter model with code adjustments fitted to Circulatory causes of death in England for the period 1961-2010.

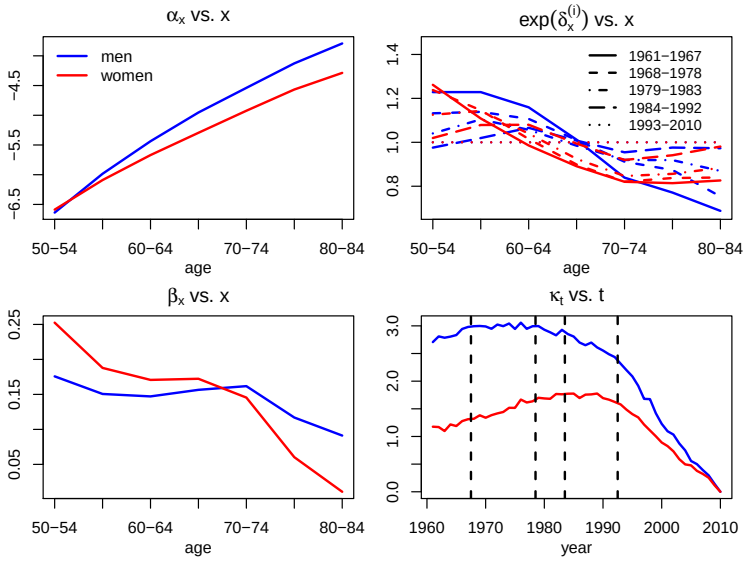


FIGURE 3.8: Fitted parameters for the Lee-Carter model with code adjustments fitted to Neoplasms in England for the period 1961-2010.

- The contrasting age patterns evidenced by the α_x parameters in Mental and Behavioural causes and External causes, with death rates staying approximately constant from age 50 to age 60 in men and increasing very slowly in women.

- The steady and significant decline in death rates from Circulatory diseases captured by the κ_t parameters, with mortality improvements peaking at age 65-69 as shown by the β_x parameters (see Figure 3.7). This considerable reduction in mortality has resulted in Circulatory causes passing from being the main cause of

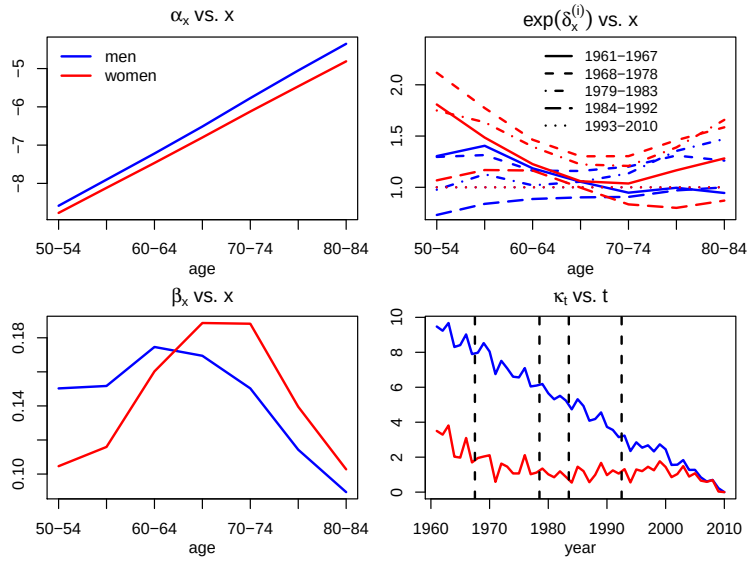


FIGURE 3.9: Fitted parameters for the Lee-Carter model with code adjustments fitted to Respiratory diseases in England for the period 1961-2010.

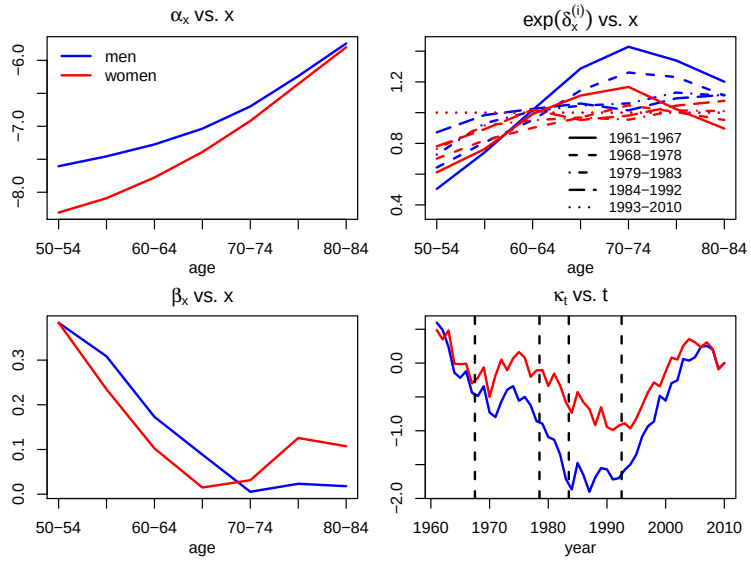


FIGURE 3.10: Fitted parameters for the Lee-Carter model with code adjustments fitted to Digestive diseases in England for the period 1961-2010.

death in 1961 to being in 2010 the second cause of death after Neoplasms for both sexes.

- The evident changes in trend in the κ_t parameters for Neoplasms, showing a noticeable faster improvement of cancer mortality from 1993 onwards (see Figure 3.8). It is also worth noticing that while cancer mortality for men started to decline in the late 1970s, for women the decline in cancer mortality was delayed until the end of the 1980s. These gender differentials in cancer mortality can be partly attributed to gender differences in smoking behaviour with cigarette consumption

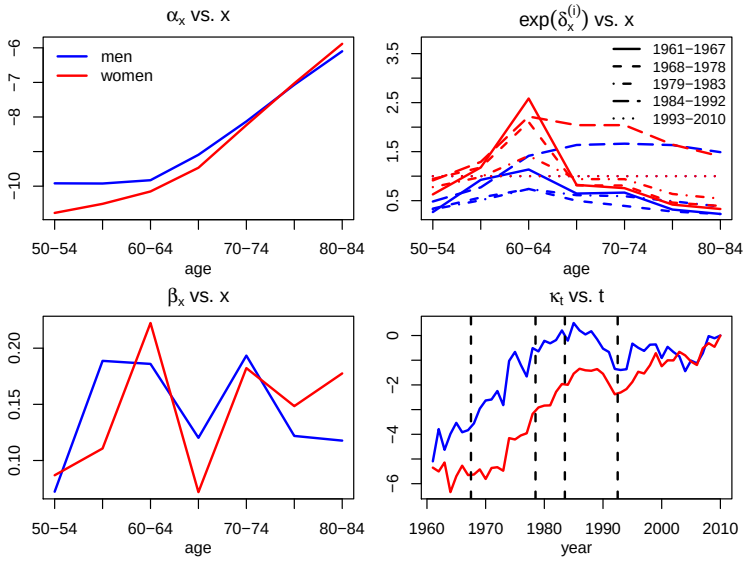


FIGURE 3.11: Fitted parameters for the Lee-Carter model with code adjustments fitted to mental & behavioural diseases in England for the period 1961-2010.

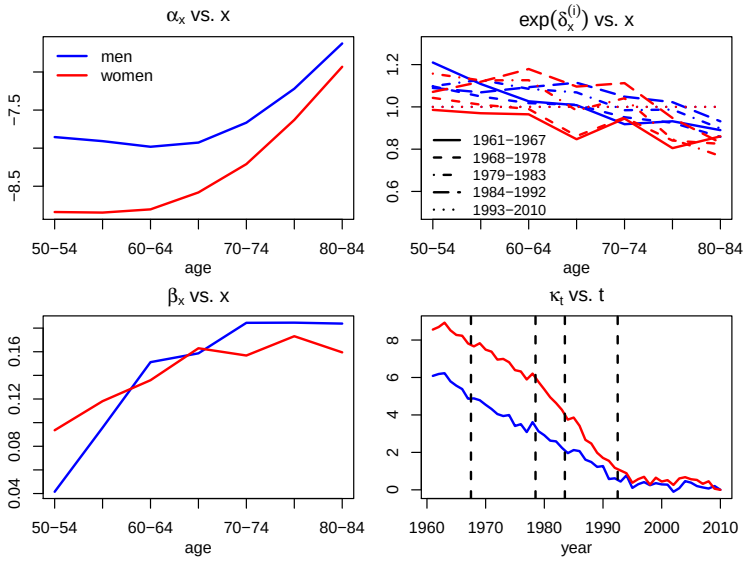


FIGURE 3.12: Fitted parameters for the Lee-Carter model with code adjustments fitted to External causes of death in England for the period 1961-2010.

peaking in the 1950s for men and in the 1960s for women (Willets, 2004; Di Cesare and Murphy, 2009).

- The contrasting trends by gender observed in κ_t for Respiratory diseases, with men showing a steady decline in mortality while women show a stagnation of the mortality rates (see Figure 3.9). These differentials in trends among men and women can also be partly attributed to gender differences in smoking prevalence trends and offer a possible explanation for the recent gender convergence in life expectancy reported in Mayhew and Smith (2014).

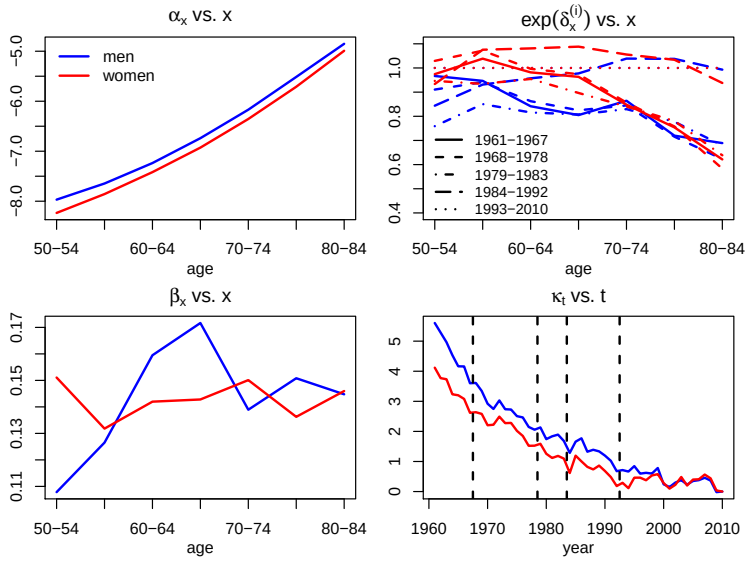


FIGURE 3.13: Fitted parameters for the Lee-Carter model with code adjustments fitted to "Other" causes of death in England for the period 1961-2010.

- The sharp increase in mortality rates from Digestive causes after 1994 indicated by the trends in κ_t for both genders (see Figure 3.10). As a result of this increase, Digestive diseases now represent a significant proportion of the total death rates. We note however that the increase of mortality rates from Digestive causes appears to be reverting in the most recent years.

- The complex patterns observed in κ_t and β_x for Mental and Behavioural causes (see Figure 3.11). As pointed out by Griffiths and Rooney (2006), trends in mortality from mental conditions have been strongly affected by changes in cause of death coding rules; hence trends from these conditions need to be analysed with care. For instance, as discussed before, the decline in mortality observed between 1984 and 1993 appears to be the result of a gradual transfer of deaths from dementia to Alzheimer's disease.

3.3.5 Historical cause-specific mortality differentials: England 1981-2007

In the Three-Way Lee-Carter model we can assess level and trend differentials in mortality by examining parameters α_{xg} and λ_g, respectively. Figure 3.14 plots the estimated values of $\exp(\alpha_{xg})$ for the different causes of death in the England deprivation subpopulations. Recalling the discussion in Section 3.2.4 and since in we are using year 2010 as the

TABLE 3.3: Relative mortality level differentials for the different causes of death between the most and least deprived quintiles of England ($\exp(\alpha_{x,Q5})/\exp(\alpha_{x,Q1})$).

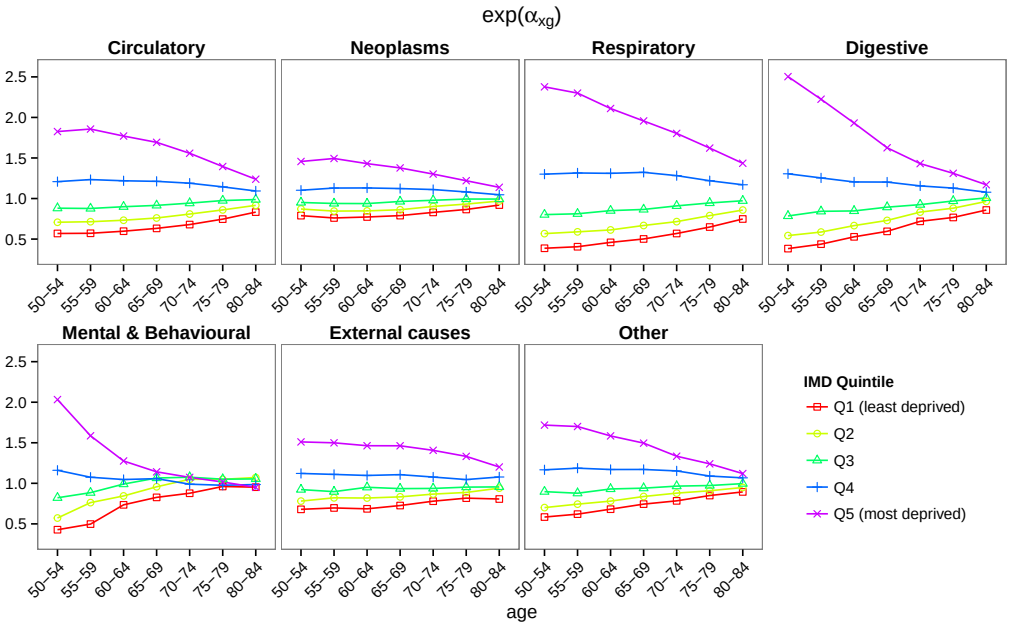
	Men						
	50-54	55-59	60-64	65-69	70-74	75-79	80-84
Circulatory	3.21	3.25	2.96	2.68	2.29	1.87	1.49
Neoplasms	1.85	1.97	1.85	1.74	1.57	1.41	1.24
Respiratory	6.14	5.67	4.59	3.90	3.17	2.50	1.92
Digestive	6.54	5.09	3.66	2.73	1.99	1.71	1.36
Mental & Behavioural	4.75	3.19	1.74	1.38	1.22	1.06	1.00
External causes	2.22	2.15	2.13	2.01	1.80	1.63	1.49
Other	2.94	2.74	2.33	2.01	1.70	1.46	1.25
	Women						
	50-54	55-59	60-64	65-69	70-74	75-79	80-84
Circulatory	4.05	4.12	3.71	3.07	2.42	1.81	1.41
Neoplasms	1.50	1.61	1.61	1.61	1.48	1.30	1.16
Respiratory	4.41	4.54	4.01	3.55	2.87	2.06	1.58
Digestive	3.82	3.18	2.67	2.27	1.88	1.69	1.37
Mental & Behavioural	3.20	2.23	1.96	1.30	1.33	1.23	1.13
External causes	1.73	1.66	1.64	1.60	1.54	1.42	1.28
Other	2.26	2.31	2.06	1.81	1.60	1.35	1.15

reference year in our modelling (i.e. $\kappa_{2010} = 0$), $\exp(\alpha_{xg})$ can be interpreted as the estimated percentage deviation in 2010 of mortality at age x in subpopulation g relative the England and Wales reference population. Similarly, the values of $\exp(\alpha_{x,Q5})/\exp(\alpha_{x,Q1})$, which are reported in Table 3.3, represent the estimated mortality rate ratio at age x between the most and least deprived quintiles of England. These quantities give an indication of relative mortality differentials for the different causes of death. From Figure 3.14 and Table 3.3 we note the following:

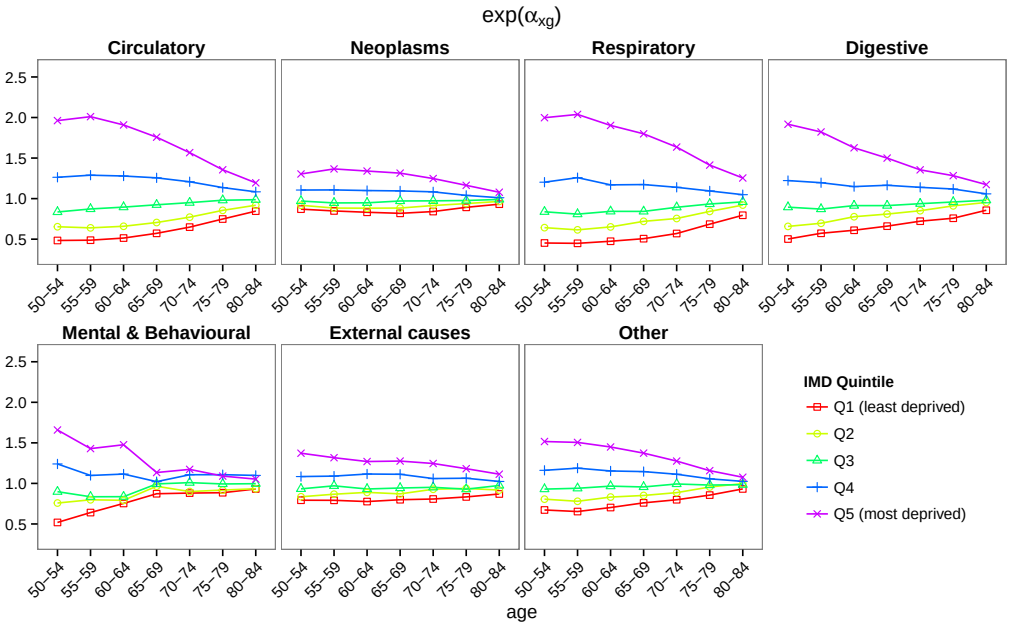
- There is consistency in level differentials in mortality for all seven groups of causes, with more deprived quintiles showing significantly higher death rates than less deprived quintiles. However, these differentials show important variations across causes, sexes and ages.

- Circulatory, Respiratory and Digestive causes show the most marked relative mortality differentials while Neoplasms and External causes show the smallest differentials across deprivation quintiles. In particular, the relative mortality differentials for Respiratory and Digestive causes in the male population are substantial. For example, men aged 50-54 in the most deprived quintile have for these two causes of death more than six times the mortality of men aged 50-54 in the least deprived quintile.

- In terms of age, there is a general decrease in relative mortality differentials as people get older. For instance, men and women aged 50-54 in the most deprived



(A) Men



(B) Women

FIGURE 3.14: Mortality level differentials for the different causes of death across deprivation quintiles of England ($\exp(\alpha_{xg})$)

quintile have, respectively, 3.21 and 4.05 higher mortality from Circulatory diseases than persons of the same sex and age in the least deprived quintile. By contrast, at age 80-84 the mortality ratio from Circulatory causes between the most deprived quintile and the least deprived quintile reduces to 1.49 and 1.41 for men and women, respectively.

TABLE 3.4: Mortality trend differentials for the different causes of death for the deprivation quintiles of England (λ_g). AI: average percentage annual improvement in England and Wales.

Men

Cause	λ_{Q1}	λ_{Q2}	λ_{Q3}	λ_{Q4}	λ_{Q5}	$AI_{1981-2010}$	$AI_{1994-2010}$
Circulatory	1.10	1.08	1.02	0.92	0.81	4.21	5.25
Neoplasms	0.82	0.90	0.94	0.96	0.97	1.42	1.96
Respiratory	1.09	1.06	0.98	0.84	0.77	2.62	2.10
Digestive	0.62	0.71	0.94	1.17	1.48	-0.56	-1.34
Mental & Behavioural	1.70	1.86	1.45	0.32	-0.09	-0.16	-1.22
External causes	1.08	1.07	1.02	0.94	0.76	1.29	0.69
Other	1.09	1.23	1.11	0.75	0.61	0.90	0.59

Women

Cause	λ_{Q1}	λ_{Q2}	λ_{Q3}	λ_{Q4}	λ_{Q5}	$AI_{1981-2010}$	$AI_{1994-2010}$
Circulatory	1.09	1.08	0.99	0.92	0.85	4.26	5.35
Neoplasms	0.92	0.98	0.98	0.91	0.83	0.83	1.33
Respiratory	1.31	0.95	1.01	1.00	0.86	0.42	0.50
Digestive	0.31	0.80	0.93	1.37	1.72	-0.08	-0.86
Mental & Behavioural	0.71	0.45	0.85	1.46	1.50	-1.39	-1.94
External causes	1.00	0.97	1.02	0.95	0.96	2.44	0.79
Other	1.06	1.09	0.98	0.69	0.85	0.55	0.10

- In terms of gender, there are bigger mortality differences in men than in women for all causes with the notable exception of Circulatory diseases, where women show slightly higher differentials than men.

Figure 3.15 depicts 95% confidence intervals of parameters λ_g for the different causes of deaths. The corresponding central estimates are reported in Table 3.4. Since the interpretation of λ_g is closely linked to the direction and pace of mortality change for a specific cause (see Section 3.2.4), Table 3.4 also includes average annual cause-specific improvement rates in England and Wales for the periods 1981-2010 and 1994-2010. This latter period is of interest as several causes of death seem to have experienced a trend change after 1994. The average improvement (AI) rates have been computed as $AI_{1981-2010} = \frac{1}{k} \times \frac{\kappa_{1981}}{2010-1981}$ and $AI_{1994-2010} = \frac{1}{k} \times \frac{\kappa_{1994}}{2010-1994}$, where $k = 7$ is the number of age bands used in the fitting of the models.

Figure 3.15 and Table 3.4 reveal the following in relation to mortality trend differentials across deprivation quintiles of England:

- Trend differentials are not as clear and consistent as level differentials, with the plots of λ_g showing a variety of patterns for the different causes of death.

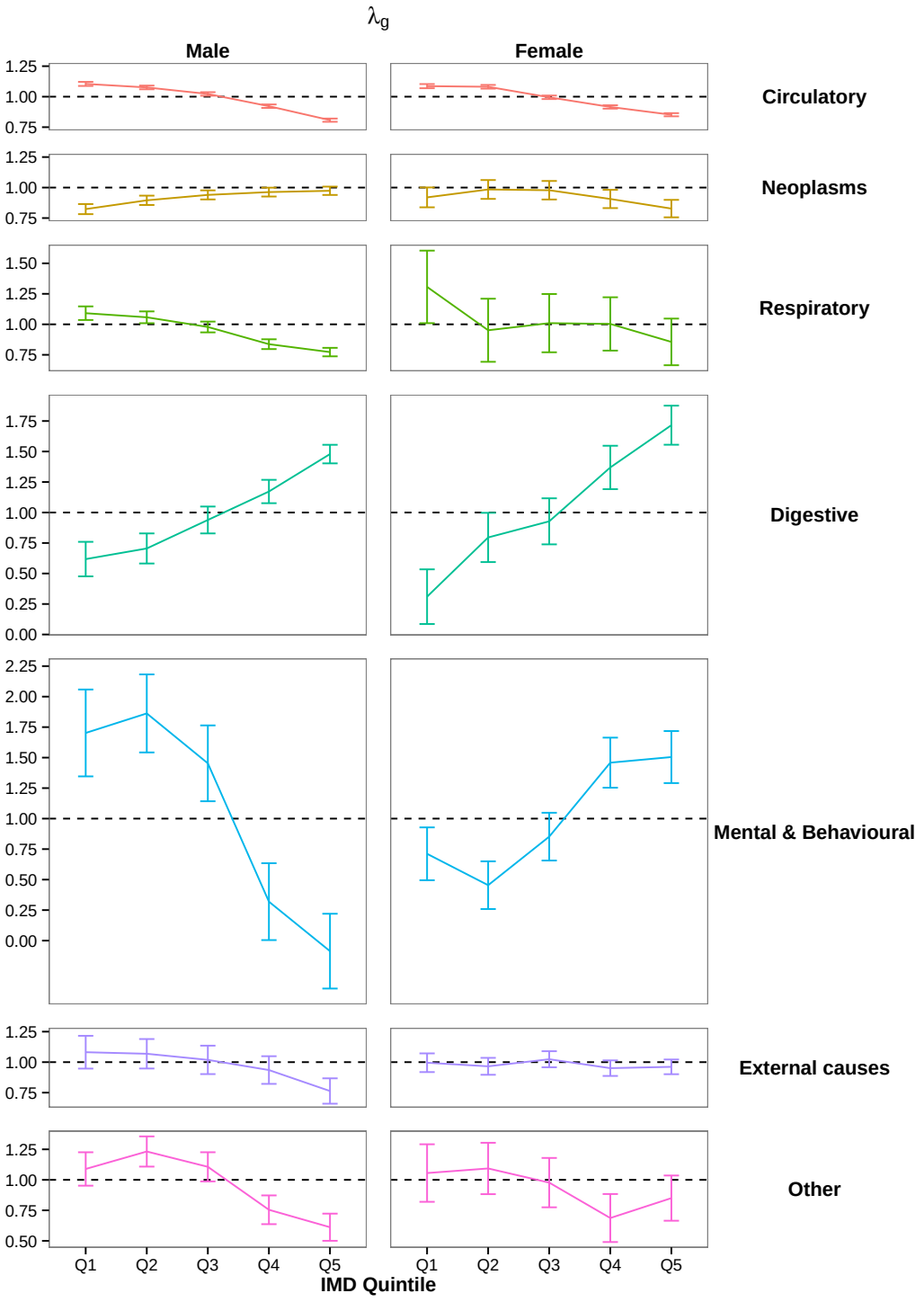


FIGURE 3.15: Mortality trend differentials for the different causes of death for the deprivation quintiles of England (λ_g).

- For Circulatory causes, for instance, there is a clear socio-economic gradient in mortality improvements for both genders, with least deprived quintiles experiencing significantly faster mortality improvements than most deprived quintiles.

Since Circulatory diseases were the main cause of death over the 1981-2007 period, improvement differentials from this cause are the main determinant of trend differentials in all-cause mortality. Therefore, it is not surprising to see that the magnitude of improvement differentials in mortality from Circulatory diseases is in close agreement with the all-cause trend differentials across deprivations quintiles reported in Chapter 2.

- In contrast, Neoplasms show less clear-cut mortality trend differences. Whilst the least deprived quintile, Q1, has experienced the slowest cancer mortality improvements in the male population, it is the most deprived quintile, Q5, which shows the slowest improvements among women. This suggests that relative differentials in cancer mortality are decreasing for males, but are widening, although very weakly, for females.

- The contrasting patterns in mortality improvement differences across genders for Respiratory diseases is noteworthy. While the mortality improvement gradient among men is very clear with men in the most deprived quintile of England showing a significantly slower reduction in mortality from Respiratory causes than men in less deprived areas, the differences in mortality improvements for the different deprivation quintiles of the female population are not statistically significant (i.e. error bars for λ_g crossing 1).

- The marked increase in mortality from Digestive causes seen by both sexes after 1994 has been accompanied by very marked socio-economic differences, with mortality for the most deprived quintiles worsening at a much faster pace than for the least deprived quintiles. This mortality improvement gradient is substantial. For instance, mortality for the most deprived quintile has worsened 2.4 and 5.5 times faster than in the least deprived quintile for men and women, respectively. We note, however, that these stark trend differences also mean that Digestive diseases have the highest potential for reducing inequalities: if the overall increase in digestive mortality starts to reverse as seen in the most recent years, it will be the most deprived parts of England which will benefit the most from this mortality reduction.

- There appears to be a very pronounced mortality trend gradient for Mental and Behavioural causes, but with opposite directions for men and women. While relative mortality differences in the male population seem to be reducing with men in most deprived areas worsening at a slower pace than men in least deprived areas, relative differences in the female population appear to be widening with women in most deprived areas worsening at a faster pace than women in least deprived areas. These apparent strong differentials need to be analysed carefully though. As discussed before, mortality trends for Mental conditions are complex due to the important impact that changes in cause of death coding have had on these diseases. Moreover, the observed big relative differences in improvements do not necessarily imply big differences in absolute terms because the increases of mortality rates from Mental and Behavioural causes seems to be decelerating in recent years as indicated by κ_t in Figure 3.11.

- Improvement differences in mortality are in general less clear in women than in men. In the male population all the seven groups of causes of death show statistically significant improvement differences. By contrast, in the female population Neoplasms, Respiratory diseases, External causes and "Other" causes of death show non-significant improvement differences (i.e. error bars for λ_g crossing 1). This in agreement with the less clear differences in all-cause mortality by deprivation observed in Chapter 2 for the English female population.

- The very small differences between the λ_g values for two least deprived quintiles, Q1 and Q2, for most causes of death indicate that these two quintiles have experienced similar mortality improvements. This fact was previously noticed in Chapter 2 where it was found that improvement differences in all-cause mortality for these two quintiles over the 1981-2007 period were negligible.

3.3.6 Cause-specific mortality differential projections: England 2008-2030

In order to project cause-specific mortality rates and examine the possible trend of mortality differentials, we use ARIMA models to extrapolate the time index κ_t for each cause of death. Table 3.5 specifies the ARIMA(p, d, q) model used for each of the seven causes of death. These specifications have been selected among all possible

TABLE 3.5: ARIMA model used to forecast the time index κ_t for the different causes of death.

ICD Chapter	**Men**	**Women**
Circulatory	ARIMA$(0, 2, 2)$	ARIMA$(2, 2, 1)$
Neoplasms	ARIMA$(1, 2, 1)$	ARIMA$(1, 2, 1)$
Respiratory	ARIMA$(2, 1, 0)$ with drift	ARIMA$(2, 1, 2)$ with drift
Digestive	ARIMA$(0, 1, 0)$	ARIMA$(0, 1, 0)$
Mental Behavioural	ARIMA$(0, 1, 1)$ with drift	ARIMA$(0, 1, 0)$ with drift
External	ARIMA$(1, 1, 0)$ with drift	ARIMA$(0, 1, 0)$ with drift
Other	ARIMA$(0, 2, 2)$	ARIMA$(0, 2, 2)$

combinations of $p, d, q \in \{0, 1, 2\}$ so to optimise the BIC of the model coupled with some subjective judgement to ensure consistence between genders. In particular, even though an ARIMA$(0, 2, 1)$ was the optimal model for Digestive causes in the male population, we have decided to use instead an ARIMA$(0, 1, 0)$, the optimal model of order $d = 1$ of integration, so that the models for both sexes have the same order of differencing. For similar reasons, we have chosen to use an ARIMA$(0, 2, 2)$ to forecast mortality from "Other" causes of death in the female position as opposed to the ARIMA$(0, 1, 0)$ with drift dictated by the BIC criterion.

Figures 3.16 and 3.17 show fancharts depicting 50%, 80% and 95% prediction intervals of age-specific mortality rates for the different causes of death at selected ages in the England and Wales male and female reference populations. These figures show that forecasts for most causes of death an age groups look visually consistent with the historical evolution of mortality, both in terms of central projections and of forecasted variability. A noticeable exception is mortality from Digestive diseases for men aged 75-79 and women aged 65-69, where the forecasted levels of uncertainty look very low relative to the historical variability.

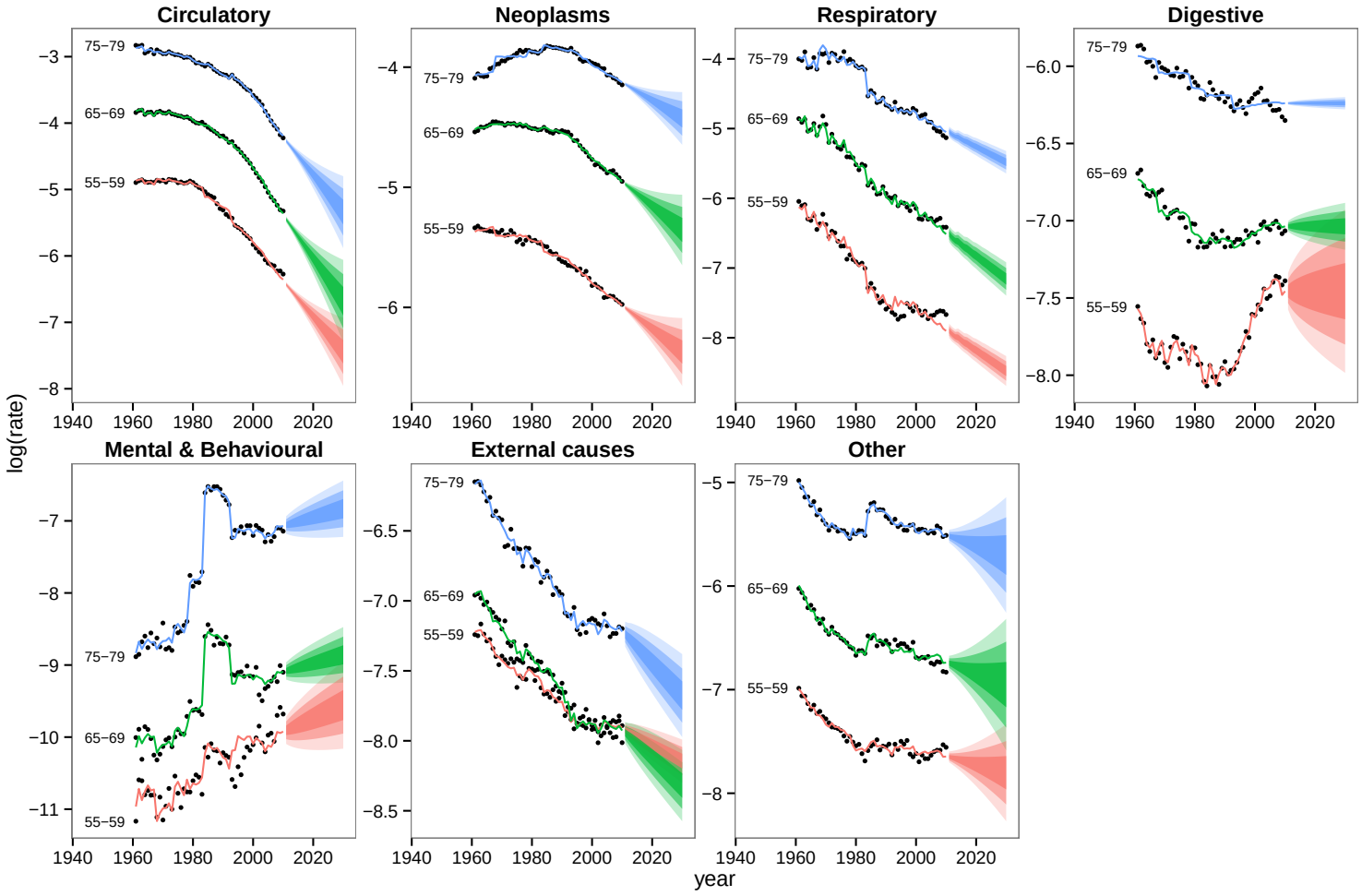


FIGURE 3.16: Fan charts for mortality rates μ_{xt} at selected ages for the different causes of death in the England and Wales male population. The dots show historical mortality rates for 1961-2010. Shades in the fan represent prediction intervals at the 50%, 80% and 95% level.

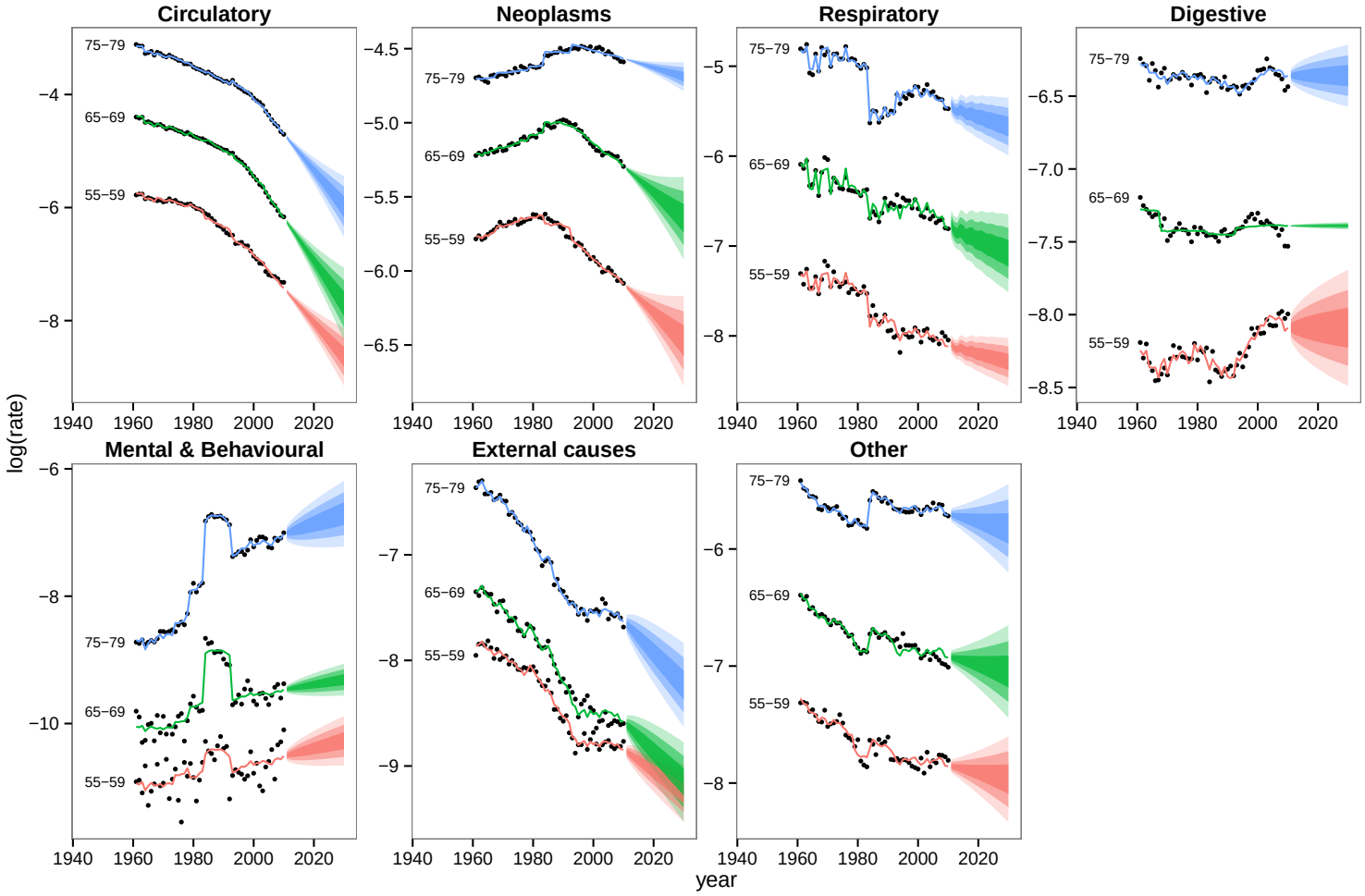


FIGURE 3.17: Fan charts for mortality rates μ_{xt} at selected ages for the different causes of death in the England and Wales female population. The dots show historical mortality rates for 1961-2010. Shades in the fan represent prediction intervals at the 50%, 80% and 95% level.

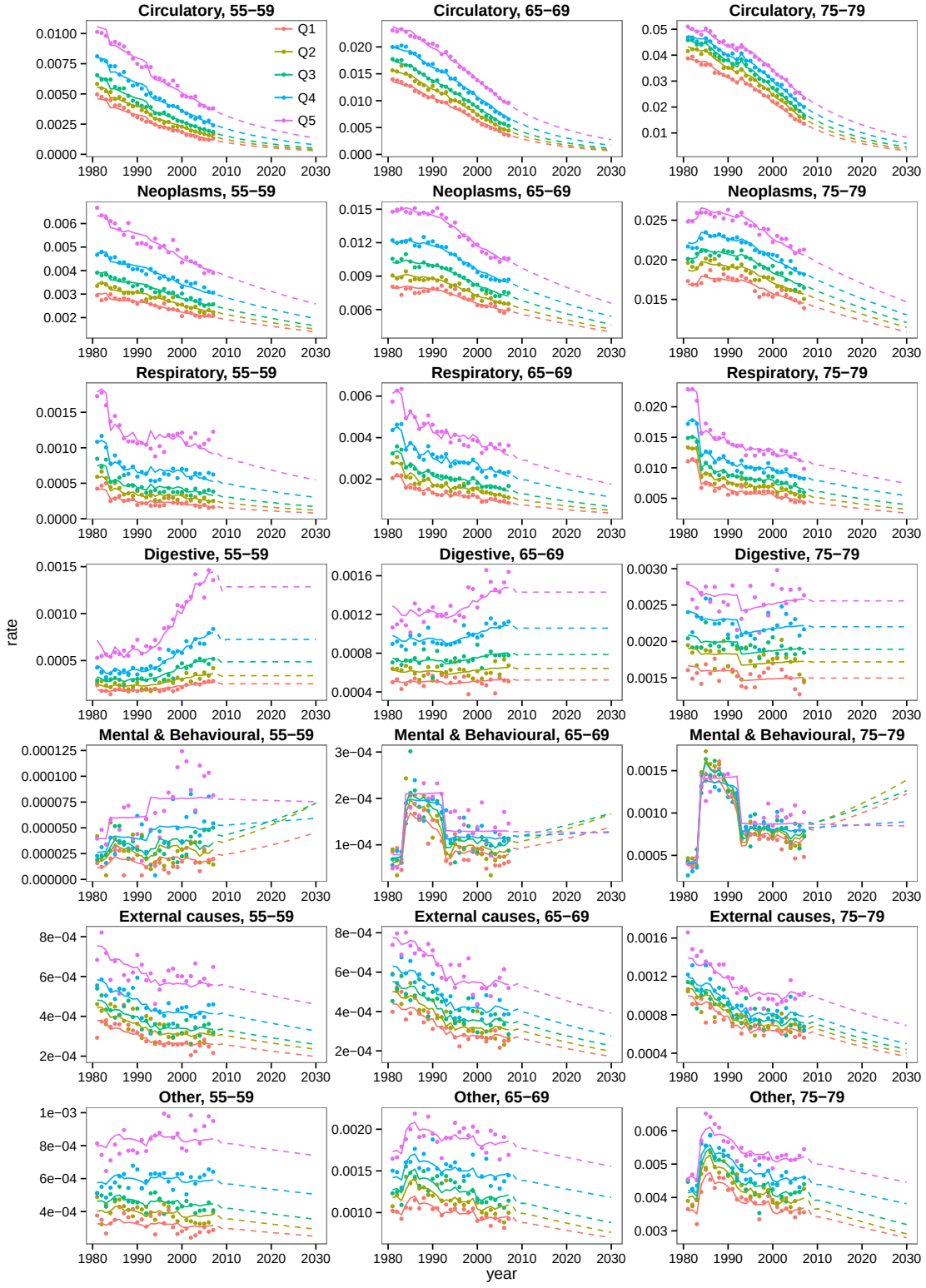


FIGURE 3.18: Time series of fitted and forecasted cause-specific mortality rates μ_{xtg} for the deprivation subpopulation of England males.The dots show observed rates for the period 1981-2007.

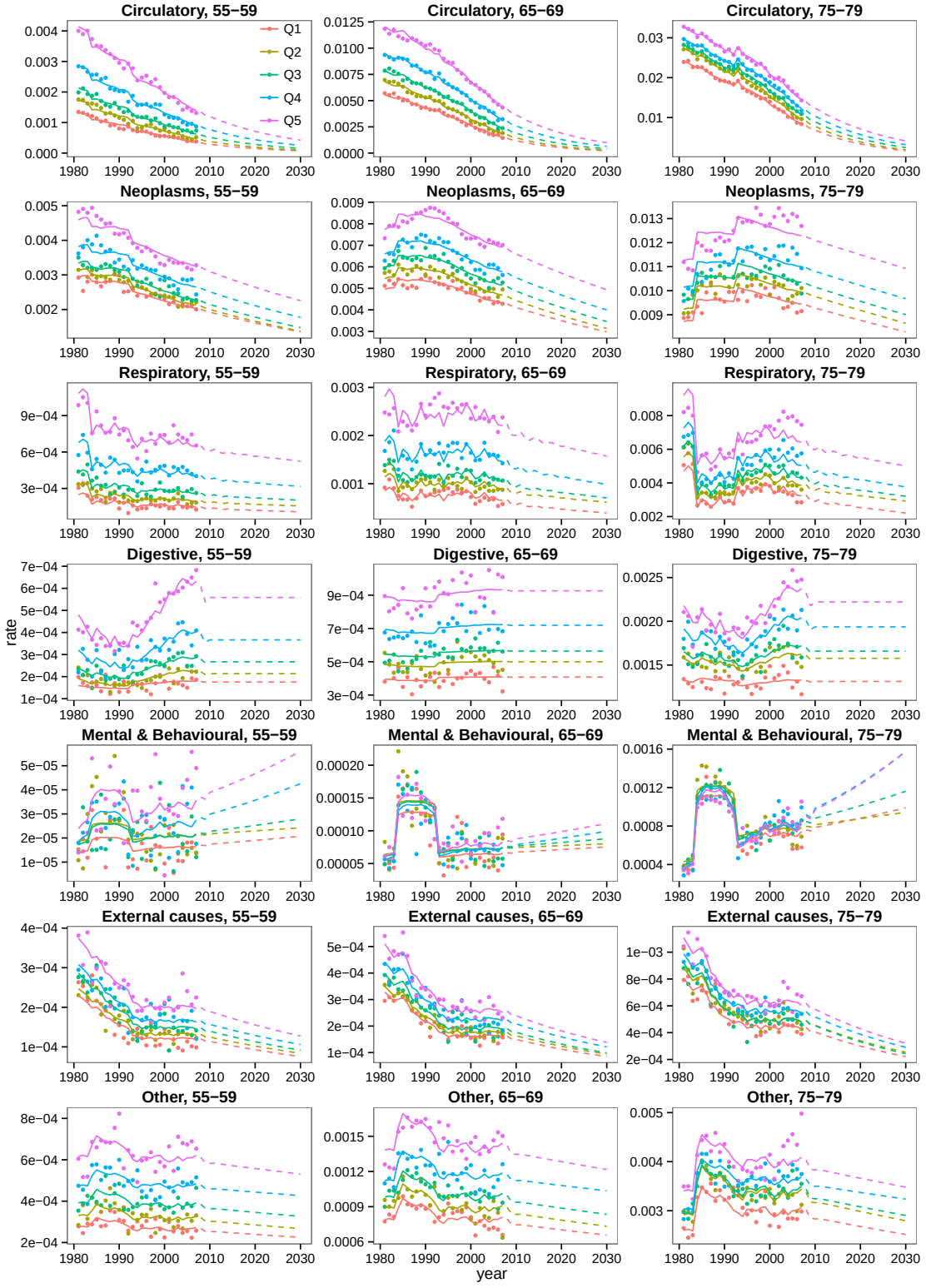


FIGURE 3.19: Time series of fitted and forecasted cause-specific mortality rates μ_{xtg} for the deprivation subpopulation of England females.The dots show observed rates for the period 1981-2007.

The exceptional performance of the Lee-Carter model with code adjustments in capturing the variety of mortality trends observed among the different causes of death is noteworthy, especially in light of the work of Di Cesare and Murphy (2009) who suggest that different modelling approaches may be required for different causes of death. In particular, Di Cesare and Murphy (2009) have found that the traditional Lee-Carter model was unable to capture the inverted U-Shape observed in cancer mortality. The contrasting good performance of the Lee-Carter with code adjustments in this case is explained by an unexpected side effect of the inclusion of the coding adjustments parameters, $\delta_x^{(i)}$, which are capturing some of the cohort effects inherent to cancer mortality.

Figures 3.18 and 3.19 depict for men and women in England, respectively, the project evolution of deprivation-specific mortality rates at selected ages for the the different causes of death. Both figures clearly show that the Three-way Lee-Carter model with code-adjustments is in general able to fit and forecast satisfactorily the mortality trends across deprivation subgroups and causes of death. We note, however, that Mental and Behavioural causes are a noticeable exception with the forecasted deprivation-specific mortality trends looking clearly implausible.

In order to examine the possible future evolution of mortality differentials, we plot in Figure 3.20 historical trends and projections of the relative age-specific mortality differences between the most and least deprived quintiles, $\mu_{x,t,Q5}/\mu_{x,t,Q1}$, for the seven groups of causes of death. Matching absolute differences, $\mu_{x,t,Q5} - \mu_{x,t,Q1}$, are plotted in Figure 3.21. These two figures also present differences in cause-specific ASDRs (thicker black lines), whose values for selected years are tabulated in Tables 3.6 and 3.7 for relative and absolute differences, respectively. In Figures 3.20 and 3.21 we can distinguish four main patterns in the trends of mortality differentials among the different causes:

- First, Circulatory and Respiratory diseases show a reduction of absolute mortality differentials while at the same time, due to the faster mortality improvement experienced by the least deprived subpopulation, experiencing a widening of mortality differentials in relative terms. For instance, the ASDR for Circulatory causes among men in the most deprived quintile passed from being 1.37 times the ASDR of men in the least deprived quintile in 1981, to being 1.74 times in 2005. In addition, if this widening were to continue at the same pace observed between 1981 and 2007, the ratio between the ASDR from Circulatory causes among men in the

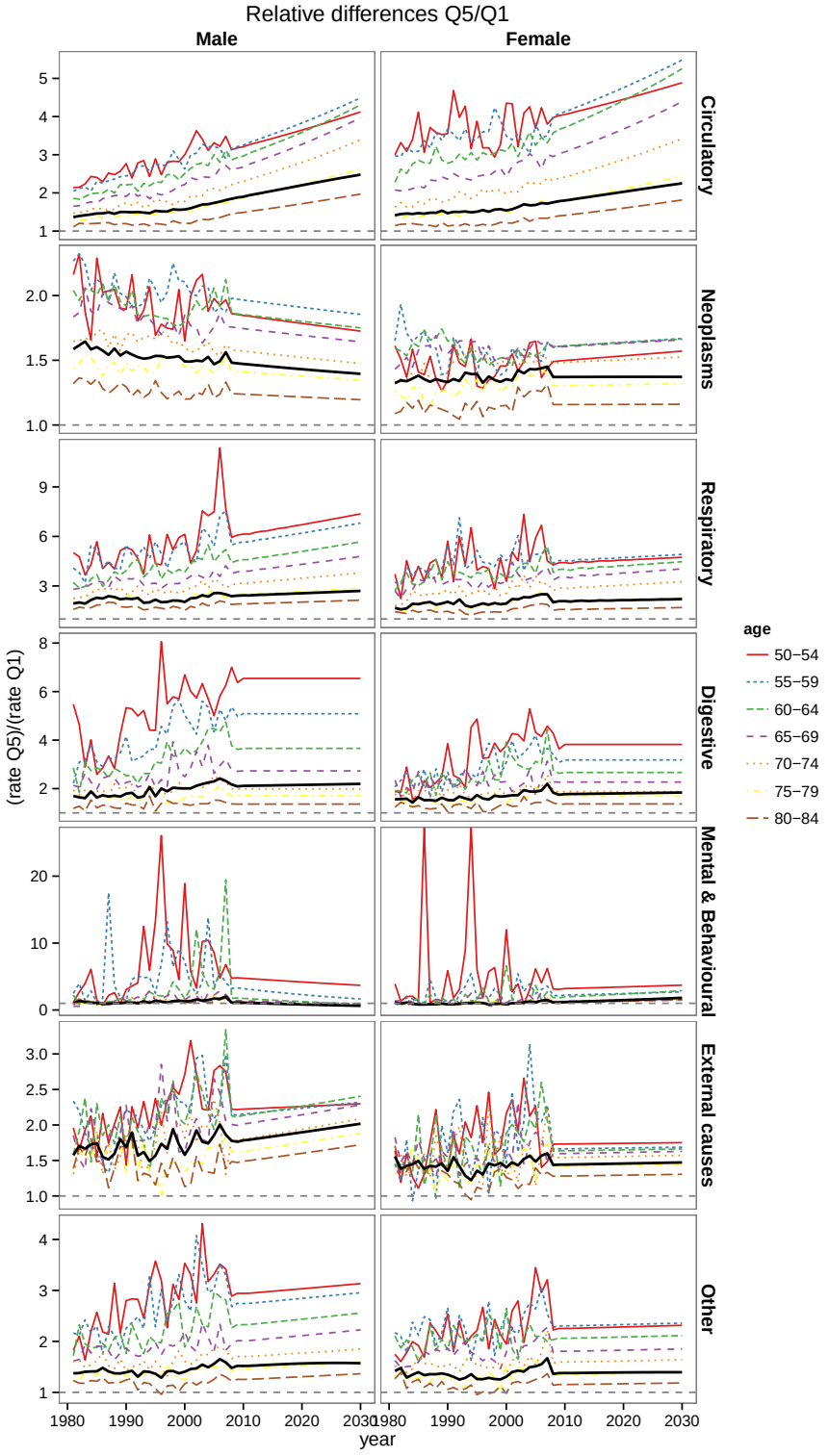


FIGURE 3.20: Time series of cause-specific relative mortality differentials between the most and least deprived quintiles of the English population. Values prior to 2007 are observed differentials and values post 2007 are projected differentials. The thicker black lines represent differentials in age standardised death rates.

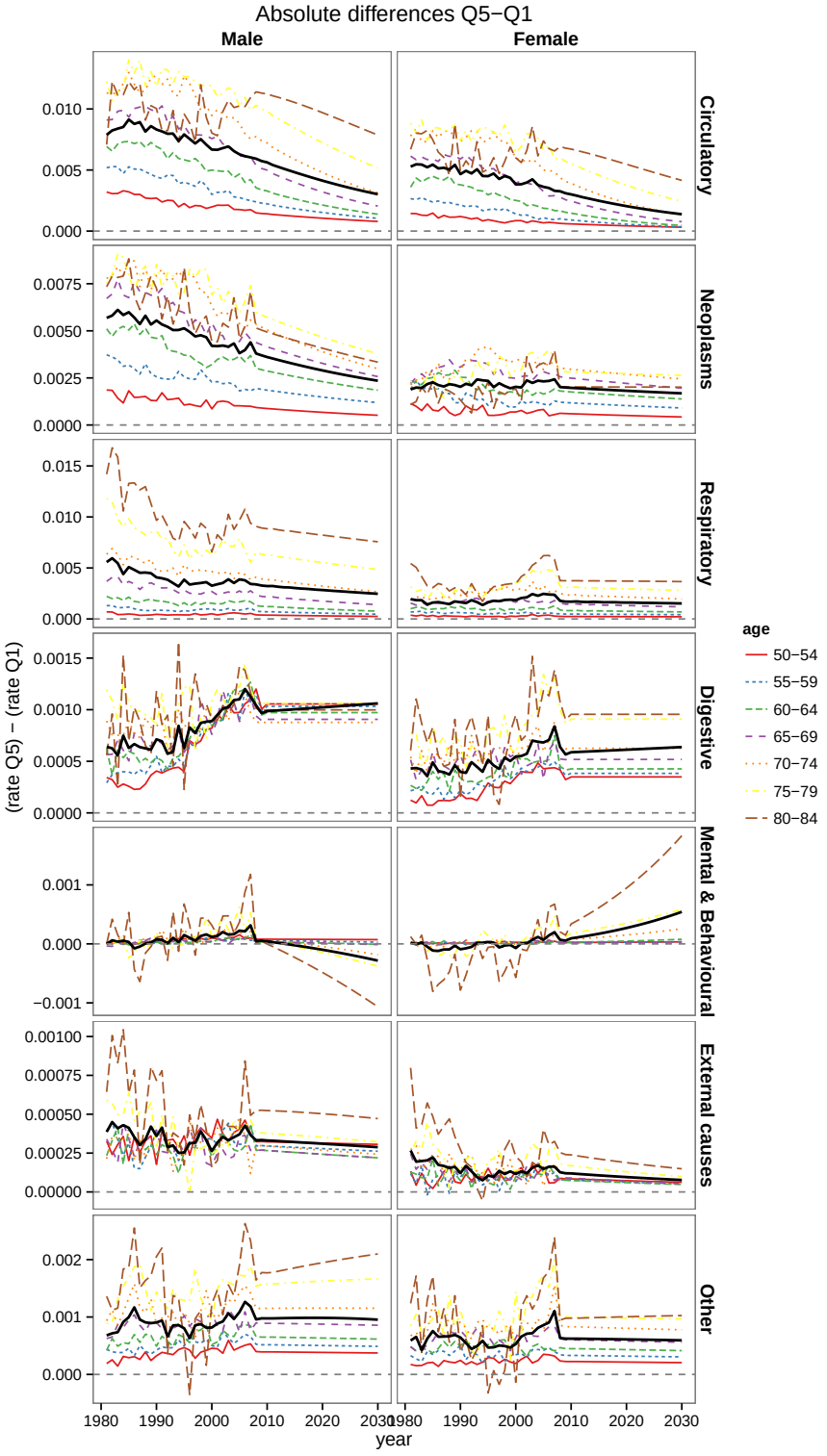


FIGURE 3.21: Time series of cause-specific absolute mortality differentials between the most and least deprived quintiles of the English population. Values prior to 2007 are observed differentials and values post 2007 are projected differentials. The thicker black lines represent differentials in age standardised death rates.

TABLE 3.6: Ratio of age standardised mortality rates between the most and least deprived populations for the different causes of death. Values up to 2005 are observed ratios and post 2005 are projected ratios.

Men

	1981	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030
Circulatory	1.37	1.46	1.50	1.53	1.57	1.74	1.90	2.05	2.20	2.34	2.48
Neoplasms	1.59	1.60	1.57	1.53	1.49	1.47	1.47	1.45	1.43	1.41	1.40
Respiratory	1.93	2.28	2.24	2.17	2.03	2.57	2.43	2.49	2.56	2.63	2.70
Digestive	1.70	1.63	1.79	1.68	2.01	2.30	2.11	2.13	2.15	2.17	2.20
Mental & Behavioural	0.99	1.05	1.12	1.38	1.31	1.71	1.10	0.95	0.83	0.72	0.63
External causes	1.57	1.74	1.68	1.57	1.58	1.86	1.78	1.84	1.90	1.96	2.02
Other	1.37	1.41	1.39	1.36	1.40	1.56	1.52	1.54	1.57	1.58	1.57

Women

	1981	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030
Circulatory	1.41	1.47	1.47	1.51	1.54	1.68	1.80	1.91	2.03	2.14	2.25
Neoplasms	1.32	1.38	1.33	1.39	1.35	1.43	1.37	1.37	1.37	1.37	1.37
Respiratory	1.67	1.90	2.04	1.83	1.93	2.42	2.08	2.10	2.14	2.17	2.20
Digestive	1.56	1.64	1.53	1.73	1.69	1.88	1.77	1.79	1.81	1.82	1.84
Mental & Behavioural	1.12	0.83	0.86	1.09	0.85	1.17	1.19	1.32	1.47	1.64	1.82
External causes	1.56	1.49	1.35	1.36	1.40	1.48	1.44	1.45	1.46	1.46	1.47
Other	1.42	1.40	1.35	1.26	1.31	1.51	1.38	1.38	1.39	1.39	1.39

TABLE 3.7: Absolute difference in age standardised mortality rates between the most and least deprived populations for the different causes of death. Values are reported in deaths per 1000 persons. Values up to 2005 are observed differences and post 2005 are projected differences..

Men

	1981	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030
Circulatory	7.88	9.14	8.32	7.92	6.68	6.28	5.63	4.81	4.13	3.53	3.02
Neoplasms	5.69	5.99	5.55	5.14	4.20	3.81	3.61	3.23	2.90	2.60	2.35
Respiratory	5.57	5.08	4.08	3.85	3.26	3.88	3.22	3.03	2.83	2.64	2.47
Digestive	0.64	0.63	0.71	0.63	0.89	1.10	0.99	1.00	1.02	1.04	1.06
Mental & Behavioural	-0.00	0.03	0.08	0.15	0.12	0.22	0.04	-0.02	-0.10	-0.19	-0.28
External causes	0.39	0.41	0.36	0.25	0.26	0.37	0.33	0.32	0.31	0.30	0.29
Other	0.68	1.00	0.90	0.80	0.82	1.09	0.98	0.98	0.98	0.98	0.95

Women

	1981	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030
Circulatory	5.28	5.44	4.66	4.50	3.82	3.62	3.01	2.48	2.04	1.68	1.38
Neoplasms	1.89	2.27	2.03	2.34	2.03	2.32	1.97	1.89	1.82	1.75	1.68
Respiratory	1.94	1.61	1.64	1.63	1.85	2.46	1.69	1.66	1.60	1.57	1.52
Digestive	0.43	0.49	0.40	0.51	0.55	0.68	0.59	0.60	0.61	0.62	0.64
Mental & Behavioural	0.02	-0.11	-0.09	0.03	-0.07	0.07	0.09	0.17	0.27	0.39	0.54
External causes	0.27	0.22	0.12	0.11	0.12	0.15	0.12	0.11	0.10	0.09	0.08
Other	0.59	0.76	0.64	0.47	0.53	0.86	0.62	0.62	0.61	0.60	0.59

most deprived quintile and the least deprived quintile could reach a value of 2.48 in 2030.

- Second, Neoplasms show a decrease of mortality differentials among men both in absolute and relative terms and a fairly constant evolution of both relative and absolute differentials among women. For example, as a result of the faster mortality improvements in cancer mortality seen by the the most deprived quintile of the English male population, the relative difference in the ASDR among men for this

cause decreased from 1.59 in 1981 to 1.47 in 2005 and is forecasted to continue reducing to 1.4 in 2030.

- Third, Digestive and "Other" causes of death experienced a widening of differentials both in absolute and relative terms over the 1992-2007 period, but with these differentials forecasted to stay constant around their 2010 levels. For Digestive causes in particular, the increase in differentials was very significant. For instance, at age 60-64 the ratio of mortality from Digestive causes among the men in the most and least deprived areas of England increased steadily from 2.23 in 1992 to a maximum of 4.74 in 2007 and is forecasted to stay around 3.66 from 2010 onwards.

- Finally, Mental and Behavioural diseases and External causes show very erratic historical trends in differentials, partly due to the fact that these two causes of death account for the smallest proportion of the total number of deaths for people over age 50.

3.4 Discussion

In this chapter we have introduced new Lee-Carter based techniques for the modelling of mortality by cause of death and socio-economic stratification. An application to mortality data for deprivation subpopulations in England showed that these modelling techniques offer a simple, yet effective, approach for the assessment of the magnitude of historical mortality differentials for the main causes of death and for the projection of the possible future evolution of these differentials. Key to the success of our approach was the introduction of new modelling tools to control for data production changes in cause of death data which, if not taken into account appropriately, could hinder the validity of any analysis of trends in cause-specific mortality.

The empirical part of this chapter revealed a clear association between area deprivation and mortality rates, with people living in more deprived areas of England consistently having higher mortality rates for the leading causes of death than those living in less deprived areas. The mortality differentials shown in this study are substantial although with varying magnitude across causes, sexes and ages. Circulatory, Respiratory and Digestive causes of death showed the highest differences over the 1981-2007 period of study, with mortality in the most deprived quintile for some ages and causes of death

being more than six times the mortality rates of the least deprived quintile. Neoplasms and External causes, by contrast, showed the smallest relative mortality differences.

In Chapter 2 it was found that all-cause relative mortality differentials in England by deprivation widened over the 1981-2007 period, mainly due to the slower mortality improvements experienced by the most deprived socio-economic subgroups. However, this present study shows a variety of patterns in the evolution of relative differentials by cause of death. While relative differentials for Circulatory, Respiratory, Digestive and "Other" causes widened for both sexes over the 1981-2007 period, differentials for Neoplasms narrowed slightly for men and remained mainly unchanged for women. Moreover, whereas the increase of relative differentials for Circulatory, Respiratory and "Other" causes was associated with the faster improvement of mortality amongst the least deprived quintiles, for Digestive causes the widening of differentials was due to a significant increase of mortality amongst the most deprived quintiles of England. Digestive causes, however, offer the greatest potential for reduction in mortality differentials and greater attention needs to be given to alcohol consumption which is a predominant factor in the mortality increase seen for this cause (Ashton et al., 2010).

With the noticeable exception of Circulatory diseases, we have also found that mortality differentials are higher among men than among women. Furthermore, clearer trend mortality differentials were observed in the male population. These results are consistent with previous studies using area-based deprivation measures in the UK (Romeri et al., 2006; Villegas and Haberman, 2014) and other developed countries (Singh, 2003; Turrell and Mathers, 2001; Windenberger et al., 2011) and with other UK (White et al., 2003) and international (Saurel-Cubizolles et al., 2009) studies using alternative markers of socio-economic conditions (e.g. occupation and educational attainment).

There are a number of possible limitations with the modelling approach introduced in this chapter. The first and most important one is that, both in the fitting and forecasting of the models, we have assumed independence among the different causes of death. This is clearly an unrealistic assumption as, for instance, the same risk factors can affect several causes at the same time. An alternative approach to allow for the dependence between the causes of death is offered by Vector Autoregression (VAR) and Vector Error Correction Models (VECMs) as suggested by Arnold (-Gaille) and Sherris (2013, 2015). Hence, one could use such multivariate time series techniques to model and forecast

simultaneously the time series indices, κ_t, for the different causes of death, as opposed to the independent univariate ARIMA models used in this chapter.

Second, our modelling approach is an age-period model which assumes that the mortality evolution for each cause of death is driven exclusively by a single period index, κ_t. Although this has proven to be sufficient for the restricted age range we have considered in this study, additional period factors may be required if looking at a wider age range encompassing younger ages whose mortality change may be driven by different risk factors. Furthermore, some causes of death characterised by clear cohort effects may require the inclusion of a cohort term and could be better modelled using an age-period-cohort approach (Di Cesare and Murphy, 2009). We note, however, that in spite of the lack of an explicit cohort term, the Lee-Carter model with coding adjustment showed a reasonable goodness-of-fit and forecasting performance for Neoplasms and Circulatory diseases, causes of death for which Willets (2004) has reported clear cohort effects among the England and Wales population. The surprising good performance of the Lee-Carter with code adjustments when applied to these two groups of causes is explained by an unexpected side effect of the inclusion of the coding adjustments parameters, $\delta_x^{(i)}$, which seem to be partially capturing some of the cohort effects associated to these causes of death. These encouraging results also suggests that the piece-wise technique we have introduced to account for coding adjustments may be useful in other settings not related to cause-specific mortality, such as when modelling all-cause mortality in populations which have undergone structural changes in their mortality trend (see e.g. Renshaw and Haberman (2003) and van Berkum et al. (2014)).

Finally, the Three-way Lee-Carter model used to capture socio-economic differences in mortality assumes, for the sake of parsimony, that mortality trend differentials are constant across ages. This is despite the fact that trend differentials may show important variations by age as reported, for instance, in Bajekal et al. (2013) for coronary heart diseases. Hence, a subject of further investigation is the development of modelling approaches which allow for a more complex age-structure of mortality trend differentials.

We also recognise additional limitations of our study stemming from the dataset used in the analysis of socio-economic difference in cause-specific mortality in England. First, we have used an ecological measure of socio-economic deprivation instead of individual level socio-economic measures. This leads to the usual problem of any ecological study

whereby outcomes obtained at the group level do not necessarily apply at the individual level (Greenland, 2001). This ecological design also limits the ability of associating a causal pathway between socio-economic conditions and mortality, for which individual level information would be required.

Second, we note that our data measure mortality rates for those living in a particular area at a particular point in time, ignoring any past exposure to risk factors whilst living in other areas. It is plausible that healthier people will tend to migrate from more deprived areas to less deprived ones and that less healthy people are less likely to move, resulting in a potential bias towards higher mortality inequalities (Norman et al., 2005).

Finally, our analysis is based on quintile classification of LSOAs using the IMD 2007 which refers to area indicators in 2005. Thus, we have assumed time stability in the membership of the LSOAs in the deprivation subgroups. However, Norman (2009) and Lu et al. (2014) have shown that the majority of small areas in England have remained in the same deprivation quintile over the past three decades.

Acknowledgements

We are grateful for comments received at the Insurance: Mathematics and Economics Conference in Copenhagen, Denmark in 2013; at the Actuarial Research Conference in Philadelphia, USA in 2013; at the First Congress on Actuarial Science and Quantitative Finance in Bogotá, Colombia in 2014; and at the Institute and Faculty of Actuaries Mortality and Longevity Symposium in Birmingham, UK in 2014. We thank Paul Norman from University of Leeds who compiled the deprivation-specific population denominators used in this study. We would also like to thank Andrew Hunt for numerous discussions on the identifiability of mortality models, which were central to the development of some of the the techniques introduced in this chapter.

Bibliography

Ahmad, O. B., Boschi-pinto, C., Lopez, A. D., Murray, C. J., Lozano, R., Mie, I., 2001. Age standardization of rate : A new who standard (GPE Discussion Paper Series: No . 31 EIP/GPE/EBD).

Arnold (-Gaille), S., Sherris, M., 2013. Forecasting Mortality Trends Allowing for Cause-of-Death Mortality Dependence. North American Actuarial Journal 17 (4), 273–282.

Arnold (-Gaille), S., Sherris, M., 2015. Causes-of-death mortality: What do we know on their dependence? North American Actuarial Journal.

Ashton, C., Bajekal, M., Raine, R., 2010. Quantifying the contribution of leading causes of death to mortality decline among older people in England, 1991-2005. Health Statistics Quarterly (45), 100–127.

Bajekal, M., Scholes, S., O'Flaherty, M., Raine, R., Norman, P., Capewell, S., 2013. Unequal trends in coronary heart disease mortality by socioeconomic circumstances, England 1982-2006: An analytical study. PLoS ONE 8 (3).

Brock, A., Griffiths, C., Rooney, C., 2006. The impact of introducing ICD-10 on analysis of respiratory mortality trends in England and Wales. Health Statistics Quarterly (29), 9–17.

Continuous Mortality Investigation, 2004. Working Paper 3 – Projecting future mortality: A discussion paper.

Cook, L., Griffiths, C., Rooney, C., 2002. The implementation of ICD-10 for cause of death coding – some preliminary results from the bridge coding study. Health Statistics Quarterly (13), 31–41.

Di Cesare, M., Murphy, M. J., 2009. Forecasting mortality, different approaches for different cause of deaths? The cases of lung cancer; influenza, pneumonia, and bronchitis; and motor vehicle accidents. British Actuarial Journal 15 (Supplement), 185–211.

GAD, 2001. National population projections : Review of methodology for projecting mortality. Tech. Rep. 8, Government Actuary's Department.

Goldacre, M. J., Duncan, M. E., Cook-Mozaffari, P., Griffith, M., Sep. 2003. Trends in mortality rates comparing underlying-cause and multiple-cause coding in an English population 1979-1998. Journal of Public Health 25 (3), 249–253.

Greenland, S., 2001. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. International Journal of Epidemiology 30 (6), 1343–50.

Griffiths, C., Rooney, C., 2006. Trends in mortality from Alzheimer's disease, Parkinson's disease and dementia, England and Wales, 1979-2004. Health Statistics Quarterly.

Johnson, B., 2011. Deriving trends in life expectancy by the National Statistics Socio-economic Classification using the ONS Longitudinal Study. Health Statistics Quarterly (49), 9–52.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Lu, J. L. C., Wong, W., Bajekal, M., 2014. Mortality improvement by socio-economic circumstances in England (1982 to 2006). British Actuarial Journal 19 (1), 1–35.

Mayhew, L., Smith, D., 2014. Gender convergence in human survival and the postponement of death. North American Actuarial Journal 18 (1), 194–216.

Meyricke, R., Sherris, M., 2013. The determinants of mortality heterogeneity and implications for pricing annuities. Insurance: Mathematics and Economics 53 (2), 379–387.

Moriyama, I. M., Loy, R. M., Robb-Smith, A. H., 2011. History of the statistical classification of diseases and causes of death. National Center for Health Statistics, Hyattsville, MD.

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Exley, S., Barnes, H., Dibben, C., 2007. The English indices of deprivation 2007. Department of Communities and Local Government, London.

Norman, P., 2009. Identifying change over time in small area socio-economic deprivation. Applied Spatial Analysis and Policy 3 (2-3), 107–138.

Norman, P., Boyle, P., Rees, P., 2005. Selective migration, health and deprivation: a longitudinal analysis. Social Science and Medicine 60 (12), 2755–71.

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

Rey, G., Aouba, A., Pavillon, G., Hoffmann, R., Plug, I., Westerling, R., Jougla, E., Mackenbach, J., 2011. Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes. Population Health Metrics 9 (52), 1–11.

Richards, S., 2009. Selected issues in modelling mortality by cause and in small populations. British Actuarial Journal 15 (Supplement 1), 267–283.

Ridsdale, B., Gallop, A., 2010. Mortality by cause of death and by socio-economic and demographic stratification 2010. In: International Congress of Actuaries 2010.

Romeri, E., Baker, A., Griffiths, C., 2006. Mortality by deprivation and cause of death in 1999-2003. Health Statistics Quarterly 32, 19–34.

Rooney, C., Devis, T., 1996. Mortality trends by cause of death in England and Wales 1980-94: the impact of introducing automated cause coding and related changes in 1993. Population Trends (86), 29–35.

Russolillo, M., Giordano, G., Haberman, S., 2011. Extending the Lee-Carter model: a three-way decomposition. Scandinavian Actuarial Journal (2), 96–117.

Saurel-Cubizolles, M.-J., Chastang, J.-F., Menvielle, G., Leclerc, A., Luce, D., 2009. Social inequalities in mortality by cause among men and women in France. Journal of Epidemiology and Community Health 63 (3), 197–202.

Singh, G. K., 2003. Area deprivation and widening inequalities in US mortality, 1969-1998. American Journal of Public Health 93 (7), 1137–1143.

The Marmot Review, 2010. Fair society, healthy lives: Strategic review of health inequalities in England post 2010.

Turrell, G., Mathers, C., 2001. Socioeconomic inequalities in all-cause and specific-cause mortality in Australia: 1985-1987 and 1995-1997. International Journal of Epidemiology 30 (2), 231–239.

van Berkum, F., Antonio, K., Vellekoop, M., 2014. The impact of multiple structural changes on mortality predictions. Scandinavian Actuarial Journal.

Villegas, A. M., Haberman, S., 2014. On the modeling and forecasting of socioeconomic mortality differentials: an application to deprivation and mortality in England. North American Actuarial Journal 18 (1), 168–193.

White, C., van Galen, F., Chow, Y. H., 2003. Trends in social class differences in mortality by cause, 1986 to 2000. Health Statistics Quarterly (20), 25–37.

Willets, R., 2004. The cohort effect: Insights and explanations. British Actuarial Journal 10 (4), 833–877.

Windenberger, F., Rican, S., Jougla, E., Rey, G., 2011. Spatiotemporal association between deprivation and mortality: trends in France during the nineties. European Journal of Public Health, 1–7.

Wood, S., 2015. Package mgcv.
URL http://cran.r-project.org/web/packages/mgcv/index.html

Part II

Longevity Basis Risk

4

A Comparative Study of Two-population Models for the Assessment of Basis Risk in Longevity Hedges

The majority of this chapter has been carried out within the framework of a joint research project between Cass Business School and Hymans Robertson LLP. This project was commissioned by the Longevity Basis Risk Working Group of the Institute & Faculty of Actuaries and the Life & Longevity Markets Association with the objective of developing a practical methodology for assessing basis risk for longevity transactions. Therefore, part of this chapter draws from the report documenting this research. The full reference of this report follows:

– Haberman, S., Kaishev, V. K., Millossovich, P., Villegas, A. M., Baxter, S., Gaches, A., Gunnlaugsson, S., Sison, M., 2014. Longevity Basis Risk: A methodology for assessing basis risk. Institute and Faculty of Actuaries Sessional Research Paper.

Previous versions of this chapter were presented at the following conferences and events:

- November 2013. Life Conference and Exhibition, Edinburgh, UK. "Towards an industry standard to assess longevity basis risk".
- February 2014. ACA Members' Conference, London, UK. "Developing new longevity solutions: understanding longevity basis risk".
- September 2014. International Mortality and Longevity Symposium, Birmingham, UK. "Towards an industry standard to assess Longevity Basis Risk".
- December 2014. Institute and Faculty of Actuaries Sessional Research Event, London, UK. "Longevity Basis Risk: A methodology for assessing basis risk". Video available at http://openchannel.multichanneltv.com/the-actuarial-profession/longevity-basis-risk-methodology/platform.php?noCPD
- January 2015. Perspectives on Actuarial Risks in Talks of Young Researchers, Liverpool, UK. "A comparative study of two-population models for the assessment of basis risk in longevity hedges".

A Comparative Study of Two-population Models for the Assessment of Basis Risk in Longevity Hedges

Andrés M. Villegas, Steven Haberman, Vladimir Kaishev, Pietro Millossovich

Cass Business School, City University London, United Kingdom

Abstract

Longevity swaps have been one of the major success stories of pension scheme de-risking in recent years. However, with some few exceptions, all of the transactions to date have been bespoke longevity swaps based upon the mortality experience of a portfolio of named lives. In order for this market to start to meet its true potential, solutions will ultimately be needed that provide protection for all types of members, are cost effective for large and smaller schemes, are tradable, and enable access to the wider capital markets. Index-based solutions have the potential to meet this need; however concerns remain with these solutions. In particular, the lack of a well-established approach for assessing the basis risk emerging from the potential mismatch between the underlying forces of mortality for the index reference portfolio and the pension fund/annuity book being hedged is the stand out issue that has, to date, prevented many schemes progressing their consideration of index-based solutions. Two-population stochastic mortality models offer an alternative to overcome this obstacle as they allow market participants to compare and project the mortality experience for the reference and target populations and thus assess the amount of demographic basis risk involved in an index-based longevity hedge. In this chapter, we systematically assess the suitability of several multi-population stochastic mortality models for assessing basis risks and provide guidelines on how to use these models in practical situations, paying particular attention to the data requirements for the appropriate calibration and forecasting of such models.

Keywords: Two-population model; basis risk; mortality-linked derivatives

4.1 Introduction

Recent years have seen a huge growth in longevity risk transfer, both in the insurer to reinsurer market, and from pension schemes to the insurance market. For example during 2014 in the UK £36.6bn of longevity risk liabilities were transferred from pension schemes to insurers and reinsurers via buy-ins, buy-outs and longevity swaps. Of this, £25.4bn related to longevity only transactions (longevity swaps), more than double the volume written in the preceding 3 years (Hymans Robertson LLP, 2015). An effective, growing market with sufficient capacity to meet demand would be to the benefit of all participants, whether to enable business to be done, or to manage risk.

To date most transactions have been "bespoke" deals, with the payouts linked directly to the actual experience or lifespans of the individuals being covered. But index-based solutions – where the payouts are linked to a longevity index or metric based on an external reference population – are possible. They have the potential to provide important benefits: lower costs, faster execution, potential for liquidity, and greater transparency.

In its simplest form an index based longevity swap involves a payment to the pension scheme or insurer that is based on the longevity experience of a reference index. An index-based swap provides a means to obtain (partial) protection from longevity risk both for pensioners but also deferred pensioners who are generally not covered by the "bespoke" transactions. In the case of life insurers they offer a potentially flexible way to manage exposure to longevity risk, or to facilitate a more capitally optimal balance between longevity and mortality risk. However, index-based swaps do not provide a perfect risk reduction due to the presence of basis risk, which arises from the differences in the mortality experiences of the reference population of the index and of the target population being hedged. As a result, the index based payments will not exactly match the actual annuity payments being made by the insurer or pension scheme.

There are three primary sources of basis risk driving the mismatch between the insurer or pension scheme liabilities and the longevity index hedge (LLMA, 2012):

- *Structuring risk* due to the payoff of the hedging instruments being different to that of the portfolio: for example the hedging instrument making annual payments

whereas the portfolio pays annuities or pensions monthly, or the hedge may be of shorter duration than the liabilities.

- *Sampling risk* arising from the random outcomes of the individual lives within the portfolio and the index population meaning the actual mortality experienced by the two populations will not be the same, other than by chance.

- *Demographic risk* owing to demographic and socio-economic differences in the composition of the actual portfolio being hedged and the index population referenced in the hedge, leading to different underlying mortality rates now – and in the future.

Well-established approaches for modelling the first two of these exist. Structuring risk can be assessed by simulating the cashflows under the portfolio and the payoffs under the instrument, whilst sampling risk can be modelled by simulating the outcomes for the respective populations.

In contrast there is no well-established approach for assessing demographic basis risk. Yet it is this risk which worries (re)insurers and pension schemes when they consider entering index-based longevity transactions (LLMA, 2012). The absence of an appropriate approach for quantifying such risk makes it very difficult to assess whether such a transaction looks good value for money, or what impact the transaction would have on the insurer's or pension scheme's overall risk profile and hence capital/funding requirements.

In the academic literature there have been a few contributions setting out possible approaches for quantifying longevity basis risk. Coughlan et al. (2011) propose a comprehensive framework for assessing the effectiveness of a longevity hedge, in which the first and key step entails a careful analysis of the historical experiences of the reference and target population to get an informed understanding of the mortality differences between the two populations. Li and Hardy (2011) investigate the use of a number of multipopulation extension of the Lee-Carter model (Lee and Carter, 1992) for the assessment of basis risk and use the Augmented Common Factor model of Li and Lee (2005) to quantify the hedge effectiveness of an index-based q-forward longevity hedge. Li et al. (2015) propose a systematic approach for the construction of two-population mortality models that can be used for the quantification of the population basis risk in a standardised longevity hedge. In addition, recent years have seen a boom in the actuarial and

demographic literature looking at the modelling of mortality in two (or more) related populations (e.g. Li and Lee (2005); Jarner and Kryger (2011); Plat (2009b); Cairns et al. (2011a); Dowd et al. (2011)). These two-population models, although not always proposed with the specific aim of assessing longevity basis risk, have the potential for allowing market participants to compare and project the mortality experience for the reference and target populations and thus assess the amount of demographic basis risk involved in an index-based longevity hedge. However, often the portfolio experience data will be sparse posing a challenge for the accurate calibration and projection of the two-population model.

In light of the increasing attention that two(multi)-population mortality models are receiving and their potential for assessing the basis risk and hedge effectiveness of standardised longevity transactions, in this chapter we aim to systematically assess the suitability for assessing basis risks of existing multipopulation mortality models and provide guidelines on the use of these models in practical situations. The results from these investigations set the theoretical grounds for the longevity basis risk assessment methodology described in Haberman et al. (2014b,a).

In this chapter we proceed as follows. In Section 4.2 we introduce our notation. In Section 4.3 we provide an overview of the multipopulation mortality models that have been proposed in the literature. Then, to facilitate the comparison of models, we discuss in Section 4.4 a general modelling framework under which most two-population mortality models can be accommodated. In Section 4.5 we draw from the literature comparing single population mortality models to introduce a number of criteria that a good and practical two-population model for basis risk assessment should satisfy. We use these criteria in Section 4.6 to systematically evaluate the appropriateness of the possible two-population models for basis risk assessment. First, in Section 4.6.1, we evaluate the models against those criteria which are theoretical properties of a model and can be evaluated without reference to a specific dataset. Then, in Section 4.6.2, we focus on those criteria which can only be evaluated after a model has been fitted to data. This systematic evaluation of the models will allow us to identify the main features of a good model for basis risk assessment and discuss the data requirements for the appropriate calibration and forecasting of such a model. Having identified some reasonable models for basis risk assessment, we examine in Section 4.7 the performance of these model in some illustrative hedge-effectiveness evaluation exercises, paying particular attention

to the impact that different volumes of data may have on the assessment of basis risk. Finally, we conclude in Section 4.8 with a discussion of our main findings and future areas of research.

4.2 Notation

We denote by R the reference population backing the hedging instrument and by B the book population whose longevity risk is to be hedged. We assume that the following data are available. For the reference population,

- D_{xt}^R number of deaths aged x last birthday in calendar year t.
- E_{xt}^R initial exposed to risk for age x and calendar year t.

The corresponding 1-year death rate for an individual in the reference population aged x and in calendar year t, denoted q_{xt}^R, can be estimated as $\hat{q}_{xt}^R = D_{xt}^R \big/ E_{xt}^R$. Similarly, the corresponding quantities for the book population are denoted D_{xt}^B, E_{xt}^B and $\hat{q}_{xt}^B = D_{xt}^B \big/ E_{xt}^B$. We assume that these data are available for a given set of ages and given numbers of years that can differ in the reference and the book. More precisely, we assume that D_{xt}^R, E_{xt}^R are available for consecutive ages $x = x_1^R, \ldots, x_{k_R}^R$ and consecutive calendar years $t = t_1^R, \ldots, t_{n_R}^R$, while in the book they are available for ages $x = x_1^B, \ldots, x_{k_B}^B$ and calendar years $t = t_1^B, \ldots, t_{n_B}^B$. Typically, data for the reference population will be available over a longer horizon than in the book (i.e. $t_1^R \leq t_1^B$, $t_{n_R}^R \geq t_{n_B}^B$) and the ages available within the book may be a subset of those available in the reference (i.e. $x_1^R \leq x_1^B$, $x_{k_R}^R \geq x_{k_B}^B$).

4.3 Overview of available two-population mortality models

In order to be able to assess basis risk, we need a model that is able to capture the mortality trends in the reference population backing the hedging instrument and in the book population whose risk is to be hedged. That is to find a suitable two-population model for q_{xt}^R and q_{xt}^B which produces consistent stochastic forecasts of future mortality.

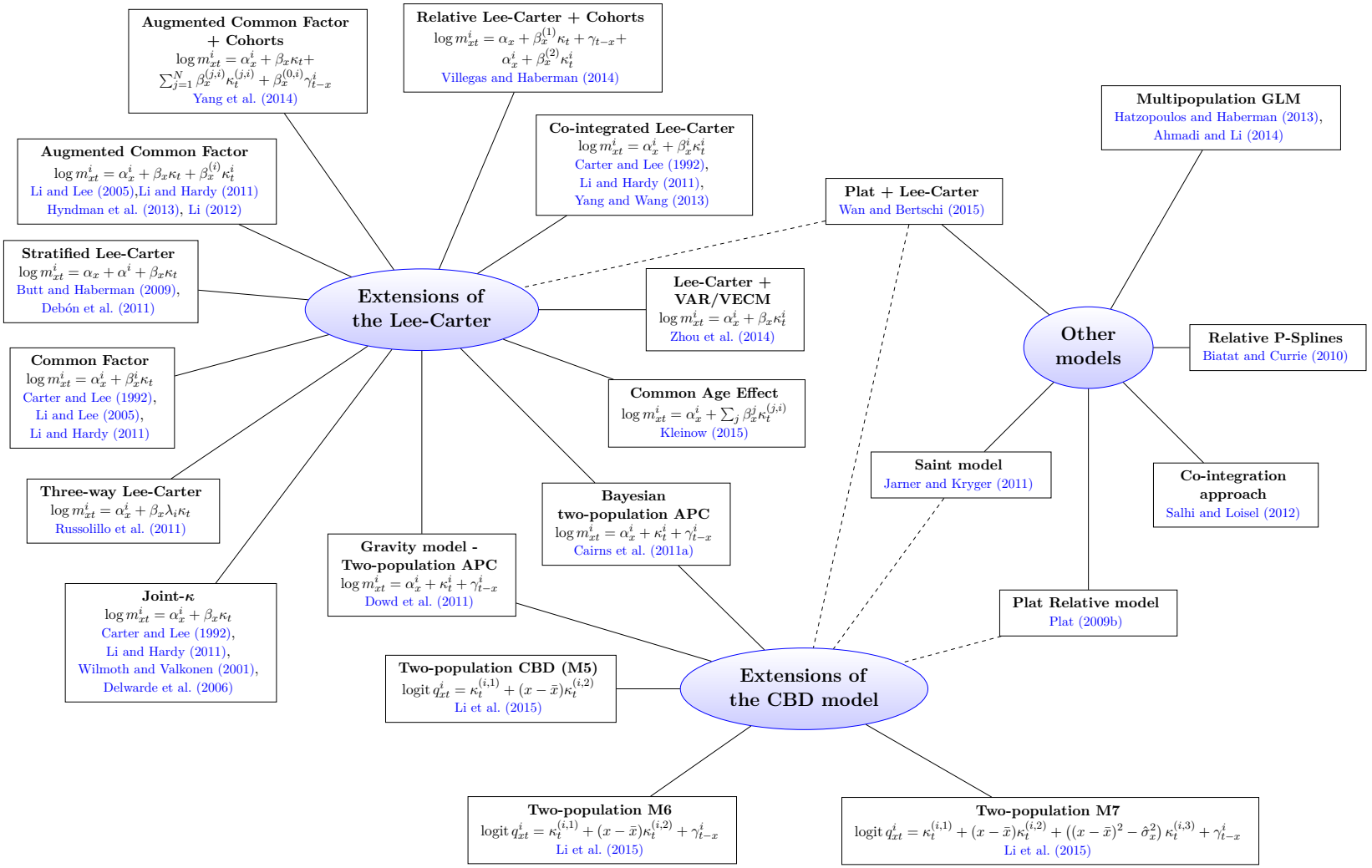


FIGURE 4.1: Overview of the multipopulation mortality modelling literature.

Many models have been proposed in the literature to represent the mortality evolution of two or more related populations. The majority of such models extend known single population models by specifying the correlation and interaction between the involved populations. Figure 4.1 contains a schematic representation of the multi-population models currently available in the published literature, broadly grouped according to three main categories concerning the single population model they extend.

Although most of the academic contributions to the modelling of multi-populations are fairly recent, the first ideas go back to the seminal paper by Carter and Lee (1992), which suggested three possible ways of extending their single population model (Lee and Carter, 1992) in order to forecast differentials in US mortality between men and women. The first and simplest approach suggested by Carter and Lee (1992) is to use independent Lee-Carter models for each population, and, if desired, to study in a later stage the dependence between the population-specific period effects. A second approach, the Joint-κ model, assumes that a single period component κ_t drives the mortality change for all the populations. The third approach estimates the populations jointly using co-integration techniques.

Under the Joint-κ model the overall level of mortality is assumed to be common to all of the populations, but the age-specific mortality pattern and the age-specific responses to changes in the level of mortality are population-specific. Formally, the Joint-κ model assumes that the central death rate at time t for age x in population i, m^i_{xt}, is given by

$$\log m^i_{xt} = \alpha^i_x + \beta^i_x \kappa_t \tag{4.1}$$

Delwarde et al. (2006) have implemented a slightly modified version of this model in the analysis of the patterns of mortality decline in five developed countries, while Wilmoth and Valkonen (2001) have considered a similar model in order to describe mortality differentials across socio-economic groups within a country.

Russolillo et al. (2011) have proposed a three-way extension of the Lee-Carter model to deal with mortality data disaggregated according to a third criterion (e.g. gender, country or ethnicity) besides age and time. The proposed model, which can be thought of as a restricted version of the Joint-κ model, is given by:

$$\log m^i_{xt} = \alpha^i_x + \beta_x \lambda^i \kappa_t \tag{4.2}$$

where λ^i captures the variability in mortality improvements across populations.

A related modification of the Lee-Carter model was introduced by Li and Lee (2005) in the so-called Common Factor model. As compared to the Joint-κ model, the Common Factor model assumes that $\beta_x^i = \beta_x$, that is

$$\log m_{xt}^i = \alpha_x^i + \beta_x \kappa_t \tag{4.3}$$

Butt and Haberman (2009) consider a restricted version of (4.3), named the stratified Lee-Carter model, in which the population-specific effects are additive and constant across ages, i.e., $\alpha_x^i = \alpha_x + \alpha^i$ for all x and all i. This same structure has been employed successfully by Debón et al. (2011) in the modelling of mortality in Spanish regions.

The structure of the Common Factor model implies the same mortality improvements for all population at all times. However, this is an unrealistic assumption for most datasets. To tackle this problem, Li and Lee (2005) add a population specific factor to equation (4.3), obtaining the Augmented Common Factor model:

$$\log m_{xt}^i = \alpha_x^i + \beta_x \kappa_t + \beta_x^i \kappa_t^i$$

where the term $\beta_x^i \kappa_t^i$ captures the deviations of the rate of mortality change of population i from the long-term trend in mortality change implied by the Common Factor, $\beta_x \kappa_t$. In order to avoid divergence in the projected mortality mortality the κ_t^i factors are assumed to be stationary processes such as a first order autoregressive process (AR(1)). Under this assumption the mortality rates of the different populations may wander apart in the short and medium terms, but tend to converge in the long-run. The Augmented Common Factor has spawned several variants and extensions. Hyndman et al. (2013) have introduced the product-ratio method which extends the Augmented Common Factor model by adopting a functional data approach and allowing more than one period index for the modelling of both the common factor and of the population-specific factors. Li (2012), who also considers multiple period indexes, uses a Poisson setting to estimate the parameters of the Augmented Common Factor models instead of the singular value decomposition approach originally employed by Li and Lee (2005). Recently, Yang et al.

(2014) have extended the Poisson Augmented Common Factor to allow for possible cohort effects. Villegas and Haberman (2014) have considered a similar cohort variant of the Augmented Common Factor for the purpose of studying socio-economic differences in mortality.

Li and Hardy (2011) have investigated the performance of several multipopulation extension of the Lee-Carter model for the purpose of measuring the basis risk implied in standardised longevity hedges. Specifically, they compare the independent modelling approach, the Joint-κ model, the co-integrated Lee-Carter, the Common Factor model, and the Augmented Common Factor model. In an application to data from the female populations of Canada and US, Li and Hardy (2011) conclude that the Augmented Common Factor model has the best performance in terms of historical goodness of fit and forecasting performance.

An important contribution of Li and Hardy (2011) is the elaboration of the use of co-integration as suggested by Carter and Lee (1992). In the two-population version of the co-integrated Lee-Carter model two independent single population Lee-Carter models are first fitted to each of the populations,

$$\log m_{xt}^i = \alpha_x^i + \beta_x^i \kappa_t^i, \quad i = 1, 2, \tag{4.4}$$

and then the period effects of the populations, κ_t^1 and κ_t^2, are modelled jointly as a co-integrated bivariate process, assuming the existence of a common stochastic long-term trend linking the mortality of the two populations. In the same vein, Yang and Wang (2013) fit independent single population Lee-Carter models to multiple populations and then model simultaneously the period effects of the different populations using a Vector Error Correction Model (VECM). In order to impose further consistency in the forecast of the two-populations, Zhou et al. (2014) assume in (4.4) that both population share the same age-sensitivity term, i.e. $\beta_x^i = \beta_x$. For modelling the period indexes κ_t^1 and κ_t^2, Zhou et al. (2014) consider three methods: a random walk with drift for κ_t^1 plus an AR(1) for $\kappa_t^2 - \kappa_t^1$ (abbreviated RWAR by the authors), a vector autoregressive model (VAR); and a VECM. Similarly to Zhou et al. (2014), Kleinow (2015) has proposed a multiple population Common-Age-Effect model in which the age-sensitivity terms (age-effects) are common to all the populations.

In two closely linked studies looking at the mortality dynamics of a pair of related populations, Cairns et al. (2011a) and Dowd et al. (2011) have proposed the use of a two-population version of the Age-Period-Cohort (APC) model:

$$\log m^i_{xt} = \alpha^i_x + \kappa^i_t + \gamma^i_{t-x}, \quad i = 1, 2 \tag{4.5}$$

In both studies, the spreads between the state variables underlying the mortality models of each population are modelled as mean-reverting processes (e.g. an AR(1)) allowing different short-run trends in the mortality rates, but parallel long-run improvements. Cairns et al. (2011a) employ a Bayesian framework permitting a single stage estimation of the unobservable state variables and the parameters of the stochastic process driving them. Dowd et al. (2011) use a planetary analogy in which the mortality of the two populations are attracted to each other by a gravitational force dependent on the relative size of the populations. Although both papers use the APC model as an illustrating example, the proposed methodologies can be extended to other stochastic mortality models.

Another alternative for modelling multi-population mortality is to extend the widely used single-population Cairns-Blake-Dowd (CBD) model of mortality (Cairns et al., 2006). This approach has recently been considered by Li et al. (2015) who introduce two-population versions of the CBD model and its variants. For instance, in a full two-population version of the M7 model (the CBD model with cohort and quadratic effects proposed in Cairns et al. (2009)), the one-year death rate at for a person aged x at time t in population i, q^i_{xt}, is given by:

$$\text{logit}\, q^i_{xt} = \kappa^{(i,1)}_t + (x - \bar{x})\kappa^{(i,2)}_t + \left((x - \bar{x})^2 - \hat{\sigma}^2_x\right)\kappa^{(i,3)}_t + \gamma^i_{t-x}, \quad i = 1, 2, \tag{4.6}$$

where $\bar{x}$ is the average age in the data and $\hat{\sigma}^2_x$ is the average value of $(x - \bar{x})^2$. Li et al. (2015) also set out a systematic top-down procedure to evaluate if some of the stochastic factors in the two-population model can be shared by the two populations (e.g. by assuming in (4.6) that $\kappa^{(1,j)}_t = \kappa^{(2,j)}_t$ for some $j \in \{1, 2, 3\}$ or that $\gamma^1_{t-x} = \gamma^2_{t-x}$). For model forecasting Li et al. (2015) consider the same three approaches used by Zhou et al. (2014).

There are other studies examining the joint modelling of two populations which do not

lie under the category of pure extensions of the Lee-Carter or CBD models. Several of these studies pursue a relative approach whereby a single-population model is first fitted to one of the populations and then a separate model is fitted to the ratio of the mortality rates in the two populations. For instance, Jarner and Kryger (2011) have proposed a methodology for modelling the mortality experience of a small population in conjunction with the mortality experience of a much larger reference population. They assume that the reference population follows a deterministic long-run trend which is shared with the small population, and then model short term deviations of the small population from that trend using a multivariate stationary time series. Similarly, Wan and Bertschi (2015) model the larger population using the multi-factor single population model proposed by Plat (2009a) and then model the spread between the larger population and the smaller population with a three factor Lee-Carter model. In a related study, Plat (2009b) introduces a model for forecasting portfolio specific mortality alongside the relevant national population. In this model, portfolio specific mortality forecasts are obtained by combining national mortality projections derived from a standard single-population CBD model, with forecasts of the ratio between portfolio mortality rates and national population mortality rates. It is worth noting that Jarner and Kryger (2011), Wan and Bertschi (2015) and Plat (2009b) adopt the same approach for modelling the factors driving the dynamics of the mortality ratios and use a vector autoregressive model of order 1, VAR(1), so to avoid any long-term divergence of the mortality in the two populations.

Some authors have considered the use in a multipopulation setting of other well-known single population modelling approaches. For instance, Biatat and Currie (2010) extend to two populations the P-spline methodology (Currie et al., 2004) that has been successfully applied in the single population case, while Hatzopoulos and Haberman (2013) and Ahmadi and Li (2014) use the framework of generalised linear models (GLM) to obtain coherent mortality forecasts for multiple populations. Finally, instead of applying the co-integration methodology to the state variables of the Lee-Carter model as done by Li and Hardy (2011), Salhi and Loisel (2012) have applied the co-integration approach directly to the age-specific mortality rates of the England and Wales national population and the CMI assured lives' population.

4.4 Modelling the reference and the book population: A general formulation

Along the same lines of the general formulation of single population models considered in Chapter 5, we have identified a general framework under which most two population models that have been introduced in the literature can be accommodated. However, in order to facilitate this comparison between models, the way such models are proposed here may slightly differ from their original formulation.

As in Jarner and Kryger (2011) we choose a relative approach where the reference population is modelled first, and then the book mortality dynamics are specified so as to incorporate features from the reference population. This relative approach has some interesting features:

- It allows data mismatch between the reference and the book.
- It is well suited to the usual situation of the reference population being considerably larger than the book population.
- Single population models for the reference population are readily available and extensively studied, so this part of the model may be well established; allowing the focus to be on making an informed decision for the book part of the model, whilst retaining a good fit to the reference population.
- It provides consistency of approach when modelling several books using the same reference population.

4.4.1 Reference population

Following the discussion in Chapter 5, a general model for the reference population can be written as[1]

$$D_{xt}^R \sim \text{Bin}(E_{xt}^R, q_{xt}^R)$$

$$\text{logit}\, q_{xt}^R = \alpha_x^R + \sum_{j=1}^{N} \beta_x^{(j,R)} \kappa_t^{(j,R)} + \gamma_{t-x}^R \tag{4.7}$$

Here:

- The term α_x^R determines the reference mortality level for age group x.
- N is an integer indicating the number of age-period terms describing the mortality trend for the reference population.
- Each time index $\kappa_t^{(j,R)}$ contributes in specifying the reference mortality trend.
- Each coefficient $\beta_x^{(j,R)}$ dictates how mortality in the corresponding age group x reacts to a change in the time index $\kappa_t^{(j,R)}$.
- The term γ_{t-x}^R accounts for the cohort effect in the reference population.

4.4.2 Book population

Given the reference population model, the mortality of the book population is then specified through

$$D_{xt}^B \sim \text{Bin}(E_{xt}^B, q_{xt}^B)$$

$$\text{logit}\, q_{xt}^B - \text{logit}\, q_{xt}^R = \alpha_x^B + \sum_{j=1}^{M} \beta_x^{(j,B)} \kappa_t^{(j,B)} + \gamma_{t-x}^B \tag{4.8}$$

Note that we are modelling the difference in the (logit of) mortality in the book and the reference populations. Therefore:

[1] Here, we have chosen to work with one-year death probabilities, q_{xt}. Therefore, it is most natural to use the logit function and model deaths using a Binomial distribution. However, if one is interested in central death rates, m_{xt}, or the force of mortality, μ_{xt}, then the general modelling framework can be easily reformulated using a log link function and a Poisson Distribution. Moreover, we do not expect to see any material differences in our analysis if central death rates, m_{xt}, or the force of mortality, μ_{xt}, were considered instead.

- The term α_x^B determines the mortality level differences of the book population compared to the reference population for age group x. Hence the mortality level in the book is $\alpha_x^R + \alpha_x^B$
- M is an integer (generally less than or equal to N) indicating the number of age-period terms describing the mortality trend differences between the book population and the reference population.
- Each time index $\kappa_t^{(j,B)}$ contributes in shaping the difference in mortality trends.
- Each coefficient $\beta_x^{(j,B)}$ dictates how mortality differences for age group x react to a change in the time index $\kappa_t^{(j,B)}$.
- The term γ_{t-x}^B accounts for the differences in cohort effect in the two populations. Hence the cohort effect in the book is $\gamma_{t-x}^R + \gamma_{t-x}^B$.

Depending on how the model is specified, identification constraints may have to be added to (4.7) and (4.8) to ensure uniqueness of the parameters estimates. The estimation of the parameters of the model is performed using maximum likelihood in two stages whereby the reference population part of the model is estimated in a first stage and then, conditional on the reference population parameters, the book population part of the model is estimated in a second stage[2].

4.4.3 Time series dynamics

The modelling is completed by specifying the dynamics of the period indices and the cohort terms which are needed for forecasting and simulating future mortality. Although alternatives have been explored by some authors (see e.g. Zhou et al. (2014)) for the choice of the time series used in the dynamics, we stick to those commonly used in the literature.

[2] An alternative approach would be to estimate simultaneously the parameters in the reference and book populations along the lines of the estimation approach used in Chapter 2. This would in principle not materially change the fitted parameters as it is expected that book population has a small size relative to the reference population. Furthermore, for some models such as the two-population APC, CBD, M5 and M7 where the log-likelihood is separable, a two-stage estimation approach results in exactly the same parameter estimates as a joint estimation approach

Starting with the reference population, we assume that the period index is modelled as a multivariate random walk with drift (MRWD)

$$\boldsymbol{\kappa}_t^R = \mathbf{d} + \boldsymbol{\kappa}_{t-1}^R + \boldsymbol{\xi}_t^R, \qquad \boldsymbol{\xi}_t^R \sim N(0, \Sigma^R), \qquad \boldsymbol{\kappa}_t^R = \begin{pmatrix} \kappa_t^{(1,R)} \\ \vdots \\ \kappa_t^{(N,R)} \end{pmatrix},$$

and that the cohort index is modelled as an integrated auto-regressive process ARIMA(1, 1, 0)

$$\Delta\gamma_c^R = \phi_0 + \phi_1 \Delta\gamma_{c-1}^R + \epsilon_c^R, \qquad \epsilon_c^R \sim N(0, \sigma_R^2),$$

where $\Delta\gamma_c^R$ denotes $\gamma_c^R - \gamma_{c-1}^R$; $\mathbf{d}$ is an N-dimensional drift vector; ϕ_0, ϕ_1 are constants; and Σ^R is an $N \times N$ variance-covariance matrix of the multivariate white noise $\boldsymbol{\xi}_t^R$.

As for the book population, we follow the assumption commonly made in the literature (Li and Lee, 2005; Plat, 2009b; Cairns et al., 2011a; Jarner and Kryger, 2011; Li and Hardy, 2011; Wan and Bertschi, 2015). More precisely we assume that in the long-run the two populations experience similar improvements and therefore model the spread in the time indexes and in the cohort effects as stationary processes:

$$\boldsymbol{\kappa}_t^B = \Phi_0 + \Phi_1 \boldsymbol{\kappa}_{t-1}^B + \boldsymbol{\xi}_t^B, \qquad \boldsymbol{\xi}_t^B \sim N(0, \Sigma^B), \qquad \boldsymbol{\kappa}_t^B = \begin{pmatrix} \kappa_t^{(1,B)} \\ \vdots \\ \kappa_t^{(M,B)} \end{pmatrix} \tag{4.9}$$

$$\gamma_c^R = \psi_0 + \psi_1 \gamma_{c-1}^R + \epsilon_c^R, \qquad \epsilon_c^R \sim N(0, \sigma_B^2)$$

where Φ_0 is an M-dimension vector; Φ_1 is an $M \times M$ matrix; ψ_0, ψ_1 are constants; and Σ^B is an $M \times M$ variance-covariance matrix of the multivariate white noise $\boldsymbol{\xi}_t^B$. Thus:

- The time indices $\boldsymbol{\kappa}_t^B$ are then modelled as a vector auto-regressive process of order 1 (VAR(1)), for which we assume that the eigenvalues of the matrix Φ_1 are smaller than 1 in absolute value.
- The cohort difference γ_c^B follows an AR(1) process for which we assume that $\psi_1 < 1$.

- We are assuming independence of the time series determining the reference population and those determining the difference between the reference and the book populations. Considering correlations between $\boldsymbol{\xi}_t^R$ and $\boldsymbol{\xi}_t^B$ or between ϵ_c^R and ϵ_c^B is in principle possible, as has been done e.g. in Cairns et al. (2011a) and in Li et al. (2015). However, we refrain from considering this due to the fact that the estimation of the appropriate covariance matrix may not be straightforward. This is the case when the models involve multiple period effects; for example if the two-population M7 model defined in Equation (4.6) is used, then the covariance matrix is particularly large, containing up to 21 distinct elements associated with 6 period factors (see Li et al. (2015)). In addition, the estimation of the covariance matrix would be further complicated in the case when the time series for the reference and the book have different lengths, which is very frequent in practice.

Overall, the time series dynamics approach considered here corresponds to the RWAR approach discussed in Zhou et al. (2014) and Li et al. (2015).

4.5 Model selection criteria

With over 20 two-population models proposed in the literature at the time of writing (see Figure 4.1), our main goal is then to identify which model(s) are most likely to provide a satisfactory solution for assessing basis risk. To provide structure to this analysis, it is useful to test each model against certain criteria that a good and practical two-population model for basis risk assessment should satisfy.

Building on the literature comparing single population models (e.g. Continuous Mortality Investigation (2007); Cairns et al. (2008, 2009, 2011b); Haberman and Renshaw (2011)), we consider the following criteria:

(a) The model should produce a *non-perfect correlation* between mortality rates in the two populations.

(b) The model should produce a *non-perfect correlation* between year-on-year changes in mortality at different ages.

(c) The model should permit the *generation of sample paths* and the calculation of prediction intervals.

(d) The structure of the model should allow the incorporation of *parameter uncertainty* in simulations.

(e) The model should permit the consideration of a *cohort effect* if necessary.

(f) The model should be *compatible with the data* that is likely to be available when performing basis risk assessment exercises.

(g) The model should be *straightforward to implement* using standard statistical methods likely to be available to practitioners.

(h) The model should be *transparent* enough so that the model assumptions, limitations and outputs are understood by the users and can be easily explained to non-experts.

(i) The model should show a reasonable *goodness-of-fit* to historical data in both the reference population and the book population for a wide range of book populations.

(j) The model should show a reasonable *goodness-of-fit* for metrics involving the two populations such as differences or ratios in mortality rates or life expectancies for a wide range of book populations.

(k) The model should be relatively *parsimonious*.

(l) The model should produce *plausible and reasonable central projections* of both single-population and of two-population metrics.

(m) The model should produce *plausible and reasonable forecast level of uncertainty* in projections of both single-population and of two-population metrics, which are in line with historical levels of variability.

(n) The parameter estimates and model forecasts should be *robust* relative to the period of data and range of ages employed.

Most of the above criteria coincide with the criteria that a good single population model should satisfy; we thus refer the reader to Continuous Mortality Investigation (2007, Section 8) and Cairns et al. (2008, Section 3) for a detailed discussion of their relevance. By contrast, criteria (a), (j), (l) and (m), referring to correlations between the mortality rates in the two populations and to the performance of the models in relation to two-population metrics, are new criteria which are of prime importance to the application of

two-population models in the assessment of basis risk in standardised longevity hedges. On the one hand, if a model assumes a perfect correlation between mortality rates in the two populations then it will imply that the reference population provides a perfect match for the book population, trivially leading to no (or very little) demographic basis risk. On the other hand, since demographic basis risk emerges from the mismatch in the mortality of the reference and the book population, it is critical that the two-population model shows a good fit to metrics involving the two populations, and that forecast levels of uncertainty and central trajectories for these metrics are plausible and consistent with historical differences between the populations.

We note, however, that a two-population model which might not be suitable for basis risk assessment, may be an appropriate model for other applications in which some of the above criteria would be superfluous. For example, consider the case of valuing the liabilities of a pension book with sparse data, where we may consider a two-population model to borrow information from a larger reference population with the objective of improving the accuracy in the projections of the pension schemes' mortality. In this situation having a non-perfect correlation between the mortality of the two populations would be irrelevant for the problem and the performance of the model relative to two-population metrics would be of lesser importance.

4.6 Identifying an appropriate two-population model

Given the wealth of models available and the large number of criteria, we have followed a two-stage filtering process to identify the model structures likely to be suitable for basis risk assessment. In a first stage, we focus on criteria (a) to (h) which are theoretical properties of a model and can be evaluated without reference to a specific dataset. Then, in a second stage, we focus on criteria (i) to (n) which can only be evaluated after a model has been fitted to data. More specifically, in the second stage of filtering we evaluate the goodness of fit, the reasonableness of the output, the forecasting performance, and the robustness of those models which pass the first stage of filtering.

4.6.1 Stage 1 of filtering: Criteria requiring no data to assess

We first evaluate all the candidate models against those criteria that can be assessed independently of data or the actual fitting of the models. This process permits the identification of a number of models which could be rejected, either because their theoretical properties are not suitable for basis risk assessment or because they are unlikely to be accessible to the wider industry.

4.6.1.1 Non-perfect correlation between mortality rates in the two populations

A perfect correlation between the mortality rates q^B_{xt} and q^R_{xt} implies that the two populations move in tandem, with changes in the mortality of the book population matched by changes in the mortality of the reference population[3]. This will result in the model spuriously suggesting that there is no (or very little) demographic basis risk. This is the case for those Lee-Carter based models with a single common period effect for both populations, leading us to view the Stratified Lee-Carter, the Common Factor Model, the Three-way Lee-Carter, and the Joint-κ model as inadequate models for assessing demographic basis risk.

4.6.1.2 Non-perfect correlation between year-on-year changes in mortality at different ages

This criterion refers to the correlation between $q^R_{x,t+1} - q^R_{x,t}$ and $q^R_{y,t+1} - q^R_{y,t}$ (or between $q^B_{x,t+1} - q^B_{x,t}$ and $q^B_{y,t+1} - q^B_{y,t}$) for $x \neq y$. As noted by Cairns et al. (2008), a model that assumes a perfect correlation between changes in mortality at different ages would incorrectly suggest that holding a derivative instruments linked to a single age would provide just as good a hedge as holding instruments linked to a range of different ages. Disregarding this issue can result in a misassessment of the structuring basis risk underlying a longevity hedge.

[3]Note that the correlation between the mortality rates q^B_{xt} and q^R_{xt} may not be perfect, although it will be close to one, even when correlation is perfect on the logit scale used by the models introduced in section 4.4. Also note that having a perfect correlation between the populations does not necessarily imply that the two populations experience exactly the same mortality improvements. For instance, the Joint-κ model in Equation (4.1) and the Three-way Lee-Carter in Equation (4.2) allow for improvement rate differentials, but imply a perfect correlation between the populations.

Lee-Carter type models with a single period effect and no cohort effect, such as the Co-integrated Lee-Carter and the Lee-Carter+VAR/VECM, have a trivial age correlation structure. In addition, the two-population APC model in Equation (4.5) implies that there is perfect correlation at all ages except at the youngest ages, where there is potentially additional randomness arising from the arrival of new cohorts with an unknown cohort effect (see Cairns et al. (2009)). In contrast, two-population extensions of the CBD model allow for imperfect correlations between annual changes in mortality at different ages due to the presence of multiple period factors.

However, we do not discard any model due to its poor age correlation structure for two reasons. In many instances it may only be required to perform an indicative assessment of the demographic basis risk associated with an index-based hedge, without necessarily considering in detail the precise structuring of the hedge. Further, in order to assess model risk, it may be useful to consider an alternative model to the one used in structuring the hedge.

4.6.1.3 Generation of sample paths

Mortality sample paths are required for the assessment of the uncertainty in the cashflows of a mortality-linked security as well as for the pricing and structuring of a longevity hedge. A distinguishing feature of the P-Spline model of Biatat and Currie (2010) and the multipopulation GLM of Ahmadi and Li (2014) is that they assume that mortality follows a deterministic time trend, meaning that these models cannot generate sample paths. Hence, we do not consider these two models any further.

4.6.1.4 Parameter uncertainty

Given that most book populations have a limited amount of data, the parameters of the models may be subject to significant estimation error. It is thus important to be able to consider the impact that parameter risk can have on forecasts levels of uncertainty and on hedge effectiveness. With the exception of the Bayesian two-population model of Cairns et al. (2011a) which naturally accounts for parameter uncertainty, none of the studies we have reviewed considers explicitly parameter uncertainty. Nevertheless, for most

of the models it is possible to incorporate parameter uncertainty using bootstrapping techniques such as the ones proposed in Brouhns et al. (2005) and Koissi et al. (2006).

4.6.1.5 Cohort effect

For some countries, including England and Wales, it is important that models allow for the now well-accepted cohort effect, separating out general improvements over time to those specific to a given birth cohort. Although not all the models include a cohort effect, they can in principle be extended to include such an effect.

4.6.1.6 Compatibility with available data

The data requirements of some of the models are incompatible with the likely available data. For instance, it is unlikely that the history length in the book population will be as long as that in the reference population, hindering the application of models which cannot easily deal with such a scenario. In particular, this requirement leads to the rejection of two further Lee-Carter based models, namely the Lee-Carter VAR/VECM and the Co-integrated Lee-Carter, as well as of the Co-integration approach of Salhi and Loisel (2012). In addition, models which use several book-specific period terms are poorly rated against the data compatibility criterion, as the more period terms a model has, the longer the data history that is required to estimate appropriately the time series processes needed for forecasting.

4.6.1.7 Ease of implementation and transparency

Ease of implementation and transparency are essential for a model to be of general use by practitioners. Accordingly, these two criteria lead to the rejection of several other models. In particular, the Co-integration approach of Salhi and Loisel (2012) and the Multipopulation GLM of Hatzopoulos and Haberman (2013) are considered to be impractical for basis risk assessment as they are complex models which are computationally involved to implement and may be difficult to communicate to non-experts. In addition, we disregard the Plat+Lee-Carter model of Wan and Bertschi (2015) (apart from other reasons discussed later) because it combines a parametric structure for the reference with a non-parametric structure for the book, and we believe that for the

sake of interpretability of the parameters both parts of the model should be within the same class of models. Finally, although the Bayesian two-population APC model of Cairns et al. (2011a) is particularly amenable to the short history and modest exposures sizes of most book datasets, the implementation and transparency issues related to the underlying Bayesian approach have led us to reject this model. However, some of the features of the approach of Cairns et al. (2011a) will still be investigated subsequently in this chapter through a maximum-likelihood implementation of the two-population APC model.

4.6.2 Stage 2 of filtering: Criteria requiring data to assess

After carrying out the initial data-independent assessment, 10 models can be identified as candidates which are worth testing against the data dependent criteria. These models are: the Augmented Common Factor model and its cohort extension, the Relative Lee-Carter model with cohorts, the Common Age Effect Model, the two-population APC (Gravity model), the two-population M5, the two-population M6, the two-population M7, the Saint model, and the Plat relative model.

In the second stage of filtering we focus on those model criteria requiring data to be assessed. This involves the evaluation of the historical goodness-of-fit, the (subjective) evaluation of the reasonableness of the forecast level of uncertainty produced by the models, and the evaluation of the forecasting performance and robustness of the models.

4.6.2.1 Data

The evaluation of the criteria in this stage requires data for model fitting. We have used as the reference population data the England and Wales male mortality experience as obtained from the Human Mortality Database (2013). For the purposes of our analysis we have focused on a subset of these data covering calendar years 1961-2010 and those older ages most relevant to longevity hedging, namely ages 60-89.

For the book population we use synthetic datasets generated based on England mortality data by quintiles of the Index of Multiple Deprivation 2007 (IMD 2007)[4] and the socio-economic composition observed within individual occupational pension schemes of the Club Vita dataset[5]. Specifically, the synthetic datasets used throughout this chapter have been generated by randomly sampling from the national IMD data to obtain a dataset of exposure size, history length, and IMD profile desired. The technical details of this data sampling process are described in Appendix 4.A.

The use of synthetic data as opposed to actual pension scheme data facilitates a more thorough assessment of the models. Concretely, synthetic datasets permit us to control some key characteristics of the book population data while changing others. For instance, it allows us to vary the history length and exposure size of the book data whilst keeping the socio-economic and age composition constant. Moreover, synthetic datasets let us rely on the longer history of the national IMD mortality data to perform backtesting exercises such us those described in Section 4.6.2.7.

For the assessment of the goodness-of-fit of the models, we consider four different synthetic datasets to reflect the variety of socio-economic mixes observed in real pension schemes and annuity books. In each case, the socio-economic splits are informed by the profiles seen within the Club Vita dataset. Table 4.1 describes the socio-economic profiles of these datasets. In all cases we use sample books with historical exposures of 100,000 male lives per year, which we believe is the largest exposure any scheme or insurer is likely to have. We also assume that book data are available for the period 1981-2010 and ages 60 to 89. Finally, we use the age distribution of the English population to split by age the total exposure of each of the sample schemes. Figure 4.2 depicts the average age distribution of the exposure for the period 1981-2010.

Figure 4.3 depicts the ratio of the mortality in each of the four datasets to the mortality in England and Wales. From this we note that:

[4] A detailed analysis of the mortality data used in this chapter can be found in Chapter 2 or in Lu et al. (2014). Further information on the Index of Multiple Deprivation see Noble et al. (2007).

[5] Club Vita is an organisation which provides longevity analytics to pension schemes. The schemes in the Club Vita dataset span a wide range of sizes including some of the largest DB schemes in the UK and (as at September 2014) consists of nearly 6 million member records.

TABLE 4.1: Description of the book datasets used for model testing. Q1 represents the least deprived quintile of England and Q5 the most deprived quintile.

Dataset	**Description**	**Percentage of exposure by IMD quintile**				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**
Typical Lives	This is the typical IMD split we would expect to see in a book population weighted by lives (head-count)	23%	22%	21%	20%	14%
Typical Amounts	This uses the same split as the typical (lives) but weighted by individual pension amounts to approximate the effect of a typical portfolio's liability distribution amongst the IMDs	30%	25%	20%	15%	10%
Extreme Wealthy	This reflects the split by IMD (on an amounts weighted basis) that we would expect to see in a very affluent book population	45%	30%	20%	5%	0%
Extreme Deprived	This reflects the split by IMD (on a lives weighted basis) that we would expect to see in a book skewed towards lower socio-economic groups	10%	15%	15%	25%	35%

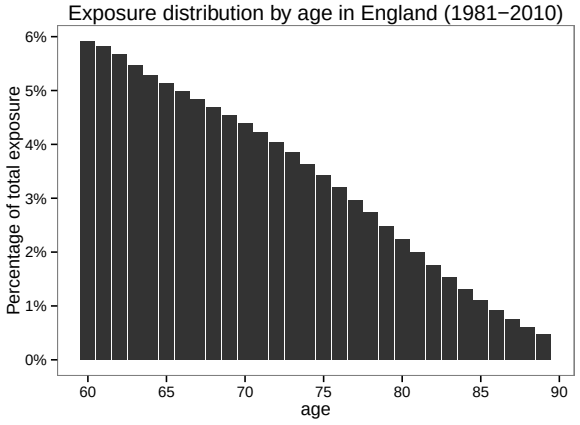


FIGURE 4.2: Average age distribution of total exposure for England males in the period 1981-2010.

- The ordering of the ratios in the four data sets is consistent with their socio-economic mixes, with the "Extreme Wealthy" dataset having below average mortality (ratio < 1) and the "Extreme Deprived" dataset having above average mortality (ratio > 1).

- The mortality ratios for both the "Typical Lives" and "Typical Amounts" dataset are close to 1 reflecting the socio-economic mix of these datasets being close to the average in England and Wales.

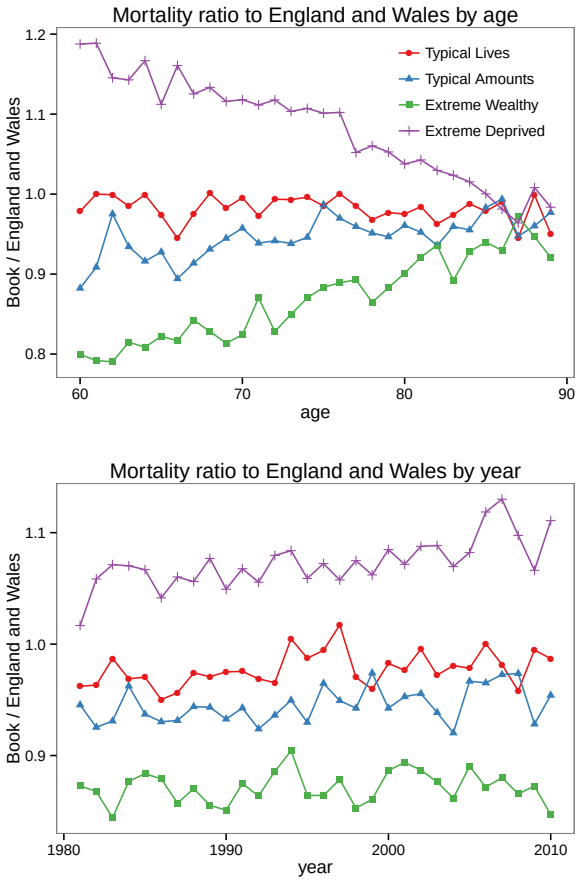


FIGURE 4.3: Ratio of the mortality in each of the four synthetic book datasets to the mortality in England and Wales. The top graph shows this ratio by age while the one on the bottom presents the time evolution of this ratio.

- In all of the datasets the mortality ratios converge with rising age. This is consistent with the commonly reported decrease in socio-economic mortality differences as people age (Hoffmann, 2005).

- None of the datasets shows any very marked increasing or decreasing time trend in the mortality ratios, albeit there is a slight upward trend in the "Extreme Deprived" dataset. This is consistent with the slower mortality improvements for the two most deprived quintiles of England reported in Chapter 2.

4.6.2.2 Model fitting

To facilitate the fitting of the 10 models that passed our first-stage filtering, we have followed the general modelling framework described in Section 4.4 whereby each model can be viewed as a model for the reference population combined with a model for the

book population (or perhaps more accurately, a model for the mortality ratio between reference and book). As such, the fitting and the assessment of the goodness-of-fit of a model can be carried out in two stages: fitting and assessing the goodness-of-fit of the reference model, followed by the fitting and the assessment of the goodness-of-fit of the book part of the model. All the model fitting performed in this chapter has been carried out using the **R** package **StMoMo** introduced in Chapter 5, which facilitates the implementation of stochastic mortality models using the unifying framework of generalised (non)-linear models.

We note that conclusions regarding the goodness-of-fit of the model to the reference may lead us to slightly modify the original formulation of certain of the two-population models before assessing the goodness-of-fit of the book part of the model. The specific modifications for each particular two-population model are described later in this section.

4.6.2.3 Goodness-of-fit for reference population

For the reference population we concentrate on the eight models described in Table 4.2 which are labelled LC, LC2, LC+Cohorts, LC2+Cohorts, APC, M5, M6 and M7. LC is the Lee and Carter (1992) model and LC2 is the two-factor extension proposed in Renshaw and Haberman (2003). The LC+Cohorts is one of the Renshaw and Haberman (2006) cohort extensions of the original Lee-Carter model while LC2+Cohorts is the LC2 model with a cohort effect[6]. The APC model is a special case of the LC+Cohorts. Models M6 and M7 are extensions of the original CBD model (M5) and were proposed in Cairns et al. (2009). These models encompass the reference population models underlying the models that passed our first stage filtering with the exception of the Saint model. For the Saint model, instead of the frailty-type model considered originally by Jarner and Kryger (2011) which we believe is too complex to be accessible to practitioners and does not permit the generation of sample paths, we will use M7 to model the mortality of the England and Wales reference population. The constraints used in fitting the models are described in Appendix 4.B.

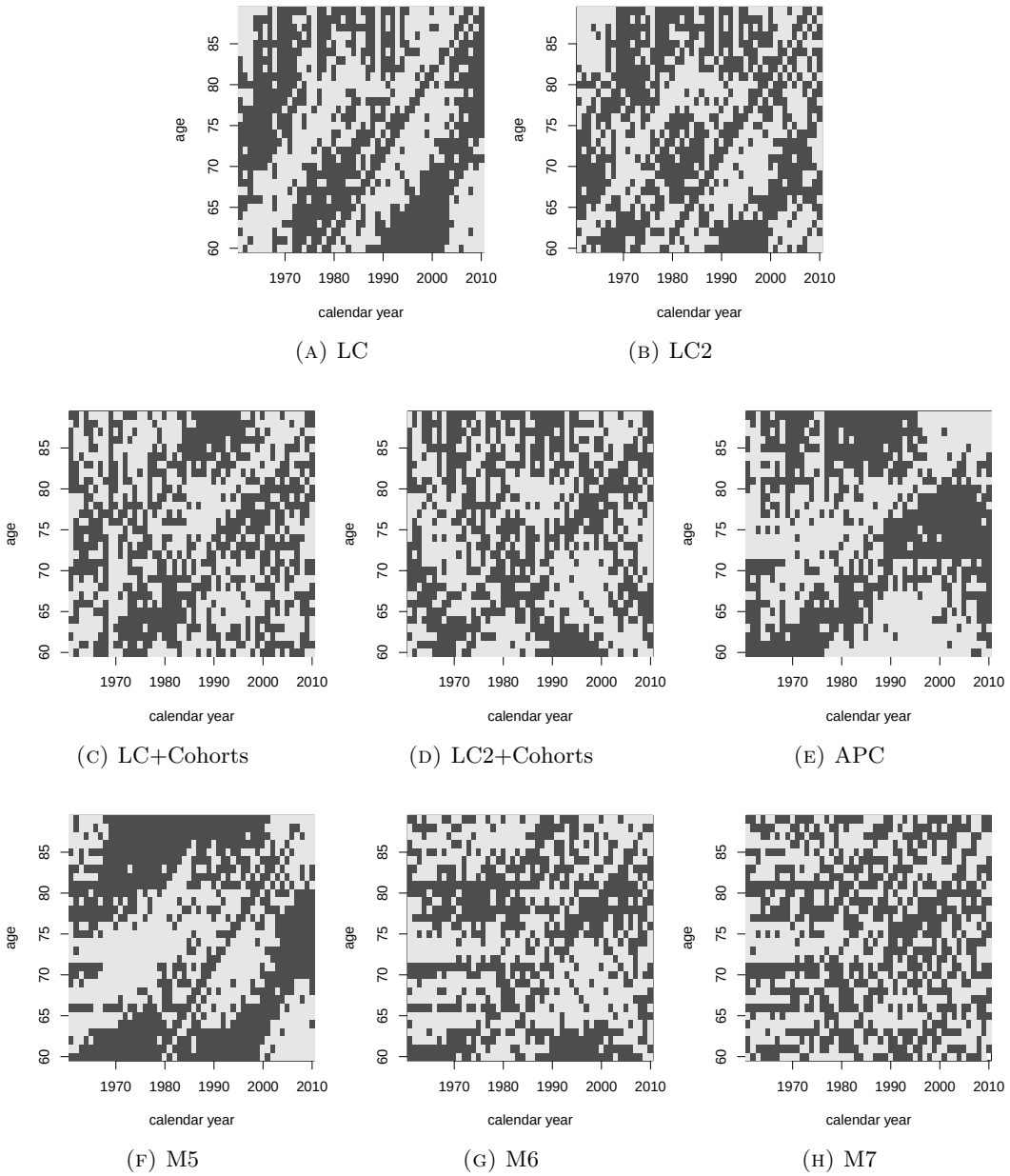
[6]It is well known that cohort extensions of the Lee-Carter model have robustness and stability issues (see e.g. Cairns et al. (2009)) with models being very sensitive to changes in the data or the fitting algorithm. Therefore, when implementing the LC+Cohorts model we follow the approach discussed in in Chapter 6 which helps resolve many of the stability issues. Similarly, for the LC2+Cohorts model we follow the approach employed by Yang et al. (2014) and first fit the LC2 model without cohorts and then, taking the parameters of the LC2 as given, estimate the cohort effect.

TABLE 4.2: Mathematical description of the eight models considered for the reference population.

Model	**Formula**
LC	$\text{logit}\, q_{xt}^R = \alpha_x^R + \beta_x^R \kappa_t^R$
LC2	$\text{logit}\, q_{xt}^R = \alpha_x^R + \beta_x^{(1,R)} \kappa_t^{(1,R)} + \beta_x^{(2,R)} \kappa_t^{(2,R)}$
LC + Cohorts	$\text{logit}\, q_{xt}^R = \alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R$
LC2 + Cohorts	$\text{logit}\, q_{xt}^R = \alpha_x^R + \beta_x^{(1,R)} \kappa_t^{(1,R)} + \beta_x^{(2,R)} \kappa_t^{(2,R)} + \gamma_{t-x}^R$
APC	$\text{logit}\, q_{xt}^R = \alpha_x^R + \kappa_t^R + \gamma_{t-x}^R$
M5	$\text{logit}\, q_{xt}^R = \kappa_t^{(1,R)} + (x - \bar{x})\kappa_t^{(2,R)}$
M6	$\text{logit}\, q_{xt}^R = \kappa_t^{(1,R)} + (x - \bar{x})\kappa_t^{(2,R)} + \gamma_{t-x}^R$
M7	$\text{logit}\, q_{xt}^R = \kappa_t^{(1,R)} + (x - \bar{x})\kappa_t^{(2,R)} + \left((x - \bar{x})^2 - \hat{\sigma}_x^2\right) \kappa_t^{(3,R)} + \gamma_{t-x}^R$

We assess the goodness-of-fit of the candidate reference population models by examining sign plots of deviance residuals. Regular patterns in the residuals are an indication of the inability of the model to describe all of the features of the data appropriately. Figure 4.4 plots the sign of the residuals in an age-period grid for the eight reference population models. From this figure we note the following:

- Models LC, LC2 and M5, which do not incorporate a cohort effect, show diagonal clusters of positive and negative residuals. This provides evidence for the need of including a cohort effect when modelling mortality in the England and Wales reference population.
- The APC model shows a strong clustering of positive and negative residuals. This is due to its inability to allow for varying improvement rates with age.
- The LC+Cohorts, LC2+Cohorts and M6 look reasonably random, but still with some clustering of positive and negative residuals. Closer inspection of scatterplot of the residuals of these three models (see Figure 4.5) reveals that the LC+Cohorts and LC2+Cohorts do not show any clear pattern while model M6 shows a strong pattern by age. This latter pattern reflects the lack of a quadratic age term in model M6 which may be necessary to capture the commonly observed curvature of the mortality rates on a logit scale.
- The M7 look reasonably random, indicating a satisfactory fit to the data.



(A) LC (B) LC2

(C) LC+Cohorts (D) LC2+Cohorts (E) APC

(F) M5 (G) M6 (H) M7

FIGURE 4.4: Sign plots of deviance residuals for the England and Wales males reference population. Positive residuals in grey and negative residuals in black.

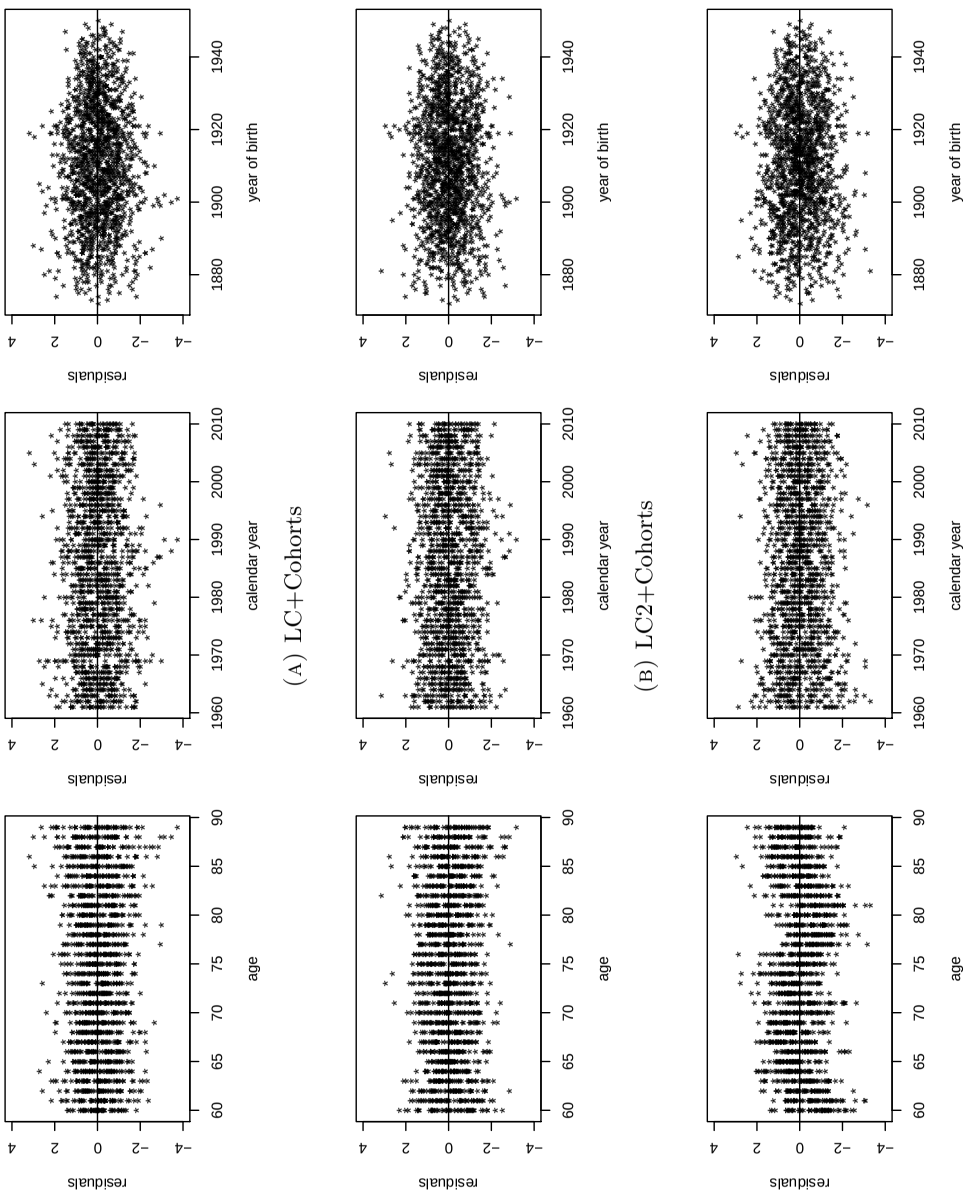


FIGURE 4.5: Scatterplots of deviance residuals for the England and Wales male population for models LC+Cohorts, LC2+Cohorts and M6.

TABLE 4.3: Effective number of parameters, and AIC for different models fitted to the England and Wales reference population. AIC rankings across models are presented in brackets

Model	**Number of parameters**	**AIC (rank)**
LC	108	24516 (7)
LC2	184	22125 (4)
LC+Cohorts	185	18590 (2)
LC2+Cohorts	262	18817 (4)
APC	156	20379 (5)
M5	100	25516 (8)
M6	177	18709 (3)
M7	226	18256 (1)

In order to assess more formally the goodness of fit of the eight reference population models, we present in Table 4.3 the Akaike Information Criterion (AIC), defined as $2\nu_R - 2\mathcal{L}_R$, where $\mathcal{L}_R$ is the log-likelihood of the reference model and ν_R is the number of parameters of the models. The AIC provides a way of assessing the balance between quality of fit and parsimony. In general, a lower value of AIC is preferable. We note that the LC+Cohorts and M7 are the best fitting models.

The previous observations are consistent with the existing literature comparing single population mortality models (e.g. Cairns et al. (2009) and Haberman and Renshaw (2011)), where the LC+Cohorts and M7 have been identified as appropriate models for modelling mortality in the England and Wales population. Accordingly, in our subsequent evaluation of two-population extensions of the Lee-Carter model, we will assume that the reference population is modelled using a LC+Cohorts. Similarly, when assessing the two-population extensions of the CBD model, we will assume that the reference population is modelled using an M7 model.

4.6.2.4 Goodness-of-fit for book population

To correct some of the goodness-fit issues discussed above, we have adapted several of the candidate two-population models before carrying out further goodness-of-fit assessments. Specifically, we have made the following adaptations:

- The Common Age Effect model, as proposed in Kleinow (2015), does not include a cohort effect. Therefore, given that there is strong evidence of a cohort effect in

England and Wales, in our testing we extend this model to include such an effect. The reference population model is then a LC+Cohorts model in the terminology of Section 4.6.2.3.

- Similarly, for the Augmented Common Factor model we should consider a cohort effect, but doing so would turn the model into the Relative Lee-Carter model with cohorts. Consequently, the Augmented Common Factor model is not considered further in the analysis.

- In the two-population M5 and the two-population M6 models we replace the corresponding M5 and M6 models for the reference population with an M7 model.

- For the Relative Plat model we assume an M7 model for the reference population as opposed to the M5 model originally assumed by Plat (2009b). In addition, while Plat (2009b) models directly the mortality ratio between the reference and the book population, $q^B_{xt}\big/q^R_{xt}$, we model the logit difference in mortality, $\text{logit}\, q^B_{xt} - \text{logit}\, q^R_{xt}$.

- For the Saint model we assume an M7 model for the reference population instead of the frailty-type model originally used by Jarner and Kryger (2011).

For comparison purposes, in some of our additional goodness-of-fit and reasonableness testing we will consider the Common Factor Model with added cohorts. This model, which was previously deemed inappropriate as it unrealistically implies zero basis risk, is useful for illustrating some of the undesired characteristics in a model.

Table 4.4 summarises the models whose goodness-of-fit will be investigated further. The parameter constraints associated with these model structures are described in Appendix 4.B. The Common Factor models with cohorts (CF+Cohorts), the Common Age Effect model with cohorts (CAE+Cohorts), and the relative Lee-Carter model with cohorts (RelLC+Cohorts) belong to the Lee-Carter family of models. The CF+Cohorts only allows for level differences between the reference and the book population, whilst the CAE+Cohorts and the RelLC+Cohorts also allow for improvement differences. Nevertheless, the latter two models differ in the specification of the age-modulating factor β^B_x accompanying the book-specific time index κ^B_t: in the RelLC+Cohorts β^B_x is estimated directly from the observed logit difference of mortality between the book and reference data while in the CAE+Cohorts β^B_x is borrowed from the reference population model, i.e., $\beta^B_x \equiv \beta^R_x$.

TABLE 4.4: Mathematical description of the two-population models considered for goodness-of-fit assessment.

Original Model	**Model tested**	$\mathbf{logit\, q^R_{xt}}$	$\mathbf{logit\, q^B_{xt} - logit\, q^R_{xt}}$
Common Factor	CF+Cohorts	$\alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R$	α_x^B
Common Age Effect	CAE+Cohorts	$\alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R$	$\alpha_x^B + \beta_x^R \kappa_t^B$
Relative Lee-Carter with cohorts	RelLC+Cohorts	$\alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R$	$\alpha_x^B + \beta_x^B \kappa_t^B$
Gravity	Gravity (APC)	$\alpha_x^R + \kappa_t^R + \gamma_{t-x}^R$	$\alpha_x^B + \kappa_t^B + \gamma_{t-x}^B$
Two-population M5	M7-M5	$\kappa_t^{(1,R)} + (x-\bar{x})\kappa_t^{(2,R)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,R)} + \gamma_{t-x}^R$	$\kappa_t^{(1,B)} + (x-\bar{x})\kappa_t^{(2,B)}$
Two-population M6	M7-M6	$\kappa_t^{(1,R)} + (x-\bar{x})\kappa_t^{(2,R)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,R)} + \gamma_{t-x}^R$	$\kappa_t^{(1,B)} + (x-\bar{x})\kappa_t^{(2,B)} + \gamma_{t-x}^B$
Two-population M7	M7-M7	$\kappa_t^{(1,R)} + (x-\bar{x})\kappa_t^{(2,R)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,R)} + \gamma_{t-x}^R$	$\kappa_t^{(1,B)} + (x-\bar{x})\kappa_t^{(2,B)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,B)} + \gamma_{t-x}^B$
Saint model	M7-Saint	$\kappa_t^{(1,R)} + (x-\bar{x})\kappa_t^{(2,R)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,R)} + \gamma_{t-x}^R$	$\kappa_t^{(1,B)} + (x-\bar{x})\kappa_t^{(2,B)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,B)}$
Plat relative model	M7-Plat	$\kappa_t^{(1,R)} + (x-\bar{x})\kappa_t^{(2,R)} + \left((x-\bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3,R)} + \gamma_{t-x}^R$	$\frac{100-x}{100-\bar{x}}\kappa_t^{(1,B)}$

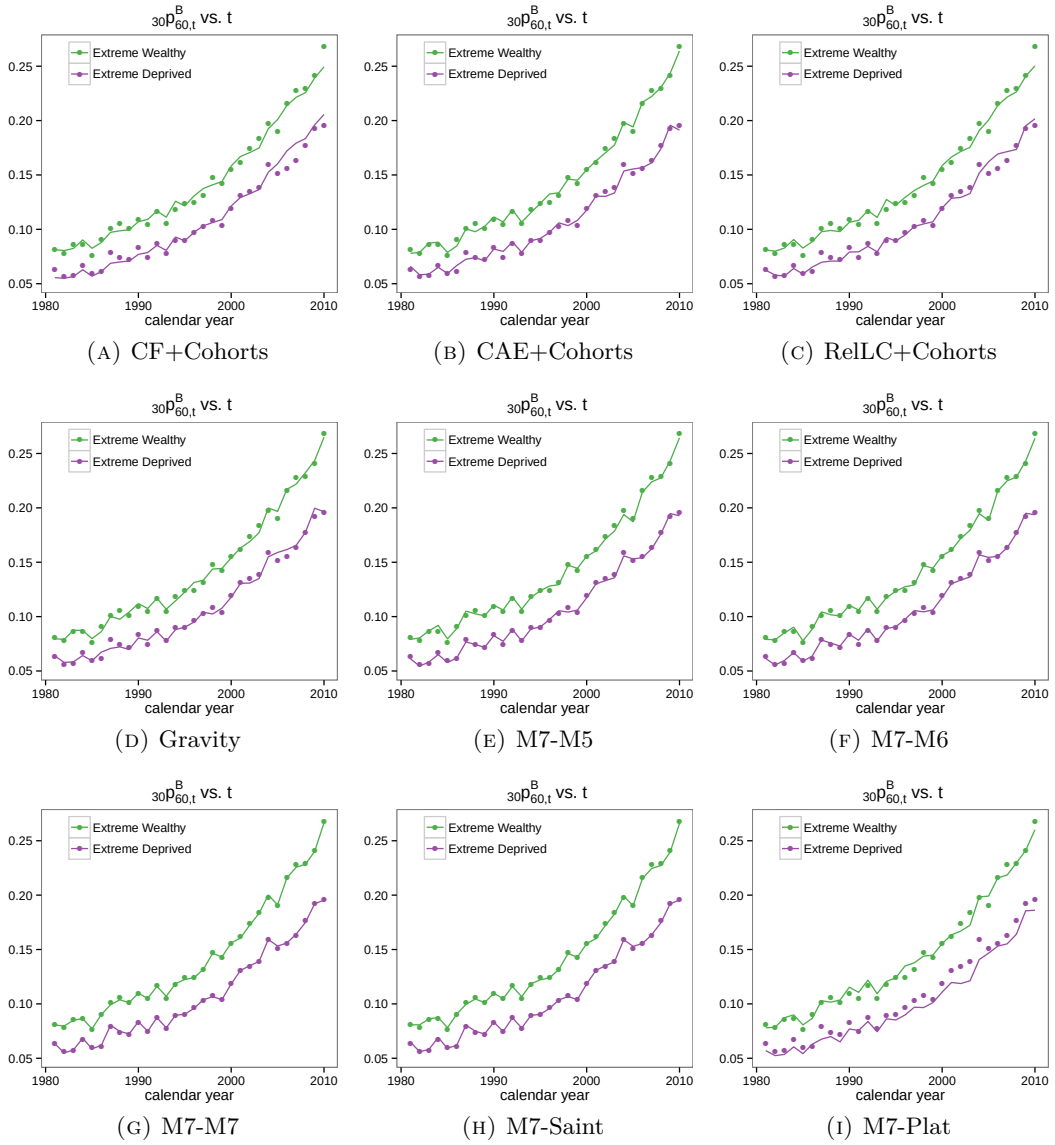
The Gravity model corresponds to a Binomial-logit version of the two-population APC introduced before in Equation (4.5).

Models M7-M5, M7-M6, M7-M7, M7-Saint, and M7-Plat (which are the implemented versions of the two-population CBD, the two-population M6, the two-population M7, the Saint model, and the Relative Plat model, respectively) all belong to the CBD family of models. These models differ in the type of differences between the book and the reference population that are allowed for in the parametric age functions:

- M7-M5 and M7-M6 allow only for level and slope differences with M6 also allowing for cohort differences;
- M7-Saint, M7-M7 allow for level, slope and curvature differences with M7 also allowing for cohort differences; and
- M7-PLAT is a constrained version of M7-M5 assuming that at age 100 there is no difference between the reference and the book.

A good two-population model should show a reasonable fit to the historical mortality rates in both the reference population and the book population. In addition, the model should show a good fit to metrics involving the two populations such as differences or ratios of mortality rates. This last criterion is very relevant as demographic basis risk emerges from the mismatch in the mortality of the reference and the book population.

When assessing the quality of the fit of the models with respect to the book population and with respect to two-population metrics, we have found that the inspection of residual sign plots is not very informative. In principle, this can be attributed to the fact that cohort and age patterns in the book residuals may be confounded with the sampling noise in the book data. As an alternative, the examination of plots of fitted vs. observed period survival probabilities in the book and the corresponding plots for ratios of period survival probabilities in the book and the reference can give useful insight into the goodness-of-fit of the models. As an illustration, Figure 4.6 depicts the fitted and observed 30 years period survival probabilities at age 60 for the "Extreme Wealthy" and the "Extreme Deprived" sample schemes using several models. Figure 4.7 plots the corresponding fitted and observed ratios of period survival probabilities between both sample schemes and the England and Wales reference population. Figures 4.6 and 4.7 are representative



(A) CF+Cohorts (B) CAE+Cohorts (C) RelLC+Cohorts

(D) Gravity (E) M7-M5 (F) M7-M6

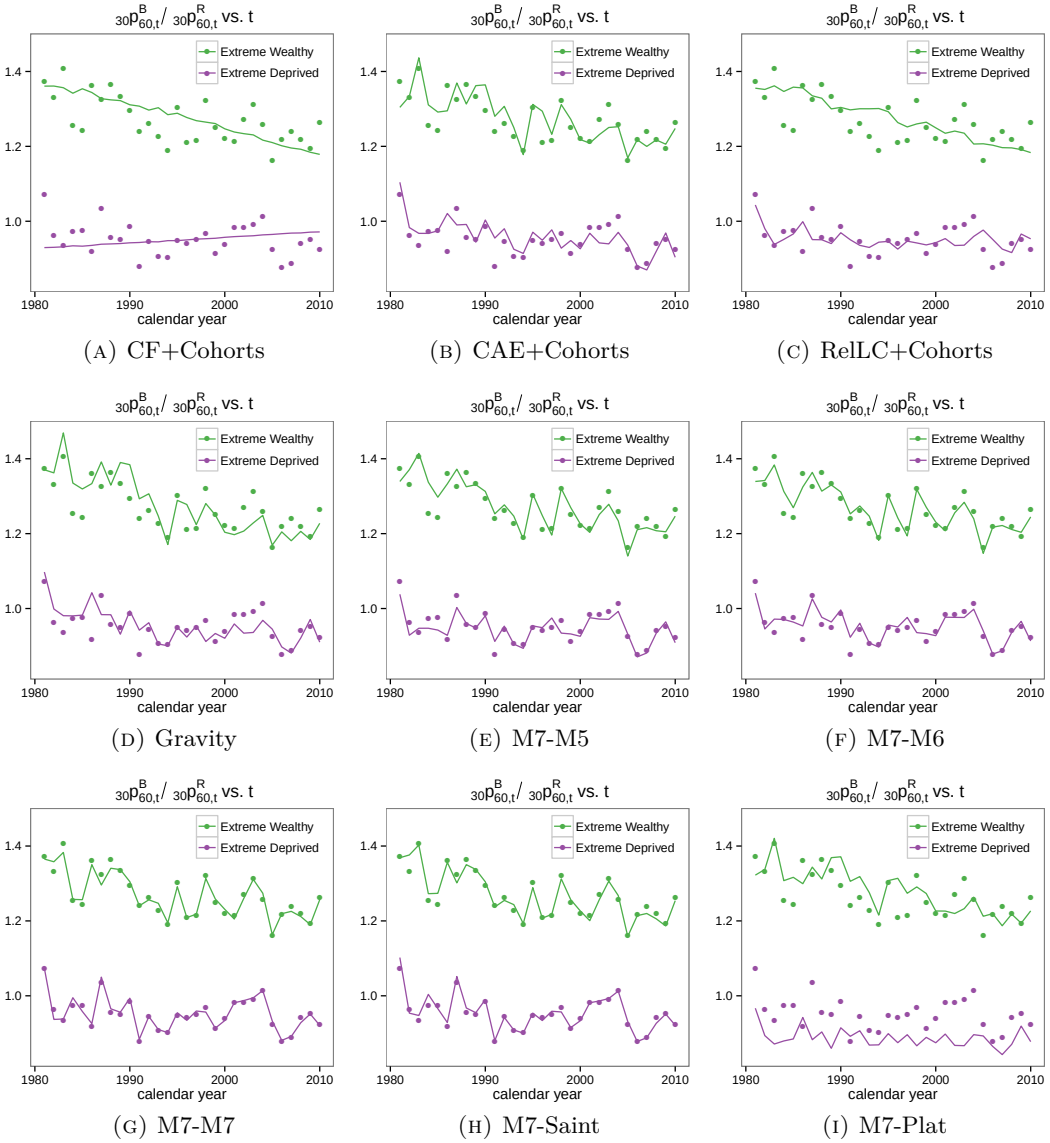
(G) M7-M7 (H) M7-Saint (I) M7-Plat

FIGURE 4.6: Fitted vs. observed 30 year period survival probabilities at age 60 for the "Extreme Wealthy" and the "Extreme Deprived" sample schemes.

of the detailed analyses we have carried out and which have helped with our assessment of the performance of the different models under consideration.

Figure 4.6 shows that, with the exception of the M7-Plat model which shows a slight underestimation in the later years when fitted to the "Extreme Deprived" scheme (see Figure 4.6i), all the models show a similar and reasonable fit to the period survival probabilities in the book. By contrast, when considering ratios of survival probabilities the models show very different performances. In particular, from Figure 4.7 we note:

- For the "Extreme Deprived" dataset the M7-Plat model shows a stark bias in



(A) CF+Cohorts (B) CAE+Cohorts (C) RelLC+Cohorts

(D) Gravity (E) M7-M5 (F) M7-M6

(G) M7-M7 (H) M7-Saint (I) M7-Plat

FIGURE 4.7: Fitted vs. observed ratio of 30 year period survival probabilities at age 60 for the "Extreme Wealthy" and the "Extreme Deprived" sample schemes.

the fitted ratios (see Figure 4.7i) consistent with the underestimation seen in the period survival probabilities in the book population. In an attempt to improve the fit of the M7-Plat model, instead of assuming that crossing of mortality between the reference and book population occurs at the prefixed age 100, we have treated the age of crossing as an additional parameter that needs to be estimated from the data. However, this has not eliminated the bias issues suggesting that the M7-Plat model might be too restrictive for some datasets. Therefore we do not consider the M7-Plat model further as a candidate for basis risk assessment.

- The CF+cohort and the RelLC+Cohort models produce very smooth ratios of

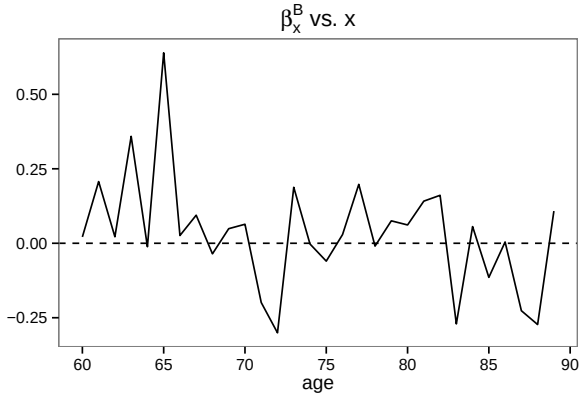


FIGURE 4.8: Fitted age modulating parameter β_x^B for the RelLC+Cohorts fitted to the "Extreme Wealthy" scheme.

survival probabilities which seem to understate the observed volatility in the ratios (see Figures 4.7a and 4.7c). Whilst the poor performance of the CF+cohort model was expected due to the perfect correlation between populations assumed by this model, that of the RelLC+Cohort was not.

- Further investigation of the parameters of the RelLC +Cohort indicates that the over-smoothed fitted ratios can be linked to the presence of a book-specific non-parametric β_x^B which needs to be estimated from the book data. The estimation of this term requires large amounts of data, and, hence, with the relatively small population sizes of the book populations, the estimated β_x^B values tend to be erratic and lack robustness. In particular, the possibility that β_x^B fluctuates around 0 (see Figure 4.8) results in mortality differentials between the book and the reference cancelling out when aggregated measures of mortality such as survival probabilities and life expectancies are calculated. Given that this over fitting of the β_x^B may result in an inappropriate perfect correlation between the reference and the book populations, we consider that the RelLC+Cohort is inadequate for basis risk assessment. This conclusion extends to other models with non-parametric β_x^B parameters such as the Augmented Common Factor model and the Plat+Lee-Carter model.

- Some models, such as M7-Saint and M7-M7 show signs of potential over-fitting to the data.

The graphic testing of the goodness-of-fit of the models leaves us with six potential candidate models for basis risk assessment. These models are: CAE+Cohorts, Gravity,

M7-M5, M7-M6, M7-M7, and M7-Saint. The balance between goodness-of-fit and parsimony of these models is investigated in Table 4.5 where we show the AIC values[7] for the book part of each model when applied to the four sample schemes, together with the corresponding ranking across models (in brackets). From Table 4.5 we note the following:

- The CF+Cohorts, which is the simplest model among all the models fitted, tops the AIC ranking for three out of four datasets. However, as noted before, this model is not suitable for basis risk assessment since it assumes that the reference and book populations are perfectly correlated. One may nevertheless consider this model for other applications where the correlation between the populations is not important.

- Among all other models, the CAE+Cohort and M7-M5 show the best compromise between goodness-of-fit and parsimony, consistently ranking in the top three places and with very similar performances.

- M7-Saint and M7-M7, which have a quadratic age term in the book model, are always outperformed by the M7-M5 models. This suggests that when considering models from the CBD-Family it is necessary to allow for differences in level of mortality and a gradient by age, but that an additional parameter for a curvature by age is not necessary, i.e., it is sufficient to inherit the curvature from the reference population. Thus, we eliminate the M7-M7 and M7-Saint models from our list of candidate models.

- The Gravity model, M7-M6 and M7-M7, which have a book-specific cohort effect, have the worst trade-off between goodness-of-fit and parsimony. This suggests that we should generally reject models with a book cohort effect on grounds of parsimony. However, for the moment, we shall retain the Gravity model (two-population APC) which, among models with book-specific cohort effect, shows the best compromise between goodness-of-fit and parsimony. This will enable us to investigate how forecasts levels of uncertainty and hedge effectiveness may be impacted by allowing for a book-specific cohort effect.

[7]Recall, the AIC value is computed as $\text{AIC} = 2\nu_B - 2\mathcal{L}_B$ where $\mathcal{L}_B$ is the Binomial log-likelihood of the book part of the model under the assumption that the reference population is treated as a known offset and ν_B is the number of book-specific parameters of the models

TABLE 4.5: Effective number of parameters and AIC for the book part of different two-population models fitted to the four test books.

Model	Reference parameters	Book parameters	Typical Lives	Typical Amounts	Extreme Wealthy	Extreme Deprived
CF+Cohorts	185	30	7008 (1)	7001 (1)	6921 (1)	7146 (4)
CAE+Cohorts	185	59	7036 (2)	7026 (2)	6950 (2)	7130 (3)
Gravity	156	116	7090 (5)	7077 (5)	7010 (5)	7182 (6)
M7-M5	226	60	7043 (3)	7049 (3)	6971 (3)	7102 (1)
M7-M6	226	117	7106 (6)	7099 (6)	7033 (6)	7166 (5)
M7-M7	226	146	7123 (7)	7128 (7)	7052 (7)	7188 (7)
M7-Saint	226	90	7069 (4)	7074 (4)	6991 (4)	7117 (2)

4.6.2.5 Plausibility of forecast central trends and levels of uncertainty

So far, we have shortlisted the CAE+Cohorts, Gravity and M7-M5 based on their theoretical properties, practicality and goodness-of-fit performance. However, the outcome of a basis risk assessment exercise will be strongly driven by the expected level of uncertainty around the central forecast of the demographic and financial quantities underlying the index-based hedge. It is then crucial to check that these models produce reasonable forecast for both single and two-population metrics. This entails judging whether or not the forecasted central trajectories and patterns of uncertainty look plausible and are in line with historical variability.

Following Cairns et al. (2011b), we assess this property by examining fan charts of the forecasts produced by the models. Fan charts allow us to examine any distinctive visual feature of the forecasts of the models, as well as the differences between models. Each fan chart presents 95% prediction intervals and depicts the forecast output from the stochastic mortality models by also presenting 80% and 50% prediction intervals. In producing the model simulations underlying the fan charts, we have considered the following sources of uncertainty (risk):

- *Process risk* (PR) arising from the possible future trajectories of the time series of the period and cohort indices and
- *Parameter uncertainty* (PU) arising from the estimation of the parameters of the model.

Process risk is taken into account by simulating trajectories of the period and cohort indices[8], while parameter uncertainty is allowed for by using a Binomial adaptation of the residual bootstrapping approach proposed by Koissi et al. (2006)[9]. We note that due to the considerable exposure of the England and Wales population (chosen as the reference population), we deliberately ignore parameter uncertainty in the reference population.

Rather than analysing forecasts of mortality rates, we concentrate on the forecast of life expectancies and survival rates. According to Coughlan et al. (2011), these two aggregate quantities are more appropriate than individual mortality rates for gaining insight into the basis risk associated with longevity hedges. On the one hand, life expectancies and survival rates are more closely related to the hedge effectiveness objective than mortality rates, as, for instance, in a pensioner population life expectancy corresponds to the number of years over which a pension needs to be paid while survival rates correspond to the fraction of pensioners who are still alive to receive pension. On the other hand, these aggregate metrics smooth out most of the noise associated with individual mortality rates.

Figure 4.9 presents fan charts of 30 year curtailed period life expectancies at age 60 for the England and Wales reference population,

$$\overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(t) = \sum_{h=1}^{30} \prod_{i=0}^{h-1} (1 - q^{R}_{60+i,t}).$$

Figure 4.10 shows equivalent fan charts of 30 year curtailed period life expectancies at age 60 for the "Extreme Wealthy" test book, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(t)$, together with fan charts of the difference between the life expectancies in the book and the reference population, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(t) - \overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(t)$. The mean and variance of the forecasted period life expectancies in 2020, $\overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(2020)$ and $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2020)$, are presented in Table 4.6.

From Table 4.6 and Figures 4.9 and 4.10 we can conclude that:

- For all the models the central forecast and their levels of uncertainty for life expectancies in the reference and the book are reasonable and consistent. We

[8] To model process risk we use a multivariate adaptation of Algorithm 2 in Haberman and Renshaw (2009) without provision for parameter error. We note that Algorithm 2 in Haberman and Renshaw (2009) is itself an adaptation of the prediction interval approach of Cairns et al. (2006).

[9] We note that in adapting the bootstrap we follow Renshaw and Haberman (2008) and solve for the observed numbered of deaths instead of the fitted number of deaths as done by Koissi et al. (2006). The details of the residual bootstrapping approach under a Binomial framework are described in Debón et al. (2010, Section 3).

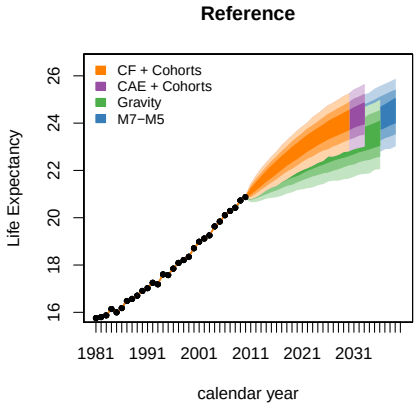


FIGURE 4.9: Fan charts of 30 year period curtailed life expectancy at age 60 for the England and Wales male reference population using different mortality models.

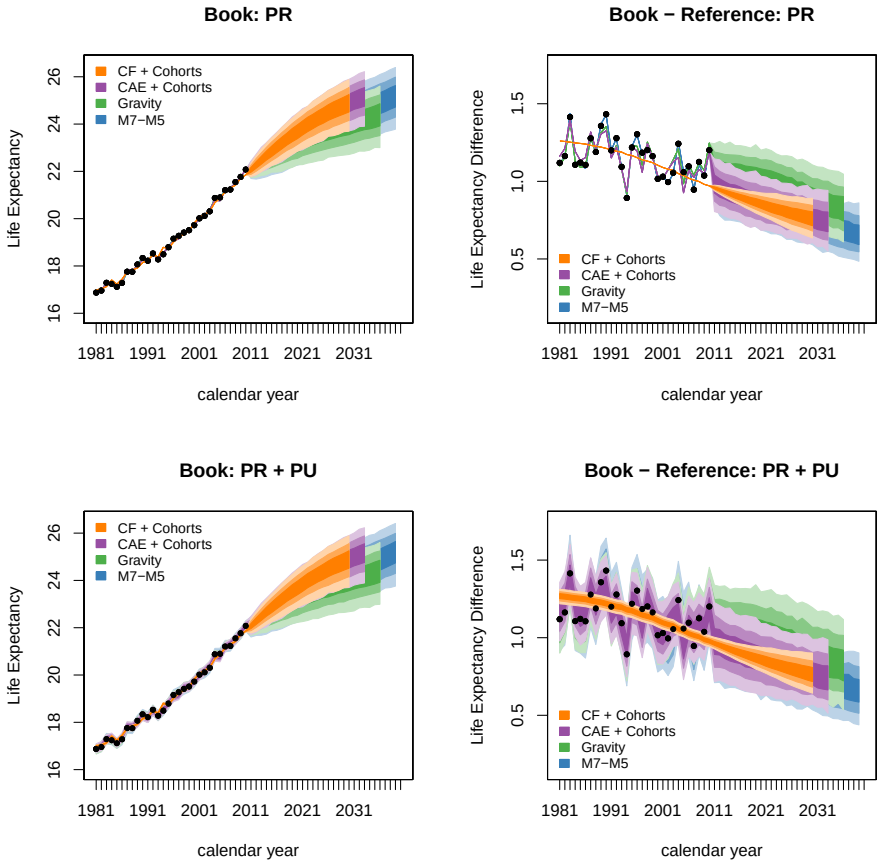


FIGURE 4.10: Fan charts of 30 year period curtailed life expectancy at age 60 for the "Extreme Wealthy" test book using different mortality models and different sources of risk (PR=process risk; PU=parameter uncertainty). Left panels present life expectancies for the book population and right panels results for the difference in curtailed period life expectancies in the book and the reference populations.

note however that there are noticeable differences between the models with the CAE+cohorts and CF+cohorts projecting longer life expectancies with slightly

TABLE 4.6: Forecasted mean and variance of 30 year period curtailed life expectancy at age 60 in 2020 for the England and Wales male reference population and for the "Extreme Wealthy" test book using different mortality models and different sources of risk (PR=process risk; PU=parameter uncertainty).

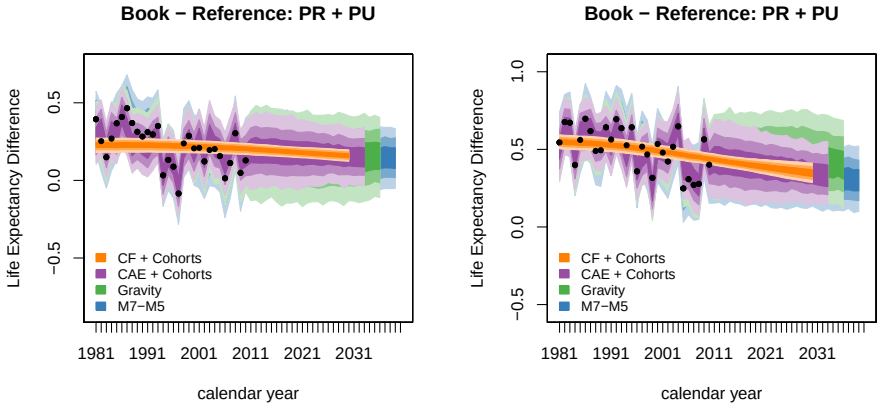
Mean of period life expectancy at age 60 in 2020					
		Book		**Difference**	
Model	**Reference**	**PR**	**PR+PU**	**PR**	**PR+PU**
CF+Cohorts	22.82	23.67	23.68	0.85	0.86
CAE+Cohorts	22.82	23.68	23.68	0.86	0.86
Gravity	22.12	23.14	23.13	1.02	1.01
M7-M5	22.33	23.19	23.19	0.86	0.86
Variance of period life expectancy at age 60 in 2020					
		Book		**Difference**	
Model	**Reference**	**PR**	**PR+PU**	**PR**	**PR+PU**
CF+Cohorts	0.2040	0.1719	0.1718	0.0014	0.0017
CAE+Cohorts	0.2040	0.1804	0.1878	0.0073	0.0135
Gravity	0.2251	0.1957	0.2065	0.0093	0.0206
M7-M5	0.2403	0.2161	0.2230	0.0091	0.0169

smaller uncertainty (narrower fan widths and smaller variances) than the other two models.

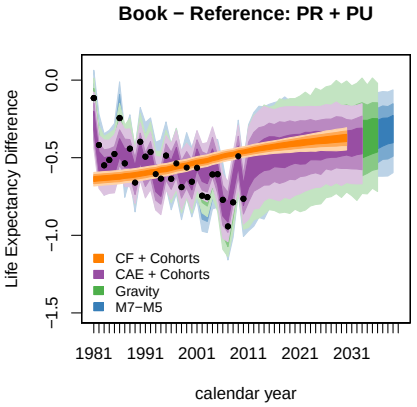
- The levels of uncertainty in the difference in life expectancy vary considerably across models and are on the low side when only process risk is contemplated. In particular, the unreasonably tight fan widths of the CF+cohorts confirm the issues with models assuming a perfect correlation between the reference and the book populations.

- The impact of parameter uncertainty on the life expectancies of the book population is minimal and almost unnoticeable to the eye. By contrast, parameter uncertainty increases significantly the levels of uncertainty of the life expectancy differences between the book and the reference populations and makes the prediction intervals start to look in-line with the historical volatility.

- In order to investigate the extent to which our conclusions generalise to the other test datasets, we show in Figure 4.11 fan charts of the difference in period life expectancies with respect to the England and Wales reference for the "Typical Lives", "Typical Amounts" and "Extreme Deprived" books. While the forecasts for the "Typical Lives" and "Typical Amounts" books look plausible, the forecasts for the 'Extreme Deprived" book look completely unreasonable, with the models



(A) Typical Lives (B) Typical Amounts



(C) Extreme Deprived

FIGURE 4.11: Fan charts of 30 year period curtailed life expectancy differences at age 60 between the different test books and the reference population including process risk (PR) and parameter uncertainty (PU).

failing to project the increase in life expectancy differences observed over the 1981-2010 period. Recalling Table 4.1, 60% of the "Extreme Deprived" population belong to the two most deprived quintiles of England which have seen a significant increase in relative mortality differentials with respect to England and Wales (see Chapter 2). This suggests that the non-divergence assumption embedded in the (vector) autoregressive process of order 1 used for forecasting the period index in the book $\boldsymbol{\kappa}_t^B$ (recall Equation (4.9)) may be inappropriate for the "Extreme Deprived" book.

Figure 4.12 presents the fan charts produced by each of the models for the value of a cohort survivor index,

$$S^R(65,t) = \prod_{i=0}^{t-1}(1 - q^R_{65+i,2011+i}),$$

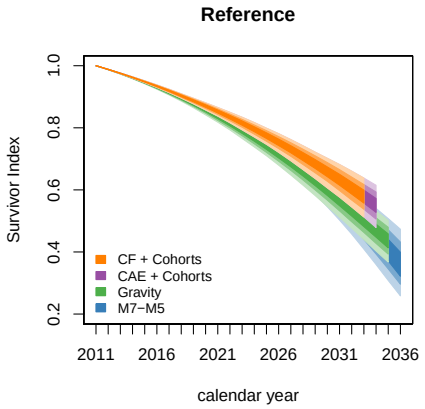


FIGURE 4.12: Fan charts of $S^R(65,t)$ for the cohort age 65 at the start of 2011 in the England and Wales male reference population using different mortality models.

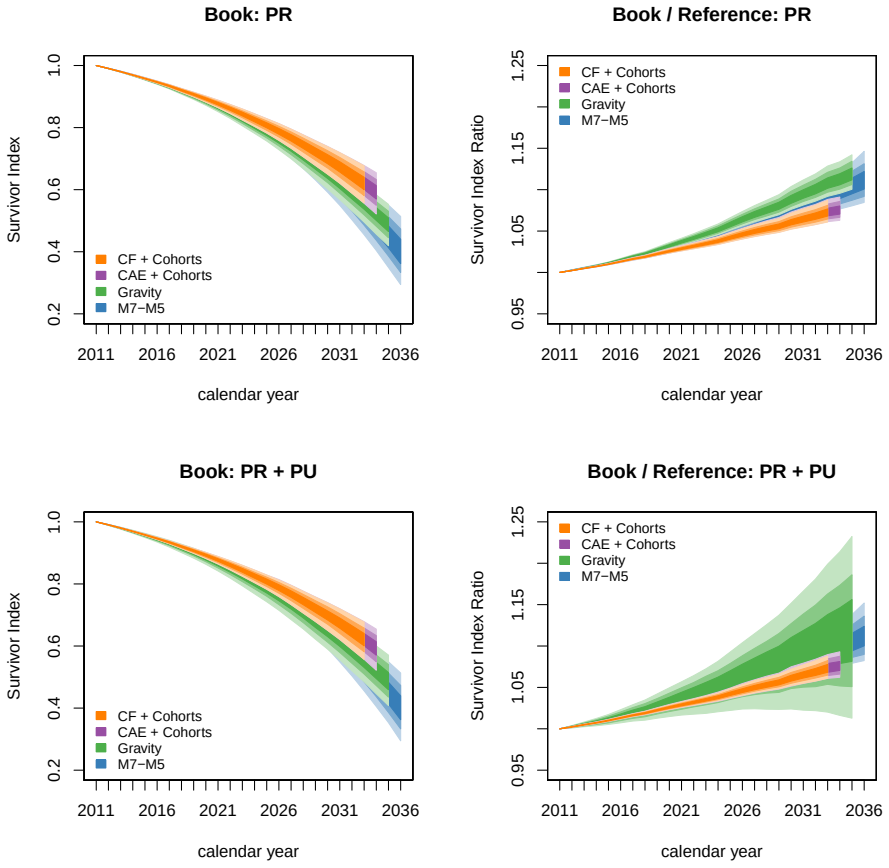
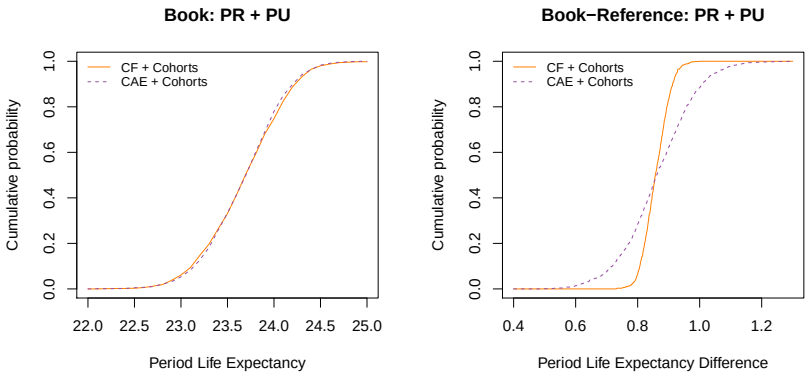
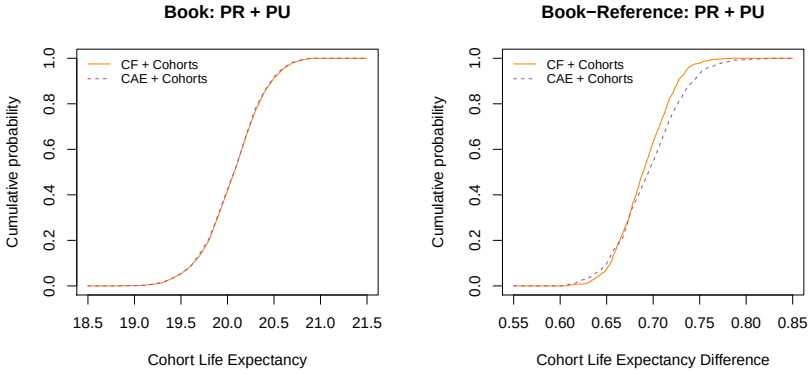


FIGURE 4.13: Fan charts of $S^B(65,t)$ for the cohort age 65 at the start of 2011 in the "Extreme Wealthy" test book using different mortality models and different sources of risk (PR=process risk; PU=parameter uncertainty). Left panels present the survivor index for the book population, $S^B(65,t)$, and right panels present results for the ratio of the survivor index in the book and the reference populations, $S^B(65,t)/S^R(65,t)$.



(A) 30 year curtailed period life expectancy at age 60 in 2020



(B) 25 year curtailed cohort life expectancy at age 65 in 2011

FIGURE 4.14: Cumulative distribution function of the curtailed period and cohort life expectancy in the "Extreme Wealthy" test book and cumulative distribution function of the corresponding difference in period and cohort life expectancies between the book and reference populations. The cumulative distribution functions account for both process and parameter risk.

computed using the forecasted mortality rates of the reference population. This index measures the proportion from a group of males aged 65 at the start of 2011 who are still alive at the start of year $2011+t$. We note that $S^R(65,t)$ does not involve any forecasts of the cohort effect as the relevant cohort effect, γ_{1946}, is known at the start of 2011. Figure 4.13 shows matching fan charts of the survivor index, $S^B(65,t)$, computed using mortality rates from the "Extreme Wealthy" test book, along with fan charts of the ratio of the book and reference population survivor indexes, $S^B(65,t)/S^R(65,t)$. From Figures 4.12 and 4.13, we note the following

- As with life expectancies, there are some differences in the trends and in the uncertainty levels of the survivor index of each of the two populations. Notably, M7-M5 produce wider fans than the other three models. This reflects the existence of

TABLE 4.7: Forecasted mean and variance of 25 year cohort curtailed life expectancy for the cohort aged 65 in 2011 for the England and Wales male reference population and for the "Extreme Wealthy" test book using different mortality models and different sources of risk (PR=process risk; PU=parameter uncertainty).

Mean of cohort life expectancy for the cohort aged 65 in 2011					
		Book		**Difference**	
Model	**Reference**	**PU**	**PU+PR**	**PU**	**PU+PR**
CF+Cohorts	20.36	21.05	21.05	0.68	0.69
CAE+Cohorts	20.36	21.06	21.06	0.69	0.70
Gravity	19.47	20.35	20.34	0.88	0.87
M7-M5	19.54	20.27	20.27	0.73	0.73
Variance of cohort life expectancy for the cohort aged 65 in 2011					
		Book		**Difference**	
Model	**Reference**	**PU**	**PU+PR**	**PU**	**PU+PR**
CF+Cohorts	0.1229	0.1062	0.1062	0.0007	0.0008
CAE+Cohorts	0.1229	0.1066	0.1066	0.0009	0.0014
Gravity	0.1252	0.1072	0.2027	0.0011	0.0905
M7-M5	0.1830	0.1670	0.1677	0.0009	0.0015

more random period effects in M7-M5 than in the CAE+cohorts, the CF+cohorts and the APC (Gravity) model.

- Parameter uncertainty has little impact on the level of uncertainty of the survivor index in the book. By contrast, parameter uncertainty can have a significant impact on the level of uncertainty observed in the ratio of survivor indexes. In particular, the right bottom panel of Figure 4.13 shows that for the APC, which is the only model that allows for a book specific cohort effect, the fans look completely unreasonable. This suggests that, unless there is strong reason to believe in the existence of a different cohort effect in the book to the reference population, the parameter uncertainty in fitting a book-specific cohort term will greatly outweigh any benefits in terms of goodness-of-fit to historical experience.

- The close alignment between the fans of models CF+cohorts and CAE+cohorts deserves further investigation. For these two models, which share the same reference population model, we plot in Figure 4.14 the simulated empirical distribution considering both process risk and parameter risk for the 30 year period curtailed life expectancy at age 60 in 2020, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30}|}(2020)$, and for the 25 year cohort life expectancy for someone aged 65 in 2011 in the book population,

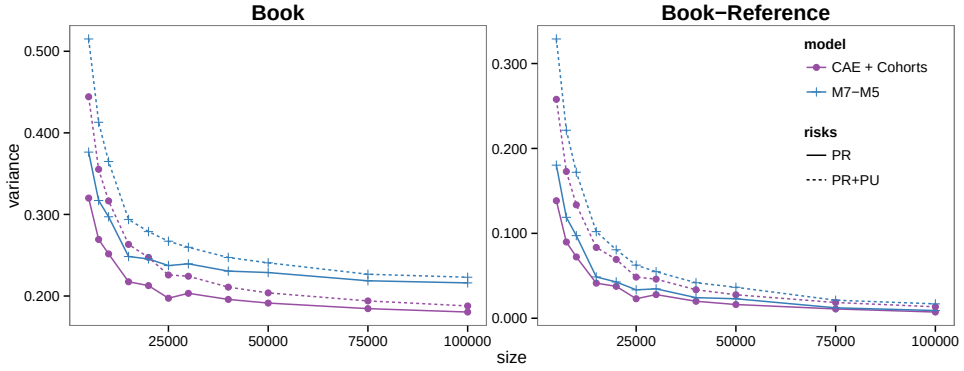
$$\overset{\nearrow}{e}{}^{B}_{65,\overline{25}|}(2011) = \sum_{t=1}^{25} S^{B}(65,t) = \sum_{t=1}^{25}\prod_{i=0}^{t-1}(1-q^{B}_{65+i,2011+i}),$$

together with the corresponding simulated empirical distribution of the difference in period and cohort life expectancies between the book and the reference populations. The associated means and variances are reported in Table 4.6 for the period life expectancies and in Table 4.7 for the cohorts life expectancies. While for the book population the empirical distributions are practically indistinguishable, there are notable dissimilarities in the empirical distributions for the difference in both period and cohort life expectancies, suggesting that the uncertainty in the book survivor index is dominated by the uncertainty in the reference part of the model. Furthermore, the discrepancies in the mean and variances of the life expectancy for the book population forecasted by both models are immaterial. These results allows us to conclude that although model CF+cohorts is unsuitable for basis risk assessment due to its implicit perfect correlation between the book and reference populations, this model might by a reasonable alternative in applications where only single population metrics are of interest such as when valuing pension liabilities or pricing annuities.

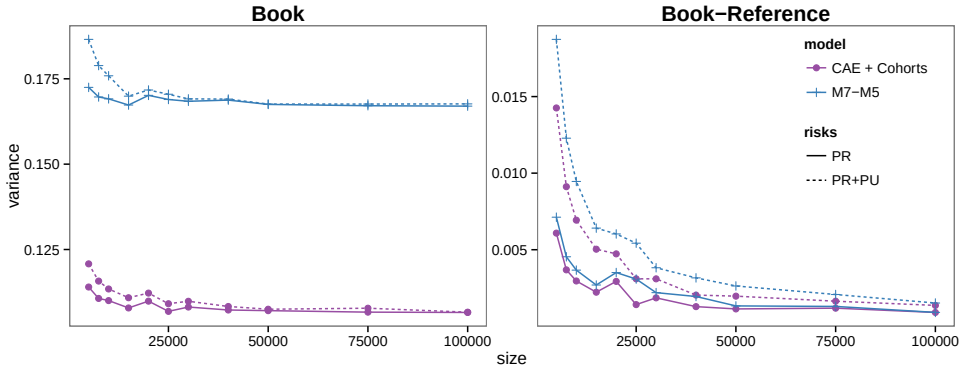
Overall, although all the models produce plausible trends and forecast levels of uncertainty for single population metrics, for two-population metrics only models CAE+Cohorts and M7-M5 produce plausible results. In addition, there are large enough differences between the models for us to acknowledge model risk as an important issue.

4.6.2.6 Forecast levels of uncertainty by book size

The analysis of the plausibility of the forecasted levels of uncertainty performed so far has been based on a fairly large book population with 100,000 exposed lives per year between ages 60 and 89. However, for smaller exposures of the book population the sampling noise in the data is bigger, leading to more uncertainty in the estimates of the parameters of the models. This additional variability arising from a smaller population size can potentially have a material impact on the plausibility of the forecasted levels of uncertainty. To explore this phenomenon, we investigate how the contribution of the different sources of uncertainty to the total level of risk varies by population size. Figure 4.15 shows for models CAE + Cohort and M7-M5 the variation by book size of:



(A) 30 year curtailed period life expectancy at age 60 in 2020



(B) 25 year curtailed cohort life expectancy at age 65 in 2011

FIGURE 4.15: Variance by population size of the curtailed period and cohort life expectancy in the "Extreme Wealthy" test book and variance by population size of the corresponding difference in period and cohort life expectancies between the book and reference populations using different models and considering different sources of risk.

- the variance of the 30 year period curtailed life expectancy at age 60 in 2020 for the "Extreme Wealthy" test book, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2020)$ (top-left panel); and
- the variance of the corresponding differences in period life expectancies with respect to the England and Wales reference population, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2020) - \overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(2020)$ (top-right panel);
- the variance of the 25 year cohort curtailed life expectancy at age 65 in 2011 for the "Extreme Wealthy" test book, $\overset{\nearrow}{e}{}^{B}_{65,\overline{25|}}(2011)$ (bottom-left panel); and
- the variance of the corresponding differences in cohort life expectancies with respect to the England and Wales reference population, $\overset{\nearrow}{e}{}^{B}_{65,\overline{25|}}(2011) - \overset{\nearrow}{e}{}^{R}_{65,\overline{25|}}(2011)$ (bottom-right panel);

From this figure we can see how:

- The differences in the levels of uncertainty produced by the models are evident, with the M7-M5 producing higher variance than the CAE+Cohorts. These differences are particularly notable for cohort life expectancies in the book population.
- The magnitude of the variance of both period and cohort life expectancies starts to stabilise around a book size of 25,000 lives. This is particularly noticeable when considering only process risk.
- For book sizes smaller than 15,000 lives, process risk is unrealistic producing artificially high variances.

These observations suggest that to avoid a distorted assessment of the levels of uncertainty, models CAE+Cohorts and M7-M5 should only be used when the book exposure is higher than 20,000-25,000 lives. Furthermore, as we show in section 4.7, insisting on using the models with modest exposure numbers may result in a misstated assessment of demographic basis risk.

4.6.2.7 Forecasting performance and robustness

A good mortality model should not only produce forecasts that appear reasonable ex-ante, but should also provide good ex-post forecast, that is forecast that do not deviated significantly from realised outcomes. In addition, these forecasts should be robust relative to the choice of period for the data employed in producing the forecasts.

To assess the forecasting accuracy of the models, we first carry out a backtesting exercise in the spirit of Booth et al. (2006) and Jarner and Kryger (2011, Section 4). This exercise entails the fitting and forecasting of the models using data for the period 1981 to 2010 for different history lengths, book sizes, and IMD compositions in the book population; and the evaluation of different metrics of forecasting performance.

Specifically, the different models were fitted to history lengths ranging from 5 years to 20 years[10], book sizes ranging from 5,000 lives to 100,000 exposed lives between ages 60 to 89 and the four test IMD compositions described before in Table 4.1. The forecasting

[10]For instance when considering a history length of 5 years the models were fitted using data for the book population covering the periods 1981-1985, 1982-1986, ..., 2004-2008, 2005-2009. In all cases, the reference population data was assumed to start in 1961 and end in the same year as the book population data.

performance of the models is evaluated by comparing the actual 30 year curtailed period life expectancies at age 60 in the book population, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30}|}(t)$, and the actual differences in 30 year curtailed period life expectancies at age 60 between the book and the reference population, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30}|}(t) - \overset{\uparrow}{e}{}^{R}_{60,\overline{30}|}(t)$, with their corresponding predicted counterparts over the rest of the period until 2010. Forecast bias (actual-fitted) is summarised by averaging across years, book sizes and forecasting horizon. The matching absolute errors are also averaged to provide a measure of forecast accuracy.

The forecast bias (mean errors) and the forecast accuracy (mean absolute error) for both period life expectancy in the book and differences in period life expectancy between the book and the reference, plotted against history length are shown in Figure 4.16. We note the following:

- Models CAE+cohorts and CF+Cohorts stand out as the best models for forecasting period life expectancies in the book with the smallest bias and with the smallest mean absolute error. The close alignment between the mean errors and mean absolute errors of these two models reflects the fact that they share the same reference population model.

- History length has a material impact on the out-of-sample performance of the models. With the exception of model CF+Cohorts which does not require the forecasting of any book specific time index, the forecasting performance of the models for history lengths shorter than 8 years is poor. The noticeably poorer performance of model M7-M5 for the shorter history lengths is explained by the fact that this model has two period indices for the book, implying a more complex and data demanding time series process for the forecasting.

- For differences in life expectancies and when we have more than 8 years of history, the models perform very similarly both in terms of bias and accuracy.

- The bias in forecasting differences between the "Extreme Deprived" population and the England and Wales reference population is considerably higher than the bias for the other three test book compositions. This higher bias gives further evidence for concluding that the non-divergence assumption may be inadequate for the "Extreme Deprived" book.

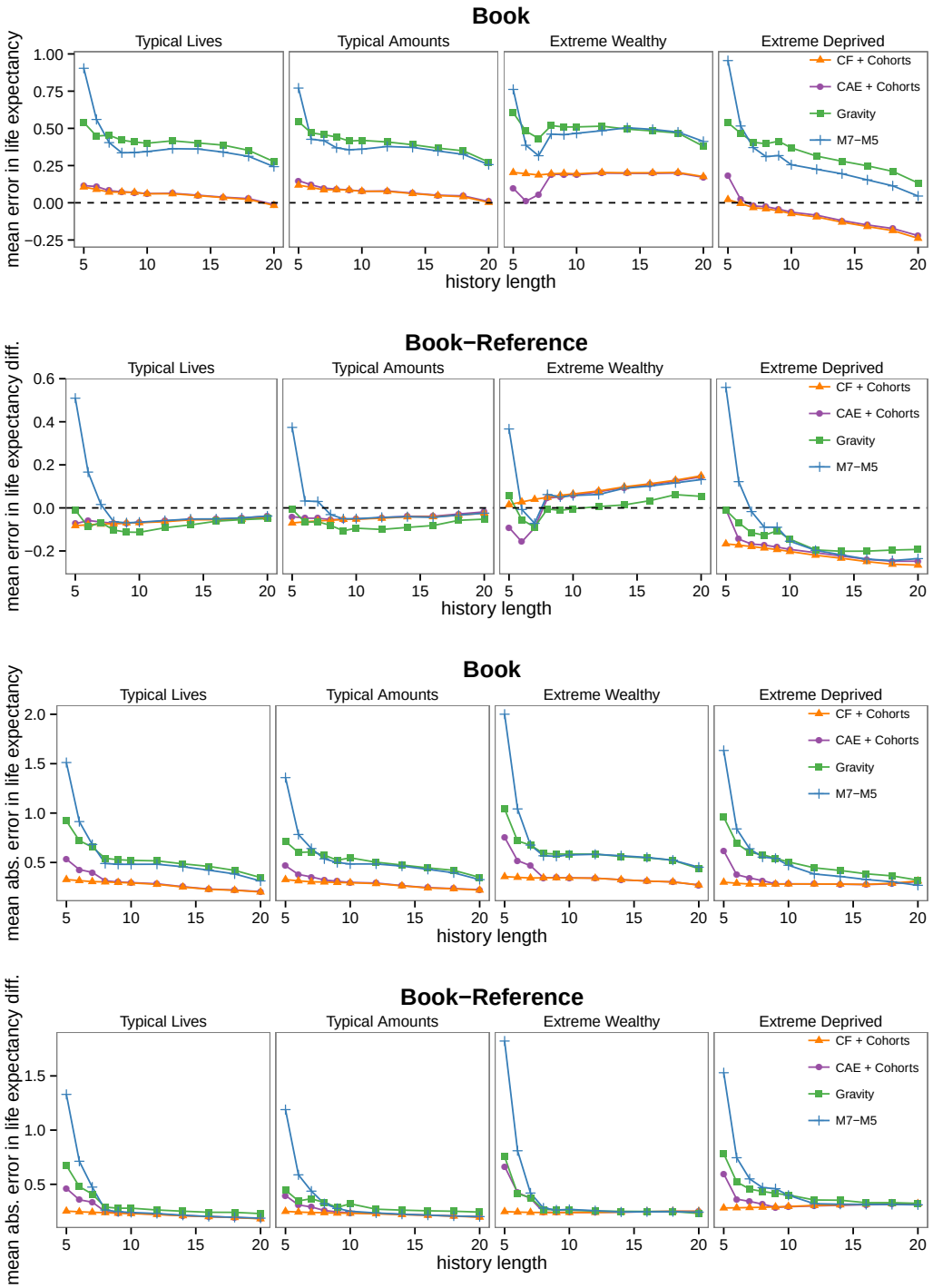


FIGURE 4.16: Mean error (actual - fitted) and mean absolute error in the forecast of 30 year period curtailed life expectancy at age 60 in the book and in differences in 30 year period curtailed life expectancy at age 60 between the book and the reference. The results are averaged across years, book sizes and forecast horizons ranging from 1 year to 15 years ahead.

In order to check the robustness of the models we will examine the stability of forecasts towards the inclusion of additional years at the right end of the data window, using a contracting horizon backtest as proposed by Dowd et al. (2010). Figure 4.17 shows plots

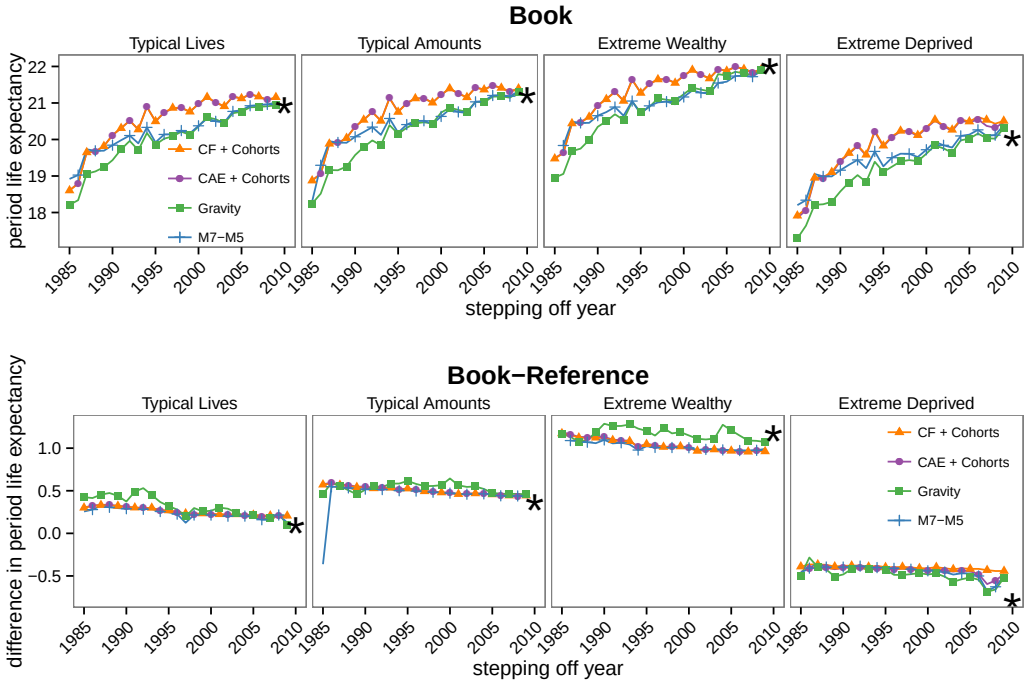


FIGURE 4.17: Forecast of the 30 year period curtailed life expectancy at age 60 in 2010 for the four different book populations using different fitting periods. The stepping-off year is the final year used in fitting the models. The realised life expectancy for 2010 is represented by a star.

of forecasts of the 30 year period curtailed life expectancy at age 60 in 2010, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2010)$, for the four test book populations with 100,000 exposed lives, made in 1985, 1986, ..., 2009. Equivalent plots for the difference between the book and reference population, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2010) - \overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(2010)$, are also included. For all book populations and models, we see that the forecasts for the book population are well-behaved, in the sense that they converge in a stable manner towards the realised outcome. For forecasts of differences in life expectancy we see a similar stable behaviour, albeit the discrepancies in the forecasts for the "Extreme Deprived" population are noticeable and consistent with our previous findings regarding the unsuitability of the non-divergence assumption for this book (recall Figure 4.11c). Similar analysis for other book sizes show that the models are robust, provided that the length of the fitting period is longer than 10 years (i.e. for stepping off years after 1990).

4.6.2.8 Summary of assessment

The main conclusions of our systematic assessment of the candidate two-population mortality models for basis risk assessment can be summarised as follows:

- We should not expect any single model to satisfy all possible desirable criteria of a practical solution to assessing basis risk.
- However, models M7-M5 and CAE+Cohorts stand out as the models which provide the most suitable balance between flexibility, simplicity, parsimony, goodness-of-fit to data, forecasting performance and robustness.
- Both M7-M5 and CAE+Cohorts produce reasonable best estimate projections with plausible levels of uncertainty, but with sufficient differences which suggests that we need to recognise model risk as an important issue.
- The paucity of book data means that it is difficult to estimate the age modulating terms $\beta_x^{(j,B)}$ without resulting in non-robust and erratic parameter estimates. Thus, any parameter which moderates the sensitivity of the book to time trends at different ages should be inherited from the reference population (i.e. $\beta_x^{(j,B)} \equiv \beta_x^{(j,R)}$)
- Unless there is strong reason to believe in the existence of a different cohort effect in the book to the reference population, the parameter uncertainty in fitting a (non-parametric) book-specific cohort term will greatly outweigh any benefits in terms of goodness-of-fit to historical experience.
- The fitting of two-population models should in principle only be pursued when:
  - the book annual exposure is over 20,000-25,000 lives– for smaller exposures, the impact of parameter uncertainty may result in a distorted assessment of basis risk;
  - there are at least 8-10 years of reliable book data- for shorter history lengths, the quality of the forecasts is likely to be poor.

The above conclusions are underpinned by this analysis based on England and Wales population data and the profile of sample schemes drawn from the Club Vita database.

We would expect many of the key conclusions to hold for other populations, although specific results (such as AIC rankings) are necessarily dependent on the choice of data.

4.7 Quantifying basis risk

In this section we examine the performance of the models when used for assessing basis risk. We also discuss the impact that different volumes of data may have on the parameter uncertainty and on the assessment of basis risk. In presenting the basis risk analysis we follow the five-step hedge effectiveness framework proposed in Coughlan et al. (2011).

4.7.1 Steps 1 and 2: Hedging objectives and hedging instruments

Most hedging exercises will either be *cash flow hedges* or *value hedges* aiming, respectively, to mitigate the variability of the *cash flows* or the variability of the *value* of these cash flows. We consider thus two separate simple examples reflecting the objectives of a *value hedge* and of a *cash flow hedge*:

Value hedge example. Noting that period life expectancy corresponds to an annuity value using a zero percent interest rate and no mortality improvements, for the value hedge case we assume that the quantity at risk to be hedged is the 30 year curtailed period life expectancy at age 60 in 2020, $\overset{\uparrow}{e}{}^{B}_{60,\overline{30}|}(2020)$, i.e. over a horizon of 10 years. The hedging instrument to reduce the liability risk is the 30 year curtailed period life expectancy at age 60 in 2020 for the England and Wales population. $\overset{\uparrow}{e}{}^{R}_{60,\overline{30}|}(2020)$. This exercise is similar in spirit to the hedge effectiveness analysis performed in Cairns (2013) and Cairns et al. (2014).

Cash flow hedge example. To reflect a cash flow hedge situation we consider that the quantity at risk to be hedged is the 25 year curtailed cohort life expectancy at age 65 in 2011, $\overset{\nearrow}{e}{}^{B}_{65,\overline{25}|}(2011)$, which can be interpreted as the arithmetic sum of the cash flows payable in 25 years to a pensioner aged 65 in 2011 and who belongs to a pension plan that pays £1 at the end of each year. The hedging instrument to reduce the liability risk is the 25 year curtailed cohort life expectancy at age 65 in 2011 for the England and Wales population. $\overset{\nearrow}{e}{}^{R}_{65,\overline{25}|}(2011)$. This exercise is similar in spirit to the hedge effectiveness analysis performed in Li and Hardy (2011).

Although very simple, these two examples should be informative of the performance of the models for hedge effectiveness assessment while avoiding the idiosyncrasies of specific pension benefit structures or more realistic hedging instruments.

4.7.2 Step 3: Method for hedge effectiveness assessment

Following Li and Hardy (2011), Cairns (2013) and Cairns et al. (2014), we use the variance as our measure of risk. Therefore, if L denotes the random unhedged liability and H represent the value of the index-linked hedging instrument, we assume that the hedger wishes to minimise the variance of $L - hH$, where h is the number of units (hedge ratio) held of the hedging instrument. We define thus the relative risk reduction (hedge effectiveness) as

$$R^2(h) = 1 - \frac{\text{var}(L - hH)}{\text{var}(L)}$$

It can be proved (see, for example, Cairns et al. (2014)) that the optimal hedge ratio is

$$h^* = \frac{\text{cov}(L, H)}{\text{var}(H)}$$

with optimal relative risk reduction

$$R^2(h^*) = 1 - \frac{\text{var}(L - h^*H)}{\text{var}(L)} = \rho^2,$$

where ρ is the correlation between L and H.

For our examples, it will thus suffice to analyse the correlation between L and H, i.e., the correlation between $\overset{\uparrow}{e}{}^{B}_{60,\overline{30|}}(2020)$ and $\overset{\uparrow}{e}{}^{R}_{60,\overline{30|}}(2020)$ for the value-hedging example and the correlation between $\overset{\nearrow}{e}{}^{B}_{65,\overline{25|}}(2011)$ and $\overset{\nearrow}{e}{}^{R}_{65,\overline{25|}}(2011)$ for the cash flow-hedging example.

4.7.3 Step 4: Calculation of hedge effectiveness

For each stochastic two-population model under consideration we compute the correlations between L and H based on 1,000 simulated mortality scenarios. In the analysis that follows, we contemplate three cases concerning the sources of risks considered in the simulations: i) only process risk (PR); ii) process risk and parameter uncertainty (PR+PU); and iii) process risk, parameter uncertainty and sampling risk (PR+PU+SR).

Process risk and parameter risk are considered using the techniques described before in Section 4.6.2.5, while sampling risk is considered by randomly sampling the number of deaths from a Binomial distribution once parameter uncertainty and process risk have been taken into account.

Specifically, for the value-hedging example, we assume that the future exposures $E^B_{x,2010+t}$, $t = 1, .., 10$, are equal to the average age-specific book exposure over the data period used in fitting the mortality model and then simulate the number of deaths using the conditional Binomial assumption:

$$D^B_{x,2010+t} | q^B_{x,2010+t} \sim \text{Bin}(E^B_{x,2010+t}, q^B_{x,2010+t}), \quad t = 1, \ldots, 10$$

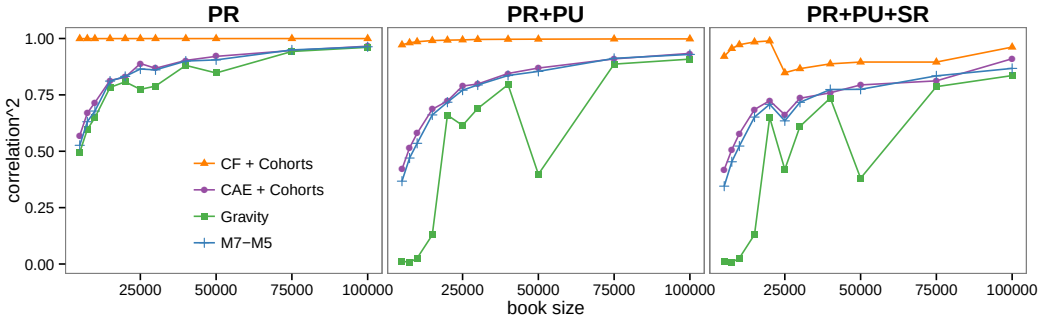
For the cash-flow hedge example, we take sampling risk into account by treating the cohort of pensioners as a random survivorship group. Thus, if l_x denotes the number of pensioners who survive to age x and given a simulated mortality scenario $\{q^B_{65,2011}, q^B_{66,2012}, \ldots, q^B_{89,2035}\}$, we model sampling risk with the following Binomial death process:

$$l_{65+t} \sim \text{Bin}(l_{65+t-1}, 1 - q^B_{65+t-1,2010+t}), \quad t = 1, \ldots, 25,$$

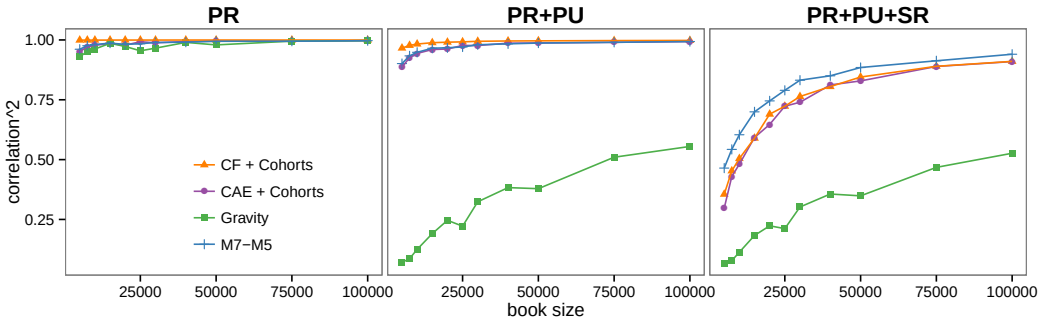
with l_{65} equal to 5% of the total exposure in the book between ages 60 to 89 (e.g. if the total book exposure is 100,000 lives we take $l_{65} = 5,000$). We have chosen 5% as from the total English male population aged 60 to 89 broadly 5% is aged 65 (see Figure 4.2).

4.7.4 Step 5: Interpretation of results

We concentrate on hedge effectiveness results for models CAE+Cohorts and M7-M5 which where identified as the best performing models after the systematic model assessment carried out in section 4.6. However, in spite of the implausible projections produced by models CF+Cohorts and APC (gravity) model, we shall also compute hedge effectiveness metrics for these two models to illustrate the issues that may arise if we insist on using them for basis risk assessment.



(A) Value hedge example: 30 year curtailed period life expectancy at age 60 in 2020



(B) Cash flow hedge example: 25 year curtailed cohort life expectancy at age 65 in 2011

FIGURE 4.18: Squared correlation, ρ^2, between the liability L and the hedging instrument H, as a function of book population size. All values correspond to the "Extreme Wealthy" socio-economic composition and using data for the period 1981-2010 to fit the models.

4.7.4.1 Hedge effectiveness by book population size

As discussed in section 4.6.2.5, population size has a material impact in the parameter uncertainty of the models. In addition, it is expected that the higher sampling risk associated with smaller populations will reduce the effectiveness of a standardised longevity-hedge. To investigate this phenomenon, we present in Figure 4.18 hedge effectiveness results (ρ^2) for the "Extreme Wealthy" test book considering population sizes ranging from 5,000 to 100,000 exposed lives between ages 60 to 89, and considering different sources of risk. In all cases, data for the period 1981-2010 (i.e. a history length of 30 years) is used for fitting the models. From Figure 4.18 we note the following:

- As expected the inherent perfect correlation of the CF+Cohorts results in an unrealistic zero or close to zero basis risk when sampling risk is ignored.
- The previously raised issues in relation to the parameter uncertainty in the estimation of book-specific cohort parameters become evident, with the APC (Gravity)

model showing implausibly low hedge-effectiveness once parameter risk is taken into account. This is especially noticeable for the cash flow-hedge example which involves cohort-type quantities. In this case, even for populations as big as 100,000 lives, the hedge effectiveness produced by the model are below 60% while other models produce hedge-effectiveness of more than 85%.

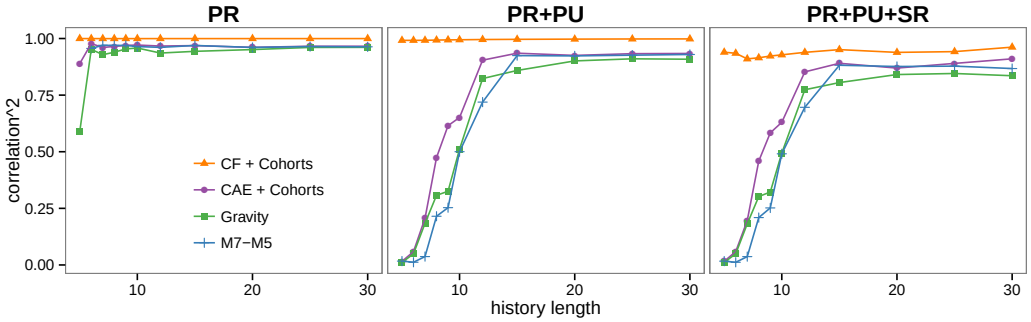
- For book sizes smaller than 15,000 lives, process risk is unrealistically high (recall Figure 4.15) distorting the assessment of basis risk and producing artificially low hedge effectiveness for the value-hedge example.

- The impact of sampling risk on correlations is material, with hedge effectiveness values for both the value-hedge and the cash flow-hedge falling rapidly for populations below 10,000 lives. We also note that for the cash flow-hedge example, sampling risk is the main determinant of basis risk. In fact, the CF+Cohorts model which implies zero basis risk before accounting for sampling risk, results in virtually the same risk reductions as models CAE+Cohort and M7-M5 once sampling risk is accounted for.

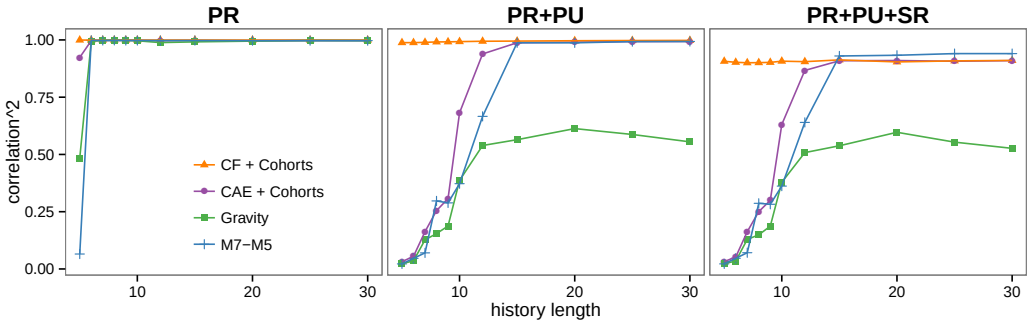
- Although models M7-M5 and CAE+Cohorts can give rather different mortality forecasts, these differences seem to attenuate in applications, with the two models producing very similar hedge effectiveness values once all risks have been taken into account.

4.7.4.2 Hedge effectiveness by history length

We now investigate how the number of years of available data in the book population impacts the evaluation of hedge effectiveness. Figure 4.19 presents hedge effectiveness results (ρ^2) for the "Extreme Wealthy" test book considering history lengths ranging from 5 years to 30 years, and considering different sources of risk. In all cases a book populations size of 100,000 lives is used for fitting the models. In this figure we can see how history length has a significant impact on hedge effectiveness assessment. For history lengths shorter than 10-12 years and once parameter uncertainty has been considered, risk reductions fall rapidly for model CAE+Cohorts, M7-M5 and APC (Gravity). This reinforces the previously discussed issues of fitting time series models when historical data are scarce.



(A) Value hedge example: 30 year curtailed period life expectancy at age 60 in 2020



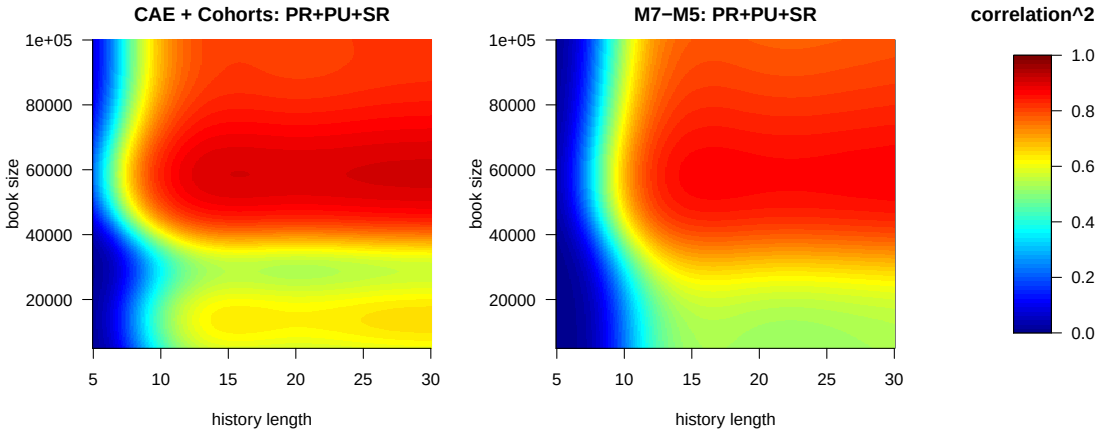
(B) Cash flow hedge example: 25 year curtailed cohort life expectancy at age 65 in 2011

FIGURE 4.19: Squared correlation, ρ^2, between the liability L and the hedging instrument H, as a function of history length of the book population data used to fit the models. All values correspond to the "Extreme Wealthy" socio-economic composition with a book size of 100,000 annual exposed lives between ages 60 to 89.

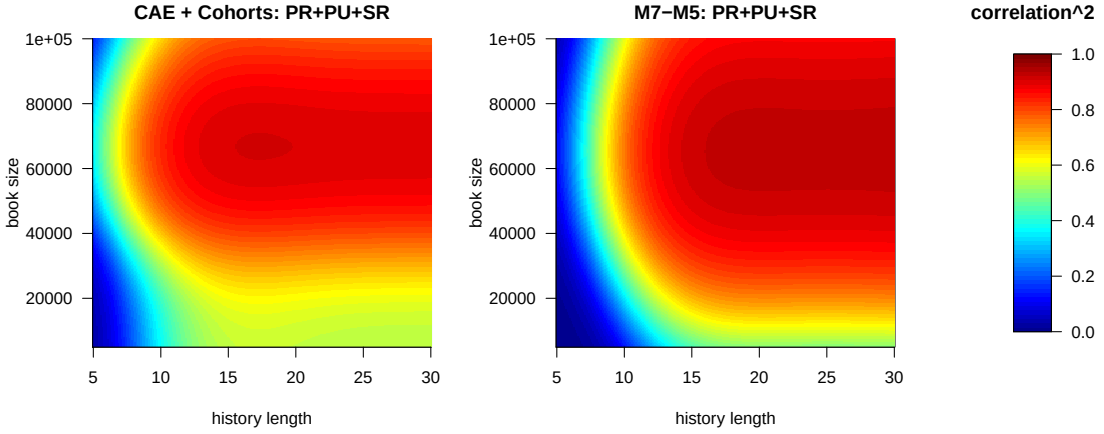
4.7.4.3 Interaction between book size and history length

We end this section by investigating the interaction between book size and history length in the assessment of hedge effectiveness. To do so, we have fitted the CAE+Cohort and M7-M5 model to the "Extreme Wealthy" test book considering all possible combinations between book sizes 5000, 7500, 10000, 15000, 20000, 25000, 30000, 40000, 50000, 75000, and 100000 and history lengths 5, 6, 7, 8, 9, 10, 12, 15, 20, 25 and 30 years. Figure 4.20 presents heatmaps depicting for both the value-hedge and cash-flow hedge examples the resulting hedge effectiveness values (ρ^2) once all sources of risks have been considered. To ease the identification of patterns, correlations have been smoothed along the book size and history length dimensions[11]. From Figure 4.20 we note the following:

[11] Smoothing has been performed using a generalised additive model of the form $\text{logit}\,\rho = s(size) + s(length)$, where, s denotes a penalised spline. For smoothing we have used the **R** package **mgcv** (Wood, 2015).



(A) Value hedge example: 30 year curtailed period life expectancy at age 60 in 2020



(B) Cash flow hedge example: 25 year curtailed cohort life expectancy at age 65 in 2011

FIGURE 4.20: Smoothed squared correlation, ρ^2, between the liability L and the hedging instrument H, as a function of the book size and history length of the book population. All values correspond to the "Extreme Wealthy" socio-economic composition.

- The interaction between history length and book size is minimal, with hedge effectiveness falling rapidly for history lengths shorter than 10-12 years and book sizes smaller than 15,000-25,000 exposed lives above age 60.

- While for model M7-M5 correlations start to fall significantly for history length below 12 years, for the the CAE+cohorts correlations only start to show a material decline for history length below 10 years. This suggest that when historical data are limited, models with fewer book specific period indexes should be preferred over models with multiple period book specific period terms. However, in all cases the fitting of two-populations models should only be pursued when book data exceeds 8-10 years of history.

- For a book size over 25,000 lives and history length above 12 years, the hedge effectiveness reductions for both the value-hedge and the cash-flow hedge examples are above 70%, suggesting that index-based hedges can be a meaningful alternative for hedging longevity risk.

4.8 Discussion

We end this chapter by looking at a number of general comments arising from our investigations.

Our previous sections have suggested that M7-M5 or the CAE+cohorts are appropriate models when undertaking the modelling of the mortality of the reference and the book populations in a basis risk assessment exercise. However, this need not preclude the consideration of additional models. Indeed, the modeller may wish to look at alternative models as part of sensitivity testing; or in order to gain a better understanding of model risk; or to err on the side of adding more features into the model than historic back-testing alone might suggest –these features may be needed as part of a personal belief on the complexity of mortality. Further as time goes on new models will enter the actuarial literature and our work can help integrate those models into a basis risk assessment context. Therefore, we next provide some general guidelines for the construction of two-population models for basis risk assessment.

When building a two-population model for assessing longevity basis risk, it is usual to find that the reference population is considerably larger and has a longer back history of data than the book population. It is therefore natural to start by selecting an appropriate model for the reference population. Once the reference population model is chosen a reasonable approach would be to select the book part of the model from within the same model family of the reference part. This will ensure a correspondence between the model parameters in the book and the reference populations which facilitates interpretation of the parameters of the models and makes the subsequent analysis more comprehensive and consistent in both populations. Our research on different models has also identified that:

- it is in general enough to include at most two book-specific time-dependent terms;

- any parameter which moderates the sensitivity of the book to these time trends at different ages should be inherited from the reference book (i.e. $\beta_x^{(j,B)} \equiv \beta_x^{(j,R)}$); and
- it is generally appropriate not to include a book specific cohort effect.

In mathematical terms, if the preferred reference population model is given by

$$\text{logit}\, q_{xt}^R = \alpha_x^R + \sum_{j=1}^{N} \beta_x^{(j,R)} \kappa_t^{(j,R)} + \gamma_{t-x}^R,$$

then a good starting point for the book model would in general be of the form:

$$\text{logit}\, q_{xt}^B - \text{logit}\, q_{xt}^R = \alpha_x^B + \sum_{j=1}^{M} \beta_x^{(j,R)} \kappa_t^{(j,B)}$$

We would usually expect M to be at most two as it is unlikely that the book population can support more than two time series i.e. $M \leq \min(2, N)$.[12] By way of example, if we choose to model the reference population using the single population model described in Börger et al. (2013):

$$\text{logit}\, q_{xt}^R = \alpha_x^R + \kappa_t^{(1,R)} + (x - \bar{x})\kappa_t^{(2,R)} + (x_{\text{young}} - x)^+ \kappa_t^{(3,R)} + (x - x_{\text{old}})^+ \kappa_t^{(4,R)} + \gamma_{t-x}^R$$

where x_{young} and x_{old} are predefined constant, then a suitable starting point for the book model would be

$$\text{logit}\, q_{xt}^B - \text{logit}\, q_{xt}^R = \alpha_x^B + \kappa_t^{(1,B)} + (x - \bar{x})\kappa_t^{(2,B)}$$

Our systematic analysis of the two-population mortality literature has focused exclusively on the use of these models for the assessment of basis risk in longevity hedges. Furthermore, we have implicitly assumed that the target book population is a subset or is closely related to the reference population on which the index is based. Hence, our conclusions may not necessarily extend directly to other applications of two-population mortality models and the evaluation of the suitability of a model will largely depend

[12] Note that the M7-M5 model and the CAE+cohorts can be derived from this form by applying the previous rules if we start by modelling the reference population using a M7 model or a LC+Cohorts model, respectively.

on the task at hand (e.g whether it is a basis risk assessment exercises or not) and on the nature of the relationship between the two populations being modelled. As we have repeatedly discussed in this chapter, simpler models that are not suitable for basis risk assessment (e.g, because of their implied perfect correlation between the populations) may be suitable for other applications such us when valuing pension liabilities or pricing annuities. In addition, the use of two-population mortality models for assessing the basis risk in longevity hedges where the mortality in one country is hedged with the mortality of another country[13] would require a deep understanding of the differences between the two countries' mortality. Such differences may not be captured by the structure of the two-population models we have proposed and the relative approach we have pursued may have to be substituted by a simultaneous modelling of the two countries' mortality, for instance along the lines of the work of Li et al. (2015) or the GLM modelling approach of Hatzopoulos and Haberman (2013).

In all our mortality projections and simulations we have employed the usual assumption that the spread between the mortality in the reference and the book will conform to the non-divergence hypothesis in the long run, i.e., that the ratio of q^B_{xt}/q^R_{xt} will tend to a constant value as $t \to \infty$. We have captured this non-divergence constraint via the use of a (vector) autoregressive process for the time series indices (κ^B_t) in the book part of the model, implying that in the long-run the spread between the logit of mortality for the book and the reference population will revert from the current level to the historical mean. Although the investigation of the appropriateness of this assumption is out of the scope of this chapter, the unsatisfactory results we have obtained when modelling the "Extreme Deprived" book population suggest that such an assumption may not be appropriate in all cases. In addition, the non-divergence assumption implies that the variance of the difference in (logit) mortality between the book and the reference population is bounded, potentially understating demographic basis risk and hence overstating the hedge effectiveness. We thus encourage further research looking at alternative choices of times series model and at the implications that such choices may have on hedge-effectiveness.

Our investigations indicate that the accurate calibration and projection of a two-population model requires that the annual exposure in the book population is over 20,000-25,000

[13] An example of this is the Kortis bond where UK mortality is hedged using US mortality, see Hunt and Blake (2015)

lives and that there are at least 10-12 years of reliable book data. However, in practice a large proportion of pension scheme books and life company portfolios will not meet these data requirements leaving open the question of how to assess longevity basis risk and hedge effectiveness for such populations. If book size is the main issue, then a Bayesian approach such as those considered in Cairns et al. (2011a) and in Antonio et al. (2015) may offer an alternative. By contrast, if the problem is the lack of sufficiently long historical data, indirect approaches where the book is modelled indirectly by reference to a bigger population with a more reliable and longer mortality experience could be the way through. Such approaches have recently been considered in the "mixing" approach proposed by Ahcan et al. (2014) and in the "characterisation" approach introduced in Haberman et al. (2014b), and we believe that this line of research deserves further consideration.

Finally, although we have only considered very stylised longevity hedges, our hedge-effectiveness results show that index-based hedge have the potential to provide an effective and flexible solution to mitigate longevity risk. We hope that our research has shed light on the assessment of basis risk and contributes to moving forward the market of standardised longevity transactions.

Acknowledgements

Part of this chapter has previously been published in the research report Haberman et al. (2014b) produced by a joint team from Cass Business School and Hymans Robertson LLP in response to research commissioned by the Longevity Basis Risk Working Group of the Institute & Faculty of Actuaries and the Life & Longevity Markets Association. We thank Steven Baxter, Andrew Gaches, Sveinn Gunnlaugsson and Mario Sison from Hymans Robertson LLP for their significant contribution to this research. We also thank the Longevity Basis Risk Working Group and Andrew Hunt for their feedback which has helped improved this work.

Appendix 4.A Generation of synthetic data

In this appendix we present a procedure for generating, based on a reference dataset, synthetic mortality datasets which have a given exposure size with a given distribution of this exposure across population subgroups.

Assume that we have a reference dataset containing observed number of deaths D_{xtg} in year t for people age x in subgroup g with matching central exposures E_{xtg} and matching death rates $\mu_{xtg} = D_{xtg}/E_{xtg}$. Let $C_t^{'}$ be the target total exposure for year t in the synthetic dataset and $(w_{tg_1}^{'}, \ldots, w_{tg_m}^{'})$ be a vector of weights adding to one which represents the desired splitting in year t of this exposure among the subgroups.

The synthetic central exposures $E_{xtg}^{'}$ in year t for people age x in subgroup g are obtained as

$$E_{xtg}^{'} = C_t^{'} \frac{\sum_g E_{xtg}}{\sum_x \sum_g E_{xtg}} w_{tg}^{'} = C_t^{'} \frac{E_{xt}}{E_t} w_{tg}^{'},$$

where $E_{xt} = \sum_g E_{xtg}$ are the total exposed to risk at age x in year t across all groups and $E_t = \sum_x \sum_g E_{xtg}$ are the total exposed to risk in year t across all groups and ages. Hence the exposure for the reference dataset is being used to obtain the split by age for a particular year and group. The corresponding synthetic number of deaths $D_{xtg}^{'}$ is generated by drawing a random sample from a Poisson distribution with mean $E_{xtg}^{'}\mu_{xtg}$.

Appendix 4.B Model fitting constraints

Some of the models require parameter constraints to ensure identifiability of the parameters. Table 4.8 presents the parameter constraints imposed to the reference population models and Table 4.9 shows the parameter constraints imposed to the book part of the two-population models.

Bibliography

Ahcan, A., Medved, D., Olivieri, A., Pitacco, E., 2014. Forecasting mortality for small populations by mixing mortality data. Insurance: Mathematics and Economics 54, 12–27.

TABLE 4.8: Parameter constraints for the reference population models.

Model	**Constraints**
LC	$\sum_x \beta_x^R = 1$, $\sum_t \kappa_t^R = 0$
LC2	$\sum_x \beta_x^{(1,R)} = 1$, $\sum_t \kappa_t^{(1,R)} = 0$, $\sum_x \beta_x^{(2,R)} = 1$, $\sum_t \kappa_t^{(2,R)} = 0$
LC2+Cohorts	$\sum_x \beta_x^{(1,R)} = 1$, $\sum_t \kappa_t^{(1,R)} = 0$, $\sum_x \beta_x^{(2,R)} = 1$, $\sum_t \kappa_t^{(2,R)} = 0$
LC+Cohorts	$\sum_x \beta_x^R = 1$, $\sum_t \kappa_t^R = 0$, $\sum_c \gamma_c^R = 0$, $\sum_c c\gamma_c^R = 0$
APC	$\sum_t \kappa_t^R = 0$, $\sum_c \gamma_c^R = 0$, $\sum_c c\gamma_c^R = 0$
M5	-
M6	$\sum_c \gamma_c^R = 0$, $\sum_c c\gamma_c^R = 0$
M7	$\sum_c \gamma_c^R = 0$, $\sum_c c\gamma_c^R = 0$, $\sum_c c^2\gamma_c^R = 0$

TABLE 4.9: Parameter constraints for the book part of the models.

Model	**Constraints**
CF+Cohorts	-
CAE+Cohorts	$\sum_t \kappa_t^B = 0$
RelLC+Cohorts	$\sum_x \beta_x^B = 1$, $\sum_t \kappa_t^B = 0$
Gravity (APC)	$\sum_t \kappa_t^B = 0$, $\sum_c \gamma_c^B = 0$, $\sum_c c\gamma_c^B = 0$
M7-M5	-
M7-M6	$\sum_c \gamma_c^B = 0$, $\sum_c c\gamma_c^B = 0$
M7-M7	$\sum_c \gamma_c^B = 0$, $\sum_c c\gamma_c^B = 0$, $\sum_c c^2\gamma_c^B = 0$
M7-Saint	-
M7-PLAT	-

Ahmadi, S. S., Li, J. S.-H., 2014. Coherent mortality forecasting with generalized linear models: A modified time-transformation approach. Insurance: Mathematics and Economics 59, 194–221.

Antonio, K., Bardoutsos, A., Ouburg, W., 2015. Bayesian Poisson log-bilinear models for mortality projections with multiple populations. Working paper.
URL https://lirias.kuleuven.be/bitstream/123456789/485564/1/AFI_1599.pdf

Biatat, V., Currie, I. D., 2010. Joint models for classification and comparison of mortality in different countries. In: Proceedings of 25rd International Workshop on Statistical Modelling, Glasgow. pp. 89–94.

Booth, H., Hyndman, R. J., Tickle, L., de Jong, P., 2006. Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. Demographic Research 15, 289–310.

Börger, M., Fleischer, D., Kuksin, N., 2013. Modeling the mortality trend under modern solvency regimes. ASTIN Bulletin 44 (1), 1–38.

Brouhns, N., Denuit, M., Van Keilegom, I., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. Scandinavian Actuarial Journal (3), 212–224.

Butt, Z., 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, Cass Business School.

Cairns, A. J., 2013. Robust hedging of longevity risk. Journal of Risk and Insurance 80, 621–648.

Cairns, A. J., Blake, D., Dowd, K., 2006. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. Journal of Risk and Insurance 73 (4), 687–718.

Cairns, A. J., Blake, D., Dowd, K., 2008. Modelling and management of mortality risk: a review. Scandinavian Actuarial Journal (2), 79–113.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., 2011a. Bayesian stochastic mortality modelling for two populations. ASTIN Bulletin 41, 29–59.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2011b. Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics 48 (3), 355–367.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1–35.

Cairns, A. J., Dowd, K., Blake, D., Coughlan, G. D., 2014. Longevity hedge effectiveness: a decomposition. Quantitative Finance 14 (2), 217–235.

Carter, L. R., Lee, R. D., 1992. Modeling and forecasting US sex differentials in mortality. International Journal of Forecasting 8 (3), 393–411.

Continuous Mortality Investigation, 2007. Working Paper 25 – Stochastic projection methodologies: Lee–Carter model features, example results and implications.

Coughlan, G. D., Khalaf-Allah, M., Ye, Y., Kumar, S., Cairns, A. J., Blake, D., Dowd, K., 2011. Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. North American Actuarial Journal 15 (2), 150–176.

Currie, I. D., Durban, M., Eilers, P. H., 2004. Smoothing and forecasting mortality rates. Statistical Modelling 4 (4), 279–298.

Debón, A., Martínez-Ruiz, F., Montes, F., 2010. A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. Insurance: Mathematics and Economics 47 (3), 327–336.

Debón, A., Montes, F., Martínez-Ruiz, F., 2011. Statistical methods to compare mortality for a group with non-divergent populations: an application to Spanish regions. European Actuarial Journal 1 (2), 291–308.

Delwarde, A., Denuit, M., Guillén, M., Vidiella-i Anguera, A., 2006. Application of the Poisson log-bilinear projection model to the G5 mortality experience. Belgian Actuarial Bulletin 6 (1), 54–68.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2010. Backtesting stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts. North American Actuarial Journal 14 (3), 281–298.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., Khalaf-allah, M., 2011. A gravity model of mortality rates for two related populations. North American Actuarial Journal 15 (2), 334–356.

Haberman, S., Kaishev, V. K., Millosaovich, P., Villegas, A. M., Baxter, S., Gaches, A., Gunnlaugsson, S., Sison, M., 2014a. A methodology for assessing longevity basis risk – User guide. Institute & Faculty of Actuaries, Life & Longevity Markets Association,

Hymans Robertson, Cass Business School.
URL http://www.hymans.co.uk/media/591834/141205-longevity-basis-risk-user-guide.pdf

Haberman, S., Kaishev, V. K., Millosaovich, P., Villegas, A. M., Baxter, S., Gaches, A., Gunnlaugsson, S., Sison, M., 2014b. Longevity Basis Risk: A methodology for assessing basis risk. Institute and Faculty of Actuaries Sessional Research Paper.
URL http://www.actuaries.org.uk/sites/all/files/documents/pdf/ifoa-llma-longevity-basis-risk-report.pdf

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

Haberman, S., Renshaw, A., 2011. A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics 48 (1), 35–55.

Hatzopoulos, P., Haberman, S., 2013. Common mortality modeling and coherent forecasts. An empirical analysis of worldwide mortality data. Insurance: Mathematics and Economics 52 (2), 320–337.

Hoffmann, R., 2005. Do socioeconomic mortality differences decrease with rising age? Demographic Research 13, 35–62.

Human Mortality Database, 2013. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL www.mortality.org

Hunt, A., Blake, D., 2015. Modelling longevity bonds: Analysing the Swiss Re Kortis Bond. Insurance: Mathematics and Economics.

Hymans Robertson LLP, 2015. Buy-outs, buy-ins and longevity hedging, Q4 2014.
URL http://www.hymans.co.uk/media/591924/150317-managing-pension-scheme-risk-q4-2014.pdf

Hyndman, R. J., Booth, H., Yasmeen, F., 2013. Coherent mortality forecasting: the product-ratio method with functional time series models. Demography 50 (1), 261–283.

Jarner, S. F., Kryger, E. M., 2011. Modelling adult mortality in small populations: The Saint Model. ASTIN Bulletin 41 (2), 377–418.

Kleinow, T., 2015. A Common Age Effect Model for the Mortality of Multiple Populations. Insurance: Mathematics and Economics.

Koissi, M., Shapiro, A., Hognas, G., 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. Insurance: Mathematics and Economics 38 (1), 1–20.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Li, J., 2012. A Poisson common factor model for projecting mortality and life expectancy jointly for females and males. Population Studies 67 (1), 111–126.

Li, J. S.-H., Hardy, M. R., 2011. Measuring basis risk in longevity hedges. North American Actuarial Journal 15 (2), 177–200.

Li, J. S.-H., Zhou, R., Hardy, M. R., 2015. A step-by-step guide to building two-population stochastic mortality models. Insurance: Mathematics and Economics.

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

LLMA, 2012. Basis risk in longevity hedging: parallels with the past. Institutional Investor Journals 2012 (1), 39–45.

Lu, J. L. C., Wong, W., Bajekal, M., 2014. Mortality improvement by socio-economic circumstances in England (1982 to 2006). British Actuarial Journal 19 (1), 1–35.

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Exley, S., Barnes, H., Dibben, C., 2007. The English indices of deprivation 2007. Department of Communities and Local Government, London.

Plat, R., 2009a. On stochastic mortality modeling. Insurance: Mathematics and Economics 45 (3), 393–404.

Plat, R., 2009b. Stochastic portfolio specific mortality and the quantification of mortality basis risk. Insurance: Mathematics and Economics 45 (1), 123–132.

Renshaw, A., Haberman, S., 2003. Lee-Carter mortality forecasting with age-specific enhancement. Insurance: Mathematics and Economics 33 (2), 255–272.

Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556–570.

Renshaw, A., 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.

Russolillo, M., Giordano, G., Haberman, S., 2011. Extending the Lee-Carter model: a three-way decomposition. Scandinavian Actuarial Journal (2), 96–117.

Salhi, Y., Loisel, S., 2012. Basis risk modelling: a co-integration based approach. Working paper.
URL http://halshs.archives-ouvertes.fr/hal-00746859/

Villegas, A. M., Haberman, S., 2014. On the modeling and forecasting of socioeconomic mortality differentials: an application to deprivation and mortality in England. North American Actuarial Journal 18 (1), 168–193.

Wan, C., Bertschi, L., 2015. Swiss coherent mortality model as a basis for developing longevity de-risking solutions for Swiss pension funds: A practical approach. Insurance: Mathematics and Economics.

Wilmoth, J., Valkonen, T., 2001. A parametric representation of mortality differentials over age and time. In: Fith seminar of EAPS Working Group on Differential in Health, Morbidity and Mortality in Europe.

Wood, S., 2015. Package mgcv.
URL http://cran.r-project.org/web/packages/mgcv/index.html

Yang, B., Li, J., Balasooriya, U., 2014. Cohort extensions of the Poisson common factor model for modelling both genders jointly. Scandinavian Actuarial Journal.

Yang, S. S., Wang, C.-W., 2013. Pricing and securitization of multi-country longevity risk with mortality dependence. Insurance: Mathematics and Economics 52 (2), 157–169.

Zhou, R., Wang, Y., Kaufhold, K., Li, J. S.-H., Tan, K. S., 2014. Modeling period effects in multi-population mortality models: applications to Solvency II. North American Actuarial Journal 18 (1), 150–167.

Part III

Tools for Modelling Mortality

5

StMoMo: An R Package for Stochastic Mortality Modelling

A stable version of the **R** package documented in this chapter is available on CRAN (the main repository for **R** packages) at:.

https://cran.r-project.org/package=StMoMo

In addition the associated source code is publicly available at:

https://github.com/amvillegas/StMoMo

StMoMo: An R Package for Stochastic Mortality Modelling

Andrés M. Villegas, Vladimir Kaishev, Pietro Millossovich

Cass Business School, City University London, United Kingdom

Abstract

In this chapter we mirror the framework of generalised (non-) linear models to define the family of generalised Age-Period-Cohort stochastic mortality models which encompasses the vast majority of stochastic mortality projection models proposed to date, including the well-known Lee-Carter and Cairns-Blake-Dowd models. We also introduce the **R** package **StMoMo** which exploits the unifying framework of the generalised Age-Period-Cohort family to provide tools for fitting stochastic mortality models, assessing their goodness of fit and performing mortality projections. We illustrate some of the capabilities of the package by performing a comparison of several stochastic mortality models applied to the England and Wales population.

Keywords: Mortality modelling; mortality forecasting; generalised linear models; generalised non-linear models

5.1 Introduction

During the last two centuries developed countries experienced a persistent increase in life expectancy. For instance, Oeppen and Vaupel (2002) estimate that during the last 160 years the world record in female life expectancy at birth has increased at an approximate steady pace of 3 months per year. This increase in life expectancy, though a sign of social progress, poses a challenge to governments, private pension plans and life insurers because of its impact on pension and health costs. Actuaries and demographers have recognised the problems caused by an ageing population and rising longevity and have thus devoted significant attention to the development of statistical techniques for the modelling and projection of mortality rates.

One of the most influential approaches to the stochastic modelling of mortality rates is the parsimonious mortality model proposed by Lee and Carter (1992). This model uses principal component analysis to decompose the age-time matrix of mortality rates into a bilinear combination of age and period parameters, with the latter being treated as time series to produce mortality projections. The Lee-Carter model has inspired numerous variants and extensions. For instance, Lee and Miller (2001), Booth et al. (2002), and Brouhns et al. (2002) have proposed alternative estimation approaches in order to improve the goodness-of-fit and the forecasting properties of the model. In particular, Brouhns et al. (2002) propose a more formal statistical approach to estimating the parameters by embedding the Lee-Carter model into a Poisson regression setting. Other authors have extended the Lee-Carter model by including additional terms, such as multiple bilinear age-period components (Renshaw and Haberman, 2003; Hyndman and Ullah, 2007), or a cohort effect term (Renshaw and Haberman, 2006; Currie, 2006).

The two factor Cairns-Blake-Dowd (CBD) model introduced by Cairns et al. (2006) is one of the most prominent variants of the Lee-Carter model. The CBD model relies on the linearity of the logit of one-year death probabilities at older ages. Specifically, it assumes that for a given year the logit of the one-year death probability is a linear function of age, and treats the intercept and slope parameters across years as stochastic processes. Cairns et al. (2009) consider three extensions to the original CBD model by incorporating combinations of a quadratic age term and a cohort effect term. Plat (2009) has combined features of the CBD and the Lee-Carter models to produce a model that is suitable for full age ranges and captures the cohort effect.

Given the abundance and rapid increase in the number of stochastic mortality models proposed in the literature, there have been some recent attempts to find the commonalities among these models. Hunt and Blake (2015) review the structure of mortality models and describe an Age-Period-Cohort model structure which encompasses the vast majority of stochastic mortality models. Currie (2014) shows that many mortality models can be expressed in terms of generalised linear models or generalised non-linear models.

In this chapter, we build upon the works of Hunt and Blake (2015) and Currie (2014) to define the family of Generalised Age-Period-Cohort stochastic mortality models by

mirroring the terminology of generalised linear models. We also introduce the **R** package **StMoMo**[1] which exploits the unifying framework of the Generalised Age-Period-Cohort family to provide computational tools for implementing many of the stochastic mortality models proposed to date. The **StMoMo** package is available at `https://cran.r-project.org/package=StMoMo`. Version 0.2 has been used for this chapter.

Several packages for mortality modelling are available in the **R** environment (R Core Team, 2014). The package **demography** (Hyndman, 2014), whose usage is explained in detail in Booth et al. (2014), implements the original Lee-Carter model along with the Lee and Miller (2001), Booth et al. (2002), and Hyndman and Ullah (2007) variants. The **ilc** package (Butt et al., 2014) implements the Renshaw and Haberman (2006) cohort extension of the Lee-Carter model together with the Lee-Carter model under a Poisson regression framework. The **LifeMetrics R** functions implement the original CBD model and the three extended CBD models considered in Cairns et al. (2009), along with the Lee-Carter model (using Poisson maximum likelihood), the Age-Period-Cohort model of Currie (2006) and the Renshaw and Haberman (2006) model. This package, which is not on CRAN, is available at `http://www.macs.hw.ac.uk/~andrewc/lifemetrics/`.

There are however several drawbacks of the existing packages which our package **StMoMo** seeks to overcome. First, the existing packages are based on model-specific fitting algorithms limiting the models available to those already predefined in the packages. By contrast, **StMoMo** allows users to easily expand the number of models available. In addition, whilst **StMoMo** provides forecasting and simulation functions for any model within the Generalised Age-Period-Cohort family, the existing packages only provide such functions for a limited number of models. For instance, the package **ilc** only includes forecasting functions for the Lee-Carter model. Similarly, simulation with the package **LifeMetrics** is limited to the Lee-Carter and the standard CBD models. Finally, **StMoMo** provides functions which are not available in existing packages, such as tools for analysing the goodness-of-fit[2] and evaluating the impact of parameter uncertainty using bootstrapping techniques.

StMoMo comes with a set of functions for defining an abstract model — specifying for instance the number of period terms, whether coefficients are parametric or not — and

[1] The acronym StMoMo, pronounced Saint Momo, stands for Stochastic Mortality Modelling. Momo is the king of Carnivals in numerous Latin American festivities (Wikipedia, 2014).

[2] We note that **ilc** also provides some graphical tools for assessing the goodness of fit of the models implemented in that package.

for fitting a given model. This is particularly useful when estimating several models on a given dataset or a given model to different datasets. The package also provides preset functions for defining the most common models available in the mortality forecasting literature. In addition, other models preferred by the user can be created in a very simple fashion, see Section 5.4 where several examples are given. Therefore, the flexibility of the package allows a user to quickly build up a battery of different models, and this is particularly useful when seeking the most appropriate mortality model, comparing different models or assessing model risk. **StMoMo** is particularly appealing for actuaries managing life and pensions portfolios exposed to longevity risk. The code backing most functions implemented in the package has been extensively used and tested in Chapter 4, where we have carried out a comparison of multipopulation mortality models in the context of assessing basis risk in longevity transactions.

In this chapter we describe the statistical framework underlying **StMoMo** and illustrate its usage. For this purpose, we use as a running example a comparison of several stochastic mortality models fitted to the England and Wales population. This example is in the spirit of the comparison exercises of Cairns et al. (2009, 2011), Haberman and Renshaw (2011) and Lovász (2011), allowing us to show how several of the analysis performed in these papers can easily be replicated using **StMoMo**. The structure of the chapter is as follows. In Section 5.2 we introduce our notation. In Section 5.3, we mirror the framework of generalised linear models to define the family of Generalised Age-Period-Cohort (GAPC) stochastic mortality models and demonstrate that many of the mortality models discussed in the literature can be framed within this family. In Section 5.4 we explain how the GAPC family of models is implemented in **StMoMo**. In Section 5.5, we describe the fitting of GAPC mortality models and illustrate how this can be accomplished using **StMoMo**. In Section 5.6 we consider the evaluation of the goodness-of-fit of GAPC models. In Section 5.7 we discuss the forecasting and simulation of GAPC models using time series techniques. Section 5.8 describes the use of bootstrapping techniques to incorporate parameter uncertainty in the estimation and forecasting of GAPC mortality models. Finally, in Section 5.9 we provide some conclusions and discuss possible extensions of the **StMoMo** package.

5.2 Notation and Data

Let the random variable D_{xt} denote the number of deaths in a population at age x last birthday during calendar year t. Also let d_{xt} denote the observed number of deaths, E^c_{xt} the central exposed to risk at age x in year t, and E^0_{xt} the corresponding initial exposed to risk. The one-year mortality rate for an individual aged x last birthday and in calendar year t, denoted q_{xt}, can be estimated as $\hat{q}_{xt} = d_{xt}/E^0_{xt}$. The force of mortality and central death rates are denoted by μ_{xt} and m_{xt}, respectively, with the empirical estimate of the latter being $\hat{m}_{xt} = d_{xt}/E^c_{xt}$. Under the assumption that the force of mortality is constant over each year of age and calendar year, i.e. from age x to age $x+1$ and t to $t+1$, then the force of mortality μ_{xt} and the death rate m_{xt} coincide. We assume that this is the case throughout.

In **StMoMo** and throughout this chapter we assume that deaths, d_{xt}, and either central exposures, E^c_{xt}, or initial exposures, E^0_{xt}, are available in a rectangular array format comprising ages (on the rows) $x = x_1, x_2, \ldots, x_k$, and calendar years (on the columns) $t = t_1, t_2, \ldots, t_n$, When only central exposures are available and initial exposures are required (or vice-versa), one can approximate the initial exposures by adding half the matching reported numbers of deaths to the central exposures, i.e, $E^0_{xt} \approx E^c_{xt} + \frac{1}{2}d_{xt}$. When the context is clear, we may write E_{xt} to refer to E^0_{xt} or E^c_{xt}.

5.3 Generalised Age-Period-Cohort stochastic mortality models

Some authors have recently sought to identify the similarities amongst stochastic mortality models. For instance, Hunt and Blake (2015) describe an Age-Period-Cohort model structure which encompasses the vast majority of stochastic mortality models. In another interesting contribution, Currie (2014) shows that many common mortality models can be expressed in the standard terminology of generalised linear or non-linear models. In this section, we build upon the aforementioned papers to define the family of Generalised Age-Period-Cohort (GAPC) stochastic mortality models.

Akin to generalised linear models (see e.g. McCullagh and Nelder (1989)), a GAPC stochastic mortality model is comprised of four components:

i. The *random component*: the numbers of deaths D_{xt} follow a Poisson distribution or a Binomial distribution[3], so that

$$D_{xt} \sim \text{Poisson}(E^c_{xt}\mu_{xt})$$

or

$$D_{xt} \sim \text{Binomial}(E^0_{xt}, q_{xt}),$$

with $\mathbb{E}\left(D_{xt}/E^c_{xt}\right) = \mu_{xt}$ and $\mathbb{E}\left(D_{xt}/E^0_{xt}\right) = q_{xt}$, respectively.

ii. The *systematic component*: following Hunt and Blake (2015) the effects of age x, calendar year t and year-of-birth (cohort) $c = t - x$ are captured through a *predictor* η_{xt} given by:

$$\eta_{xt} = \alpha_x + \sum_{i=1}^{N} \beta^{(i)}_x \kappa^{(i)}_t + \beta^{(0)}_x \gamma_{t-x}.$$

Here:

- The term α_x is a static age function capturing the general shape of mortality by age.
- $N \geq 0$ is an integer indicating the number of age-period terms describing the mortality trends, with each time index $\kappa^{(i)}_t$, $i = 1, \ldots, N$, contributing in specifying the mortality trend and $\beta^{(i)}_x$ modulating its effect across ages.
- The term γ_{t-x} accounts for the cohort effect with $\beta^{(0)}_x$ modulating its effect across ages.

The age modulating terms $\beta^{(i)}_x$, $i = 0, 1, \ldots, N$, can be either pre-specified functions of age, i.e. $\beta^{(i)}_x \equiv f^i(x)$, as in CBD type models, or non-parametric terms without any prior structure and which need to be estimated as in the Lee-Carter model. In the GAPC family we assume that the period indexes $\kappa^{(i)}_t$, $i = 1, \ldots, N$, and the cohort index γ_{t-x} are stochastic processes rather than smooth functions of time or cohort. This is the key feature that allows the stochastic projection of GAPC models and thus the generation of probabilistic forecasts of future mortality rates.

[3]More precisely, conditionally on (μ_{xt}), the numbers of deaths are independent and follow a Poisson distribution. Similarly, in the Binomial case, the numbers of deaths are independent conditionally on (q_{xt}).

iii. The *link function* g associating the random component and the systematic component so that

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}}\right)\right) = \eta_{xt}.$$

Although a number of link functions would be possible it is convenient to use the so-called canonical link and pair the Poisson distribution with the log link function and the Binomial distribution with the logit link function (see e.g. Currie (2014) for a discussion of this in the context of mortality models and McCullagh and Nelder (1989) in the wider context of GLMs).

iv. The *set of parameter constraints*: most stochastic mortality models are only identifiable up to a transformation and thus require parameter constraints to ensure unique parameter estimates. These parameter constraints are applied through a *constraint function* v which maps an arbitrary vector of parameters

$$\theta := \left(\alpha_x, \beta_x^{(1)}, ..., \beta_x^{(N)}, \kappa_t^{(1)}, ..., \kappa_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x}\right)$$

into a vector of transformed parameters

$$v(\theta) = \tilde{\theta} = \left(\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, ..., \tilde{\beta}_x^{(N)}, \tilde{\kappa}_t^{(1)}, ..., \tilde{\kappa}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x}\right)$$

satisfying the model constraints with no effect on the predictor η_{xt} (i.e. θ and $\tilde{\theta}$ result in the same η_{xt}).

Most stochastic mortality models proposed in the literature belong to the GAPC family. This includes the original Lee-Carter model, the extensions of the Lee-Carter proposed in Renshaw and Haberman (2003, 2006), the original CBD model, and the extended CBD models of Cairns et al. (2009). In addition, all the model structures considered in Haberman and Renshaw (2011), Lovász (2011) and van Berkum et al. (2014), as well as the models of Plat (2009), Aro and Pennanen (2011), O'Hare and Li (2012), Börger et al. (2013) and Alai and Sherris (2014), are part of the GAPC class of models.[4]

Next, we describe in detail some of these models highlighting how they can be framed within the GAPC family.

[4]We note however that models which rely on the smoothness of mortality over both age and time, such as the graduation approach of Renshaw et al. (1996) and the P-Spline model of Currie (2006), do not belong to the GAPC family. The P-Spline approach is implemented in the **MortalitySmooth** (Camarda, 2012) **R** package.

5.3.1 Lee-Carter model under a Poisson setting

The Lee Carter model as implemented by Brouhns et al. (2002) assumes a Poisson distribution of the deaths using a log *link function* to target the force of mortality μ_{xt}. The predictor structure proposed by Lee and Carter (1992) assumes that there is a static age function, α_x, a unique non-parametric age-period term ($N = 1$), and no cohort effect. Thus, the predictor is given by:

$$\eta_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)} \tag{5.1}$$

In order to project mortality, the time index $\kappa_t^{(1)}$ is modelled and forecasted using ARIMA processes. Typically, a random walk with drift has been shown to provide a reasonable fit, that is,

$$\kappa_t^{(1)} = \delta + \kappa_{t-1}^{(1)} + \xi_t, \qquad \xi_t \sim N(0, \sigma_\kappa),$$

where δ is the drift parameter and ξ_t is Gaussian white noise process with variance σ_κ.

The Lee-Carter model is only identifiable up to a transformation, as for arbitrary real constants c_1 and $c_2 \neq 0$ the parameters in Equation (5.1) can be transformed in the following way

$$\left(\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}\right) \rightarrow \left(\alpha_x + c_1\beta_x^{(1)}, \frac{1}{c_2}\beta_x^{(1)}, c_2(\kappa_t^{(1)} - c_1)\right), \tag{5.2}$$

leaving η_{xt} unchanged. To ensure identifiability of the model Lee and Carter (1992) suggest the following *set of parameter constraints*

$$\sum_x \beta_x^{(1)} = 1, \qquad \sum_t \kappa_t^{(1)} = 0, \tag{5.3}$$

which can be imposed by choosing

$$c_1 = \frac{1}{n}\sum_t \kappa_t^{(1)}, \qquad c_2 = \sum_x \beta_x^{(1)} \tag{5.4}$$

in transformation (5.2).

5.3.2 Renshaw and Haberman model: Lee-Carter with cohort effects

Renshaw and Haberman (2006) generalise the Lee-Carter model by incorporating a cohort effect to obtain the predictor:

$$\eta_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(0)}\gamma_{t-x} \tag{5.5}$$

Mortality projections for this model are derived using time series forecast of the estimated $\kappa_t^{(1)}$ and γ_{t-x}, generated using univariate ARIMA processes under the assumption of independence between the period and the cohort effects.

To estimate the model Renshaw and Haberman (2006) assume a Poisson distribution of deaths (*random component*) and use a log *link* function targeting the force of mortality μ_{xt}. As with the Lee-Carter model, the predictor η_{xt} is invariant to the transformation:

$$\begin{aligned}\left(\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}, \beta_x^{(0)}, \gamma_{t-x}\right) \rightarrow \Big(&\alpha_x + c_1\beta_x^{(1)} + c_2\beta_x^{(1)}, \frac{1}{c_3}\beta_x^{(1)}, \\ &c_3(\kappa_t^{(1)} - c_1), \frac{1}{c_4}\beta_x^{(0)}, c_4(\gamma_{t-x} - c_2)\Big),\end{aligned} \tag{5.6}$$

where c_1, c_2, $c_3 \neq 0$ and $c_4 \neq 0$ are real constants. Identifiability of the model can be ensured using the following *set of parameter constraints*:

$$\sum_x \beta_x^{(1)} = 1, \qquad \sum_t \kappa_t^{(1)} = 0, \qquad \sum_x \beta_x^{(0)} = 1, \qquad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0,$$

which can be imposed by setting

$$c_1 = \frac{1}{n}\sum_t \kappa_t^{(1)}, \quad c_2 = \frac{1}{n+k-1}\sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c, \quad c_i = \sum_x \beta_x^{(i)}, \quad i = 3, 4,$$

in transformation (5.6).

Renshaw and Haberman (2006) also consider several substructures of the predictor (5.5) obtained by setting to a constant one or both of the age modulating terms. Of particular interest is the substructure obtained by setting $\beta_x^{(0)} = 1$,

$$\eta_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \gamma_{t-x}, \tag{5.7}$$

which has been suggested by Haberman and Renshaw (2011) as a simpler structure that resolves some stability issues of the original model.

5.3.3 Age-Period-Cohort model

Another commonly used substructure of the Renshaw and Haberman model is the so-called Age-Period-Cohort (APC) model, corresponding to $\beta_x^{(1)} = 1$, $\beta_x^{(0)} = 1$,

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x},$$

which has a long-standing tradition in the fields of medicine and demography (see, e.g., Clayton and Schifflers (1987) and Hobcraft et al. (1982)), but has not been widely used in the actuarial literature until it was considered by Currie (2006). The APC model is known to be invariant with respect to the following two transformations:

$$\left(\alpha_x, \kappa_t^{(1)}, \gamma_{t-x}\right) \to \left(\alpha_x + \phi_1 - \phi_2 x, \kappa_t^{(1)} + \phi_2 t, \gamma_{t-x} - \phi_1 - \phi_2 (t-x)\right) \tag{5.8}$$

$$\left(\alpha_x, \kappa_t^{(1)}, \gamma_{t-x}\right) \to \left(\alpha_x + c_1, \kappa_t^{(1)} - c_1, \gamma_{t-x}\right), \tag{5.9}$$

where c_1, ϕ_1, and ϕ_2 are real constants. However, we can ensure identifiability of the model by imposing the *set of constraints*:

$$\sum_t \kappa_t^{(1)} = 0, \qquad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \qquad \sum_{c=t_1-x_k}^{t_n-x_1} c\gamma_c = 0,$$

where the last two constraints imply that the cohort effect fluctuates around zero with no discernible linear trend. Following Haberman and Renshaw (2011, Appendix A), the constraints on the cohort effect can be imposed by applying transformation (5.8) with constants ϕ_1 and ϕ_2 obtained by regressing γ_{t-x} on $t - x$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2 (t-x) + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d.}$$

The constraint on the period index can then be imposed by applying transformation (5.9) with

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}.$$

5.3.4 CBD model

Cairns et al. (2006) propose a predictor structure with two age-period terms ($N = 2$) with pre-specified age-modulating parameters $\beta_x^{(1)} = 1$ and $\beta_x^{(2)} = x - \bar{x}$, no static age function and no cohort effect. Thus, the predictor of the CBD model is given by:

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)},$$

where $\bar{x}$ is the average age in the data. Cairns et al. (2006) obtain mortality forecasts by projecting the period effects $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ using a bivariate random walk with drift.

The CBD model does not have identifiability issues and hence the *set of parameter constraints* is empty. In order to estimate the parameter of the CBD model we can follow Haberman and Renshaw (2011) and assume a Binomial distribution of deaths using a logit *link function* targeting the one-year death probabilities q_{xt}.

5.3.5 M7: Quadratic CBD model with cohort effects

Cairns et al. (2009) extend the original CBD model by adding a cohort effect and a quadratic age effect to obtain the predictor:

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \kappa_t^{(3)}\left((x - \bar{x})^2 - \hat{\sigma}_x^2\right) + \gamma_{t-x}, \tag{5.10}$$

where $\hat{\sigma}_x^2$ is the average value of $(x - \bar{x})^2$. This model, usually referred to as model M7, is not identifiable as the transformation

$$\begin{aligned}\left(\kappa_t^{(1)}, \kappa_t^{(3)}, \kappa_t^{(3)}, \gamma_{t-x}\right) \rightarrow \Big(&\kappa_t^{(1)} + \phi_1 + \phi_2(t - \bar{x}) + \phi_3\left((t - \bar{x})^2 + \hat{\sigma}^2\right), \\ &\kappa_t^{(2)} - \phi_2 - 2\phi_3(t - \bar{x}), \kappa_t^{(3)} + \phi_3, \\ &\gamma_{t-x} - \phi_1 - \phi_2(t - x) - \phi_3(t - x)^2\Big),\end{aligned} \tag{5.11}$$

for real constants ϕ_1, ϕ_2 and ϕ_3, leave the predictor unchanged. To identify the model Cairns et al. (2009) impose the *set of constraints*:

$$\sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \qquad \sum_{c=t_1-x_k}^{t_n-x_1} c\gamma_c = 0, \qquad \sum_{c=t_1-x_k}^{t_n-x_1} c^2\gamma_c = 0,$$

which ensure that the cohort effect fluctuates around zero and has no discernible linear or quadratic trend. Following Haberman and Renshaw (2011, Appendix A), these constraints can be imposed by applying transformation (5.11) with constants ϕ_1, ϕ_2 and ϕ_3 obtained by regressing γ_{t-x} on $t-x$ and $(t-x)^2$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2(t-x) + \phi_3(t-x)^2 + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d.}$$

Cairns et al. (2009) also consider the simpler predictor structures

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \gamma_{t-x},$$
$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + (x_c - x)\gamma_{t-x},$$

where x_c is a constant parameter to be estimated. These structures are typically referred to as models M6 and M8, respectively.

5.3.6 Plat model

Plat (2009) combines the CBD model with some features of the Lee-Carter model to produce a model that is suitable for full age ranges and captures the cohort effect. The proposed predictor structure assumes that there is a static age function, α_x, three age-period terms ($N = 3$) with pre-specified age-modulating parameters $\beta_x^{(1)} = 1$, $\beta_x^{(2)} = \bar{x} - x$, $\beta_x^{(3)} = (\bar{x} - x)^+ = \max(0, \bar{x} - x)$, and a cohort effect with pre-specified age-modulating parameters $\beta_x^{(0)} = 1$. Thus, the predictor is given by:

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_{t-x}. \tag{5.12}$$

Plat (2009) targets the force of mortality μ_{xt} with the log *link* and estimates the parameters of the model by assuming a Poisson distribution of the deaths. The following

parameter transformations leave the predictor in (5.12) unchanged:

$$\begin{aligned}\left(\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \tilde{\gamma}_{t-x}\right) \to \Big(&\alpha_x + \phi_1 - \phi_2 x + \phi_3 x^2, \\ &\kappa_t^{(1)} + \phi_2 t + \phi_3(t^2 - 2\bar{x}t), \kappa_t^{(2)} + 2\phi_3 t, \kappa_t^{(3)}, \\ &\gamma_{t-x} - \phi_1 - \phi_2(t-x) - \phi_3(t-x)^2\Big) \qquad (5.13)\\ \left(\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}\right) \to \Big(&\alpha_x + c_1 + c_2(\bar{x} - x) + c_3(\bar{x} - x)^+, \\ &\kappa_t^{(1)} - c_1, \kappa_t^{(2)} - c_2, \kappa_t^{(3)} - c_3, \gamma_{t-x}\Big), \qquad (5.14)\end{aligned}$$

where c_1, c_2, c_3, ϕ_1, ϕ_2, and ϕ_3 are any real constants. The following *set of parameter constraints* can be imposed to ensure identifiability:

$$\begin{aligned}&\sum_t \kappa_t^{(1)} = 0, \quad \sum_t \kappa_t^{(2)} = 0, \quad \sum_t \kappa_t^{(3)} = 0, \\ &\sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} c\gamma_c = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} c^2\gamma_c = 0 \qquad (5.15)\end{aligned}$$

The first three constraints ensure that the period indexes are centred around zero, while the last three constraints ensure that the cohort effect fluctuates around zero and has no linear or quadratic trend. Following Haberman and Renshaw (2011, Appendix A), the constraints on the cohort effect can be imposed by applying transformation (5.13) with constants ϕ_1, ϕ_2, and ϕ_3 obtained by regressing γ_{t-x} on $t-x$ and $(t-x)^2$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2(t-x) + \phi_3(t-x)^2 + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d. .} \qquad (5.16)$$

The constraints on the period indexes can then be imposed by applying transformation (5.14) with

$$c_i = \frac{1}{n}\sum_t \kappa_t^{(i)}, \qquad i = 1, 2, 3. \qquad (5.17)$$

In the cases where only older ages are of interest, Plat (2009) suggests dropping the third period term from predictor (5.12)[5]:

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_{t-x}. \qquad (5.18)$$

[5]Note that the reduced Plat model is essentially the M6 model of Cairns et al. (2009) with an added static age term α_x.

We note that this reduced Plat model has the same identifiability issues as the complete model with the omission of the transformations and constraints involving $\kappa_t^{(3)}$ and c_3.

5.4 GAPC stochastic mortality models with StMoMo

The **StMoMo** package provides an **R** implementation of the GAPC family of stochastic mortality models using the standard `S3` object-oriented system. **StMoMo** can be installed with the code:

```
install.packages("StMoMo")
```

The package is loaded within **R** as follows:

```
library(StMoMo)
```

In package **StMoMo**, GAPC stochastic mortality models are constructed using the `StMoMo` function. This function takes as input information on the *link function* (and the associated distributional assumption), the *predictor structure*, and the *set of parameter constraints* to create an object of the type `"StMoMo"` representing a GAPC mortality model:

- The argument `link` defines the *link function* and *the random component* associated with the mortality model. Setting `link = "log"` assumes that deaths follow a Poisson distribution and uses a log link targeting the force of mortality μ_{xt}, while setting `link = "logit"` assumes that deaths follow a Binomial distribution and uses a logit link targeting one-year mortality rates q_{xt}.

- The *predictor* of the model is defined via arguments `staticAgeFun`, `periodAgeFun` and `cohortAgeFun`. Argument `staticAgeFun` is a logical variable indicating whether the model has a static age function α_x or not. Argument `periodAgeFun` is a list of length N containing the definitions of the period age-modulating parameters $\beta_x^{(i)}$, $i = 1, \ldots, N$, with each entry being either `"NP"` for non-parametric age terms, `"1"` for $\beta_x^{(i)} = 1$, or a predefined parametric function of age[6]. Argument `cohortAgeFun`

[6] Note that we can define a model with no age-period terms ($N = 0$) by making `periodAgeFun = NULL`.

TABLE 5.1: Model structures considered in this chapter.

Model	**Predictor**
LC	$\eta_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)}$
CBD	$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$
APC	$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$
RH	$\eta_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \gamma_{t-x}$
M7	$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \left((x - \bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3)} + \gamma_{t-x}$
PLAT	$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_{t-x}$

defines the cohort age modulating parameter $\beta_x^{(0)}$ and can take values `"NP"` for non-parametric age terms, `"1"` for $\beta_x^{(0)} = 1$, a predefined parametric function of age, or `NULL` if the model does not have a cohort effect.

- The *set of parameter constraints* are defined via the argument `constFun` which is a user-defined implementation of the *constraint function* v mapping an arbitrary vector of parameters to a vector of transformed parameters satisfying the model constraints.

We note that due to limitations of the **R** functions used for fitting `"StMoMo"` objects to data (see Section 5.5), the current version **StMoMo** does not support models combining parametric and non-parametric age-modulating functions, $\beta_x^{(i)}$, $0 = 1, \ldots, N$. However, such model are not typically considered and the majority of models proposed in the literature are either extension of the Lee-Carter model with all age-modulating terms being non-parametric or extension of the CBD model with all age-modulating terms being parametric.[7]

In order to illustrate the creation of particular GAPC mortality models and other capabilities of **StMoMo**, in the rest of this chapter we will focus on the models summarised in Table 5.1. From now onwards, LC stands for the Lee-Carter model; CBD for the original Cairns-Blake-Dowd model; APC for the Age-Period-Cohort model of Currie (2006); RH for the cohort extension of the Lee-Carter model defined in equation (5.7) and proposed by Renshaw and Haberman (2006); M7 for the quadratic CBD model

[7]For instance, a model with predictor structure $\eta_{xt} = \alpha_x + (x - \bar{x})\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(1)}$, corresponding to `StMoMo(staticAgeFun = TRUE, periodAgeFun = c(f1, "NP"))`, with `f1 <- function(x, ages) x - mean(ages)`, is not supported.

defined in equation (5.10); and PLAT for the reduced Plat model defined previously in equation (5.18). For the sake of comparability, in all cases we will assume a Binomial distribution of deaths and use the logit function to link q_{xt} to the predictor structure η_{xt}.

Below, we show how to define each of the models in Table 5.1 using the package **StMoMo**.

Lee-Carter model. The LC model under a Binomial setting can be defined using the following code:

```
constLC <- function(ax, bx, kt, b0x, gc, wxt, ages){
  c1 <- mean(kt[1, ], na.rm = TRUE)
  c2 <- sum(bx[, 1], na.rm = TRUE)
  list(ax = ax + c1 * bx, bx = bx / c2, kt =  c2 * (kt - c1))
}
LC <- StMoMo(link = "logit", staticAgeFun = TRUE, periodAgeFun = "NP",
             constFun = constLC)
```

Recalling Section 5.3.1, we note that the constraint function `constLC` is the `R` implementation of transformation (5.2) with constants c_1 and c_2 calculated using Equation (5.4) to impose the constraints defined in equation (5.3). The **StMoMo** package also contains the function `lc` to facilitate the definition of Lee-Carter models. Hence, we could define the LC model using the much simpler predefined command:

```
LC <- lc(link = "logit")
```

CBD model. To define the CBD model we use the following commands:

```
f2 <- function(x, ages) x - mean(ages)
CBD <- StMoMo(link = "logit", staticAgeFun = FALSE,
              periodAgeFun = c("1", f2))
```

Here, we note that function `f2` defines the second age-modulating parameter $\beta_x^{(2)} = x - \bar{x}$ and that a `constFUN` argument needs not be provided since the CBD model does not have

identifiability issues. Alternatively, we can define the CBD model using the predefined function `cbd`:

```
CBD <- cbd()
```

APC model, RH model and model M7. These three models could be defined by implementing explicitly the discussions in Sections 5.3.2, 5.3.3 and 5.3.5. However, **StMoMo** includes predefined functions `apc`, `rh`, `m7` that facilitate the definition of the APC model, the RH model and model M7, respectively[8]. Thus, these models are defined with the code:

```
RH <- rh(link = "logit",  cohortAgeFun = "1")
APC <- apc(link = "logit")
M7 <- m7()
```

PLAT model. Package **StMoMo** does not include a predefined function for the Plat model. Nevertheless, recalling Section 5.3.6, we can define the reduced Plat model using the code:

```
f2 <- function(x, ages) mean(ages) - x
constPlat <- function(ax, bx, kt, b0x, gc, wxt, ages){
  nYears <- dim(wxt)[2]
  x <- ages
  t <- 1:nYears
  c <- (1 - tail(ages, 1)):(nYears - ages[1])
  xbar <- mean(x)
  #\sum g(c)=0, \sum cg(c)=0, \sum c^2g(c)=0
  phiReg <- lm(gc ~ 1 + c + I(c^2), na.action = na.omit)
  phi <- coef(phiReg)
  gc <- gc - phi[1] - phi[2] * c - phi[3] * c^2
  kt[2, ] <- kt[2, ] + 2 * phi[3] * t
  kt[1, ] <- kt[1, ] + phi[2] * t + phi[3] * (t^2 - 2 * xbar * t)
  ax <- ax + phi[1] - phi[2] * x + phi[3] * x^2
```

[8]The **StMoMo** package also includes functions `m6` and `m8` for defining models M6 and M8. We also note that the Renshaw and Haberman (2006) cohort extension of the Lee-Carter in Equation (5.5) can be defined using function `rh` with argument `cohortAgeFun = "NP"`.

```
  #\sum kt[i, ] = 0
  ci <- rowMeans(kt, na.rm = TRUE)
  ax <- ax + ci[1] + ci[2] * (xbar - x)
  kt[1, ] <- kt[1, ] - ci[1]
  kt[2, ] <- kt[2, ] - ci[2]
  list(ax = ax, bx = bx, kt = kt, b0x = b0x, gc = gc)
}
PLAT <- StMoMo(link = "logit", staticAgeFun = TRUE,
                  periodAgeFun = c("1", f2), cohortAgeFun = "1",
                  constFun = constPlat)
```

We note that the constraint function `constPlat` is the `R` implementation of transformations (5.13) and (5.14) omitting the terms involving $\kappa^{(3)}$ and c_3, and with constants ϕ_1, ϕ_2, ϕ_3 obtained via the linear regression defined in (5.16) and constant c_1 and c_2 as in Equation (5.17). Function `constPlat` imposes the constraints in Equation (5.15).

5.5 Model fitting

Parameter estimates of GAPC stochastic mortality models can be obtained by maximising the model log-likelihood, which is given by

$$\mathcal{L}(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \left\{ d_{xt} \log \hat{d}_{xt} - \hat{d}_{xt} - \log d_{xt}! \right\}$$

in the case of a Poisson distribution of deaths, and by

$$\mathcal{L}(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \left\{ d_{xt} \log \left(\frac{\hat{d}_{xt}}{E^0_{xt}} \right) + (E^0_{xt} - d_{xt}) \log \left(\frac{E^0_{xt} - \hat{d}_{xt}}{E^0_{xt}} \right) + \binom{E^0_{xt}}{d_{xt}} \right\}$$

in the case of a Binomial distribution of deaths. In both cases, ω_{xt} are weights taking the value 0 if a particular (x, t) data cell is omitted or 1 if the cell is included, and

$$\hat{d}_{xt} = E_{xt}\, g^{-1} \left(\alpha_x + \sum_{i=1}^{N} \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \right)$$

is the expected number of deaths predicted by the model, with g^{-1} denoting the inverse of the *link function* g.

In the mortality literature, maximisation of the log-likelihood is typically performed using the Newton-Raphson iterative procedure tailored for each model (see e.g. Brouhns et al. (2002), Renshaw and Haberman (2006) and Cairns et al. (2009)). This is in fact the approach implemented in the packages **ilc** and **LifeMetrics**. Nonetheless, as discussed extensively by Currie (2014), many stochastic mortality models are examples of generalised linear models or generalised non-linear model, which facilitates their fitting using standard statistical software[9]. Currie (2014) exemplifies this fact by fitting several stochastic mortality models in **R** using the standard function `glm` or the function `gnm` of the package **gnm** (Turner and Firth, 2012)[10].

StMoMo provides the generic function `fit` for estimating the parameters of GAPC mortality models. In line with the remarks of Currie (2014), the corresponding `S3` method for objects of the class `"StMoMo"` relies on function `gnm` of the package **gnm** to estimate the parameters of a GAPC model[11]. Internally, this is accomplished by constructing the equivalent **gnm** formulation of the GAPC mortality model[12]. For instance, the **gnm** formula of the Binomial `LC` model created before is

```
LC$gnmFormula


## [1] "D/E ~ -1 + offset(o) + factor(x) + Mult(factor(x), factor(t),
inst = 1)"
```

while the **gnm** formula of the Binomial `CBD` model defined before is

```
CBD$gnmFormula


## [1] "D/E ~ -1 + offset(o) + factor(t) + B2:factor(t)"
```

We now illustrate the usage of the function `fit` of the package **StMoMo** by fitting the six models defined before to England and Wales mortality data. The object `EWMaleData`

[9] Haberman and Renshaw (2011) have also noticed this fact and profit from GLM facilities in standard statistical packages when fitting CBD type models.

[10] Debón et al. (2010) also discuss the use of the package **gnm** for fitting Lee-Carter type models.

[11] The current version of **StMoMo** only includes the `S3` method `fit.StMoMo` for objects of class `"StMoMo"`, but future versions may provide tailored `fit` methods for particular stochastic mortality models.

[12] We note that when all the $\beta_x^{(i)}$ are parametric functions of age, the model is a GLM and therefore `gnm` by default resorts to the `glm` function of `R` when fitting the parameter of the model.

included in the package **StMoMo** contains data on deaths and central exposures for England and Wales males for the period 1961-2011 and for ages 0-100 obtained from the Human Mortality Database (2014). However, in our examples we concentrate on ages 55 to 89 as the CBD model and the M7 model have been particularly designed to fit higher ages. Additionally, since some models include cohort effects and in agreement with the usual practice (see e.g Cairns et al. (2009) and Haberman and Renshaw (2011)), we exclude (by setting $\omega_{xt} = 0$) all cohorts that have fewer than three observations.

Models LC, APC, CBD, M7 and PLAT can be fitted to England and Wales male mortality data for ages 55 to 89 using the code:

```
Dxt <- EWMaleData$Dxt
Ext <- EWMaleData$Ext + 0.5 * EWMaleData$Dxt
ages <- EWMaleData$ages
years <- EWMaleData$years
ages.fit <- 55:89
wxt <- genWeightMat(ages = ages.fit, years = years, clip = 3)
LCfit <- fit(LC, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
             ages.fit = ages.fit, wxt = wxt)
APCfit <- fit(APC, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
              ages.fit = ages.fit, wxt = wxt)
CBDfit <- fit(CBD, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
              ages.fit = ages.fit, wxt = wxt)
M7fit <- fit(M7, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
             ages.fit = ages.fit, wxt = wxt)
PLATfit <- fit(PLAT, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
               ages.fit = ages.fit, wxt = wxt)
```

From this code we note the following:

- In order to match the logit-Binomial setting used before in the definition of the mortality models, initial exposures are approximated by transforming the available central exposures.

- The first and last three cohorts years are excluded from the fitting via argument `wxt`. The appropriate 0-1 weighting matrix, `wxt`, is constructed using the utility function `genWeightMat` of package **StMoMo**.

As discussed in Chapter 6, the fitting of the RH model requires some care as it is well known that fitting cohort extensions of the Lee-Carter models is problematic. In particular, Currie (2014) has encountered convergence issues when using package **gnm** to fit the RH model. As a possible way to circumvent these issues, Currie (2014) suggests the use of appropriate starting values when fitting model RH. This can be achieved in function `fit` via input arguments `start.ax`, `start.bx`, `start.kt`, `start.b0x`, and `start.gc`. Using the parameters of the Lee-Carter model as starting values, model RH can be fitted with the code:

```
RHfit <- fit(RH, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
             ages.fit = ages.fit, wxt = wxt, start.ax = LCfit$ax,
             start.bx = LCfit$bx, start.kt = LCfit$kt)
```

The output from the function `fit` is an object of the class `"fitStMoMo"` including, among other things, the following information:

- `model`: the `"StMoMo"` object defining the underlying GAPC stochastic mortality model;
- `ax`, `bx`, `kt`, `b0x`, `gc`: the estimated parameters;
- `loglik`: the log-likelihood of the model;
- `deviance`: the model deviance;
- `nobs`: the number of observations in the data;
- `npar`: the effective number of parameters of the model;
- `fittingModel`: the output of the `gnm` call used to fit the model.

There are `print`, `plot`, `fitted`, and `residuals` methods for the `"fitStMoMo"` class. For instance, Figures 5.1, 5.2 and 5.3 depicting the fitted parameters of the LC model, the CBD model and the APC model, respectively, were produced with the code:

FIGURE 5.1: Parameters for the Lee-Carter (LC) model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

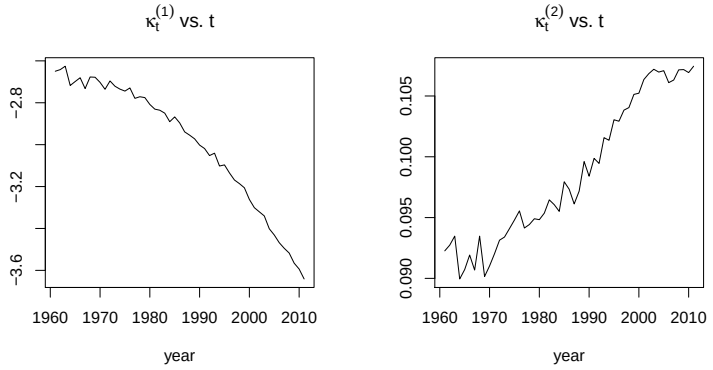


FIGURE 5.2: Parameters for the CBD model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

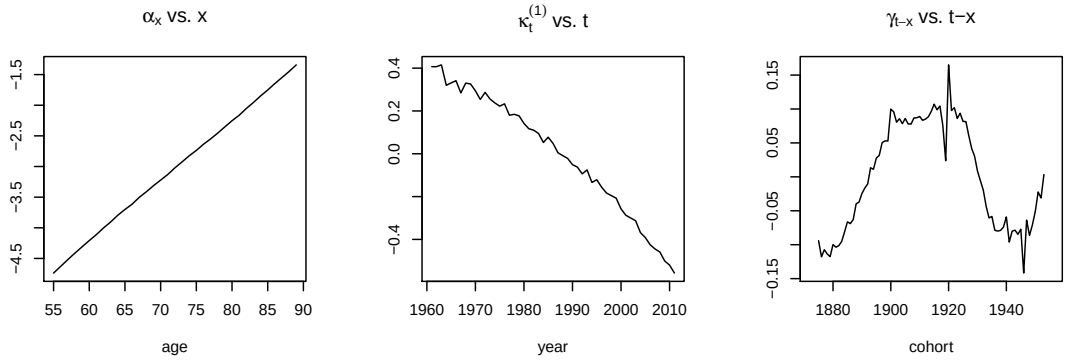


FIGURE 5.3: Parameters for the APC model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

```
plot(LCfit, nCol = 3)

plot(CBDfit, parametricbx = FALSE)

plot(APCfit, parametricbx = FALSE, nCol = 3)
```

5.6 Goodness-of-fit analysis

The goodness-of-fit of mortality models is typically analysed by inspecting the residuals of the fitted model. Regular patterns in the residuals indicate the inability of the model to describe all the features of the data appropriately. With a Poisson or Binomial random component, it is appropriate to look at the scaled deviance residuals defined as:

$$r_{xt} = \text{sign}(d_{xt} - \hat{d}_{xt})\sqrt{\frac{\text{dev}(x,t)}{\hat{\phi}}}, \qquad \hat{\phi} = \frac{D(d_{xt}, \hat{d}_{xt})}{K - \nu},$$

where

$$\text{dev}(x,t) = 2d_{xt}\log\left(\frac{d_{xt}}{\hat{d}_{xt}}\right) - (d_{xt} - \hat{d}_{xt})$$

for a Poisson random component, or

$$\text{dev}(x,t) = 2d_{xt}\log\left(\frac{d_{xt}}{\hat{d}_{xt}}\right) + (E^0_{xt} - d_{xt})\log\left(\frac{E^0_{xt} - d_{xt}}{E^0_{xt} - \hat{d}_{xt}}\right)$$

for a Binomial random component. Further,

$$D(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt}\text{dev}(x,t)$$

is the total deviance of the model, $K = \sum_x \sum_t \omega_{xt}$ is the number of observations in the data and ν is the effective number of parameters in the model.

In **StMoMo** standardised deviance residuals can be obtained with the generic function `residuals` applied to a fitted stochastic mortality model of the class `"fitStMoMo"`. For example, to obtain the residuals of the LC model and of the CBD model fitted before we use the commands:

```
LCres <- residuals(LCfit)
CBDres <- residuals(CBDfit)
```

Graphs of these residuals can be produced using the generic function `plot`. This function supports, via argument `type`, three types of plots:

- scatter plots of residuals by age, period and cohort such as those extensively used in Haberman and Renshaw (2011);

- black and white sign-plots of the residuals such as those used in Cairns et al. (2009) and Lovász (2011); and
- colour maps (heat-maps) of the residuals.

Figure 5.4 presents heat-maps of the deviance residuals for the six models fitted to the England and Wales male mortality experience. These charts were produced using function `plot` with option `type = "colourmap"`. For instance, Figure 5.4b was obtained with the code:

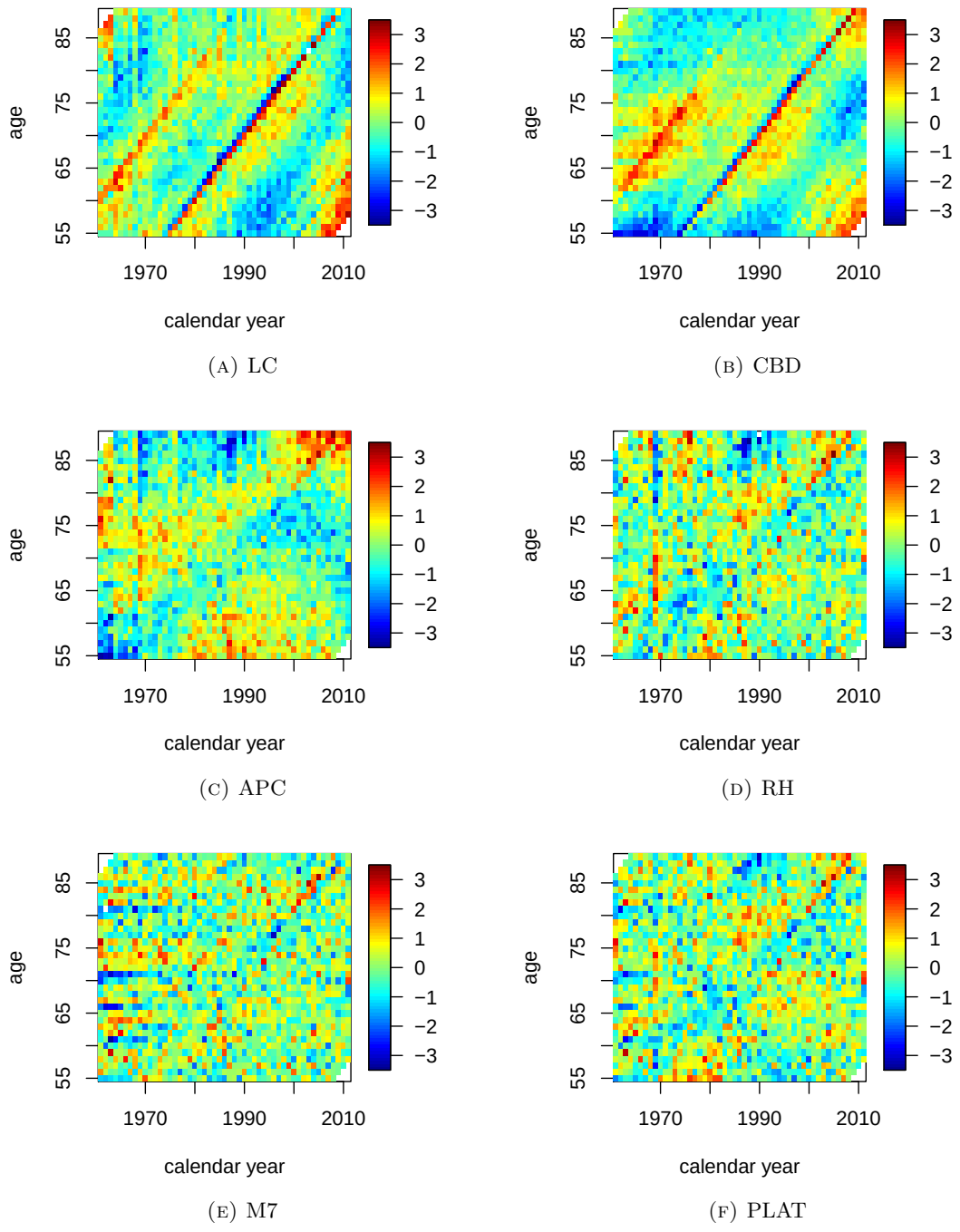
```
plot(CBDres, type = "colourmap", reslim = c(-3.5, 3.5))
```

From Figure 5.4 we see that models LC, CBD and APC display strong residual patterns while the residuals of models RH, M7 and PLAT look reasonably random. The APC model shows a strong clustering of residuals due to its inability to allow for varying improvement rates with age. The LC and CBD models, which do not incorporate a cohort effect, show very marked diagonals patterns indicating the inability of these models to capture the well-known cohort effect observed in the England and Wales population (Willets, 2004). The issues with the fit of the LC and CBD models become more evident when looking at scatter plots of the residuals by age, period and cohort. Such plots for the LC model (Figure 5.5a) and the CBD model (Figure 5.5b) can be produced using argument `type = "scatter"` of function `plot` via the commands:

```
plot(LCres, type = "scatter", reslim = c(-3.5, 3.5))
plot(CBDres, type = "scatter", reslim = c(-3.5, 3.5))
```

The right panels in Figure 5.5 clearly show that the LC and CBD models are unable to capture the cohort effect. In addition, the left panel in Figure 5.5b reveals some strong patterns by age, reflecting the the lack of a quadratic age term in the CBD which may be necessary to capture the commonly observed curvature of the mortality rates in a logit scale.

When evaluating the goodness-of-fit of different models, it is generally anticipated that models with more parameters provide a better fit to the data. To rule out the possibility that the better fit observed in a model is the result of over-parametrisation and compare the relative performance of several models, it has become common in the mortality



(A) LC (B) CBD

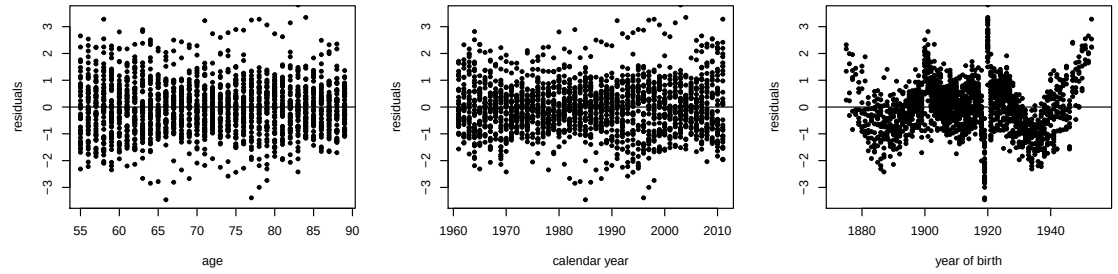
(C) APC (D) RH

(E) M7 (F) PLAT

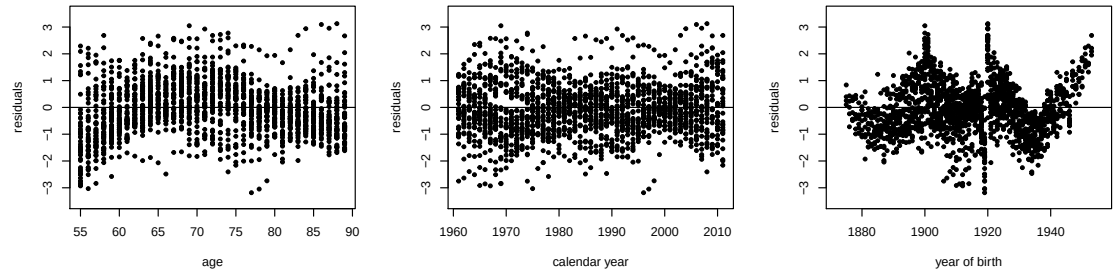
FIGURE 5.4: Heat-maps of deviance residuals for different model fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

literature to use information criteria which modify the maximum likelihood criterion by penalising models with more parameters[13]. Two of these criteria are the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC), defined as $AIC = 2\nu - 2\mathcal{L}$ and $BIC = \nu \log K - 2\mathcal{L}$, respectively, with a lower value of AIC and

[13]For examples of such analysis see Cairns et al. (2009, Section 6.1.1), Haberman and Renshaw (2011, Section 3.3), and Lovász (2011, Section 4.1).



(A) LC



(B) CBD

FIGURE 5.5: Scatter plots of deviance residuals for models LC and CBD fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

BIC being preferable. In **R** these information criteria can be computed using the generic functions `AIC` and `BIC`. For example, we can get the AIC and BIC of the CBD model as follows:

```
AIC(CBDfit)

## [1] 34697.82

BIC(CBDfit)

## [1] 35256.83
```

Table 5.2 presents AIC and BIC values for the six models fitted to the England and Wales male data. We note that both criteria lead to the same ranking of models with M7, PLAT, and RH being the best performing models. Overall, these results are consistent with the existing literature comparing single population models, where the Renshaw-Haberman extension of the Lee-Carter model and the M7 model have been identified as

good candidates for modelling mortality in the England and Wales population (Cairns et al., 2009; Haberman and Renshaw, 2011).

TABLE 5.2: Number of parameters, AIC and BIC values for different model fitted to the England and Wales males population for ages 55-89 and the period 1961-2011

	LC	CBD	APC	RH	M7	PLAT
Number of Parameters	119	102	162	197	229	211
AIC	29866	34698	24469	21779	21406	21624
BIC	30518	35257	25357	22859	22661	22780

5.7 Forecasting and simulation with stochastic mortality models

In the family of GAPC stochastic mortality models the dynamics of mortality are driven by the period indexes $\kappa_t^{(i)}$, $i = 1, \ldots, N$, and the cohort index γ_{t-x}. Therefore, the forecasting and simulation of mortality rates requires the modelling of these indexes using time series techniques.

For the period indexes we follow the standard approach (Cairns et al., 2006, 2011; Haberman and Renshaw, 2011; Lovász, 2011) and use a multivariate random walk with drift. Specifically, we assume that

$$\boldsymbol{\kappa}_t = \boldsymbol{\delta} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}_t^{\kappa}, \qquad \boldsymbol{\kappa}_t = \begin{pmatrix} \kappa_t^{(1)} \\ \vdots \\ \kappa_t^{(N)} \end{pmatrix}, \qquad \boldsymbol{\xi}_t^{\kappa} \sim N(\mathbf{0}, \Sigma), \tag{5.19}$$

where $\boldsymbol{\delta}$ is an N-dimension vector of drift parameters and Σ is the $N \times N$ variance-covariance matrix of the multivariate white noise $\boldsymbol{\xi}_t^{\kappa}$.

As pointed out by Currie (2014), the main challenge when forecasting stochastic mortality models is specifying the dynamics of the cohort effect. To have a simple starting point, we follow previous studies (Renshaw and Haberman, 2006; Cairns et al., 2011; Lovász, 2011) and assume that the cohort index, γ_{t-x}, follows a univariate ARIMA process which is independent of the the period index, $\boldsymbol{\kappa}_t$. In general, we assume that $\gamma_c \equiv \gamma_{t-x}$ follows an ARIMA(p, q, d) with drift, so that

$$\Delta^d \gamma_c = \delta_0 + \phi_1 \Delta^d \gamma_{c-1} + \cdots + \phi_p \Delta^d \gamma_{c-p} + \epsilon_c + \delta_1 \epsilon_{c-1} + \cdots + \delta_q \epsilon_{c-q}, \tag{5.20}$$

TABLE 5.3: ARIMA Models for the cohort effect for models APC, RH, M7 and PLAT.

Mortality Model	**Model for** γ_{t-x}
APC	ARIMA$(1,1,0)$ with drift
RH	ARIMA$(1,1,0)$ with drift
M7	ARIMA$(2,0,0)$ with non-zero intercept
PLAT	ARIMA$(2,0,0)$ with non-zero intercept

where Δ is the difference operator, δ_0 is the drift parameter, $\phi_1, \ldots, \phi_p$ are the autoregressive coefficients with $\phi_p \neq 0$, $\delta_1, \ldots, \delta_q$ are the moving average coefficients with $\delta_q \neq 0$ and ϵ_c is a Gaussian white noise process with variance σ_ϵ.

The time series models in (5.19) and (5.20) can be used to obtain projected (simulated) values of the period index $\dot{\boldsymbol{\kappa}}_{t_n+s} := \left(\dot{\kappa}^{(1)}_{t_n+s}, \ldots, \dot{\kappa}^{(N)}_{t_n+s}\right)'$ and cohort index $\dot{\gamma}_{t_n+s-x_1}$, $s = 1, \ldots, h$, respectively, to derive forecasted (simulated) values of the predictor

$$\dot{\eta}_{x,t_n+s} = \alpha_x + \sum_{i=1}^{N} \beta_x^{(i)} \dot{\kappa}^{(i)}_{t_n+s} + \beta_x^{(0)} \dot{\gamma}_{t_n+s-x},$$

which can in turn be used to obtain forecasted (simulated) age-specific central mortality rates, $\dot{\mu}_{x,t_n+s}$ or age-specific one-year mortality rates, $\dot{q}_{x,t_n+s}$.

In the package **StMoMo** the forecasting of GAPC stochastic mortality models is implemented via the generic method `forecast`. This function estimates and forecasts the multivariate random walk with drift in Equation (5.19) using the approach described in Haberman and Renshaw (2011, Appendix B)[14] and uses function `Arima` of package **forecast** (Hyndman and Khandakar, 2008) to estimate and forecast the ARIMA process of Equation (5.20). For instance, if we assume that the cohort indexes of the APC, RH, M7, and PLAT model follow the ARIMA processes specified in Table 5.3, 50-year ahead ($h = 50$) central projections of the period indexes, cohort index, and one-year death probabilities for the England and Wales mortality experience can be obtain with the code:

```
LCfor <- forecast(LCfit, h = 50)
CBDfor <- forecast(CBDfit, h = 50)
```

[14] This is implemented in function `mrwd` of the package **StMoMo**.

```
APCfor <- forecast(APCfit, h = 50, gc.order = c(1, 1, 0))
RHfor <- forecast(RHfit, h = 50, gc.order = c(1, 1, 0))
M7for <- forecast(M7fit, h = 50, gc.order = c(2, 0, 0))
PLATfor <- forecast(PLATfit, h = 50, gc.order = c(2, 0, 0))
```

The output from the function `forecast` is an object of the class `"forStMoMo"` including, among other things, the following information:

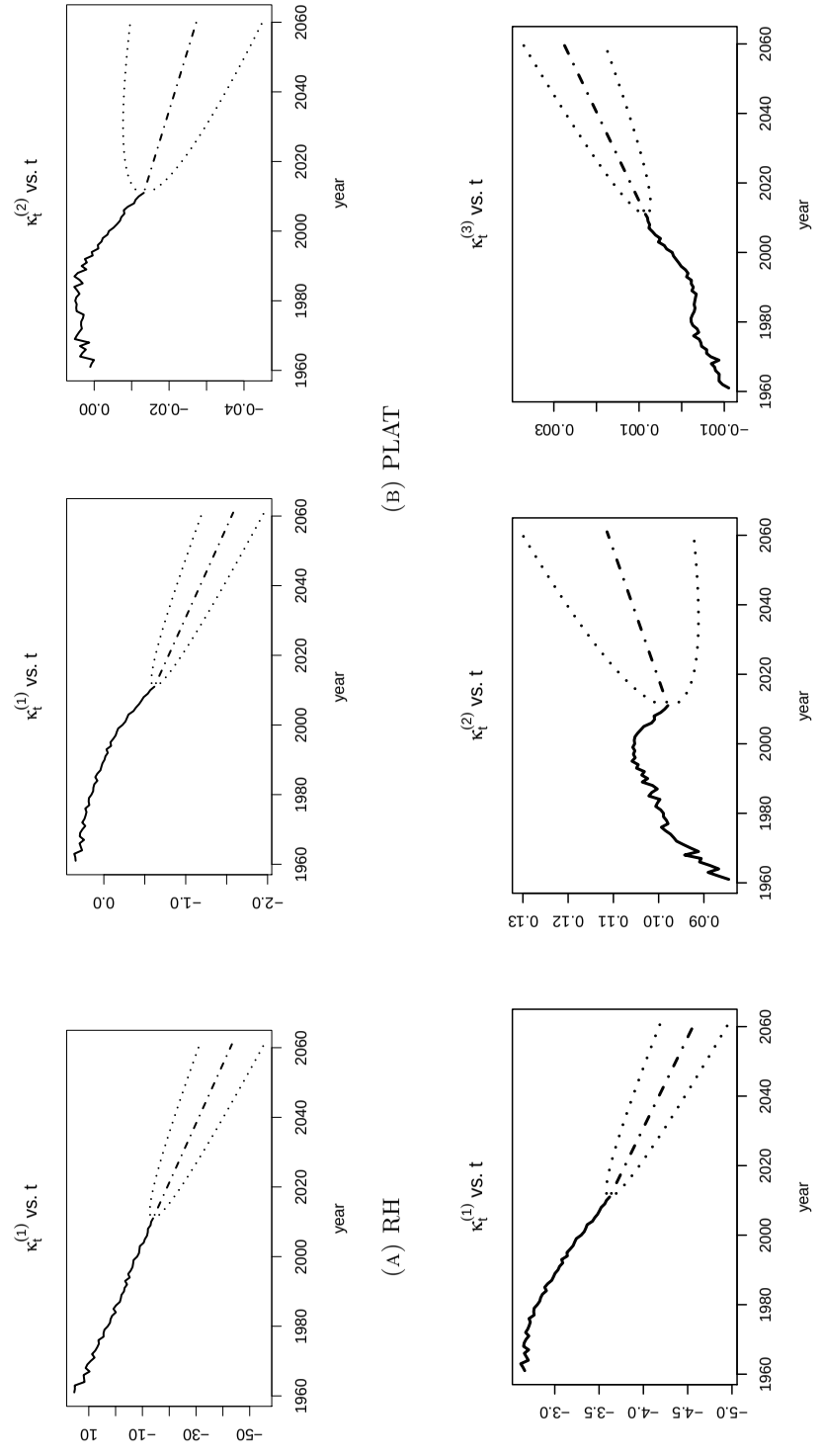
- `rates`: a matrix with the central projection of the mortality rates, $\dot{\mu}_{x,t_n+s}$ or $\dot{q}_{x,t_n+s}$, $s = 1, \ldots, h$;
- `kt.f`: a list containing information on the multivariate random walk with drift fitted to the period index $\boldsymbol{\kappa}_t$; and
- `gc.f`: a list containing information on the ARIMA model fitted to the cohort index γ_{t-x}.

There are `print` and `plot` methods for the `"forStMoMo"` class. For instance, plots of the forecast of the period indexes of models RH, M7 and PLAT (see Figure 5.6) can be produced using the code:

```
plot(RHfor, only.kt = TRUE)
plot(M7for, only.kt = TRUE, nCol = 3)
plot(PLATfor, only.kt = TRUE)
```

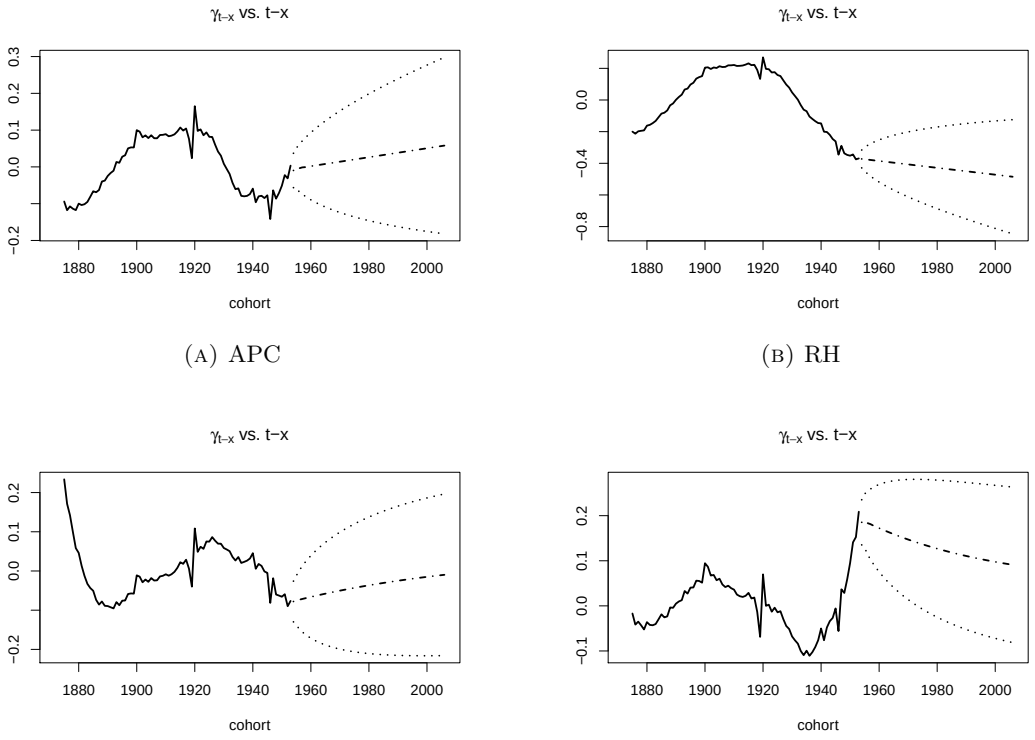
Similarly, plots of the forecast of the cohort indexes of the APC, RH, M7 and PLAT models (see Figure 5.7) can be obtained with the commands:

```
plot(APCfor, only.gc = TRUE)
plot(RHfor, only.gc = TRUE)
plot(M7for, only.gc = TRUE)
plot(PLATfor, only.gc = TRUE)
```



(A) RH (B) PLAT (C) M7

FIGURE 5.6: Forecast of the period indexes of the RH, M7 and PLAT models applied to the England and Wales males population for ages 55-89 and the period 1961-2011. Dashed lines represent central forecast and dotted lines represent 95% prediction intervals.



(A) APC (B) RH

(C) M7 (D) PLAT

FIGURE 5.7: Forecast of the cohort indexes of the APC, RH, M7 and PLAT models applied to the England and Wales males population for ages 55-89 and the period 1961-2011. Dashed lines represent central forecast and dotted lines represent 95% prediction intervals.

Package **StMoMo** also provides the function `simulate` for simulating trajectories from GAPC stochastic mortality models. To simulate the period index, $\boldsymbol{\kappa}_t$, **StMoMo** implements a multivariate adaptation of Algorithm 2 in Haberman and Renshaw (2009) without provision for parameter error[15], while to simulate the cohort index, γ_{t-x}, function `simulate` uses the equivalent `S3` method for objects of class `"Arima"` provided by the package **forecast**. For example, the code below produces 5000 simulated trajectories for the next 50 years of the six stochastic mortality models fitted previously to the England and Wales male mortality experience:

```
set.seed(1234)
nsim <- 5000
LCsim <- simulate(LCfit, nsim = nsim, h = 50)
CBDsim <- simulate(CBDfit, nsim = nsim, h = 50)
```

[15] We note that Algorithm 2 in Haberman and Renshaw (2009) is itself an adaptation of the prediction interval approach of Cairns et al. (2006).

```
APCsim <- simulate(APCfit, nsim = nsim, h = 50, gc.order= c(1, 1, 0))
RHsim <- simulate(RHfit, nsim = nsim, h = 50, gc.order= c(1, 1, 0))
M7sim <- simulate(M7fit, nsim = nsim, h = 50, gc.order= c(2, 0, 0))
PLATsim <- simulate(PLATfit, nsim = nsim, h = 50, gc.order= c(2, 0, 0))
```

The output from the function `simulate` is an object of the class `"simStMoMo"` including, among other things, the following information:

- `rates`: a three dimensional array with the future simulated mortality rates;
- `kt.s`: a list containing information on the simulated paths of the period index $\boldsymbol{\kappa}_t$; and
- `gc.s`: a list containing information on the simulated paths of the cohort index γ_{t-x}.

This output can be used to extract sample trajectories from a model. For instance, Figure 5.8, which depicts 20 trajectories of the period index, cohort index and one-year death rates at age 65 from model RH, was produced with the code:

```
#Plot period index
par(mfrow = c(1, 3))
plot(RHfit$years, RHfit$kt[1, ],
     xlim = range(RHfit$years, RHsim$kt.s$years),
     ylim = range(RHfit$kt, RHsim$kt.s$sim[1, , 1:20]),
     type = "l", xlab = "year", ylab = "kt", main = "Period index")
matlines(RHsim$kt.s$years, RHsim$kt.s$sim[1, , 1:20], type = "l",
         lty = 1)
#Plot cohort index
plot(RHfit$cohorts, RHfit$gc,
     xlim = range(RHfit$cohorts, RHsim$gc.s$cohorts),
     ylim = range(RHfit$gc, RHsim$gc.s$sim[, 1:20], na.rm = TRUE),
     type = "l", xlab = "year", ylab = "kt",
     main = "Cohort index (ARIMA(1,1,0) with drift)")
matlines(RHsim$gc.s$cohorts, RHsim$gc.s$sim[, 1:20], type = "l", lty = 1)
```

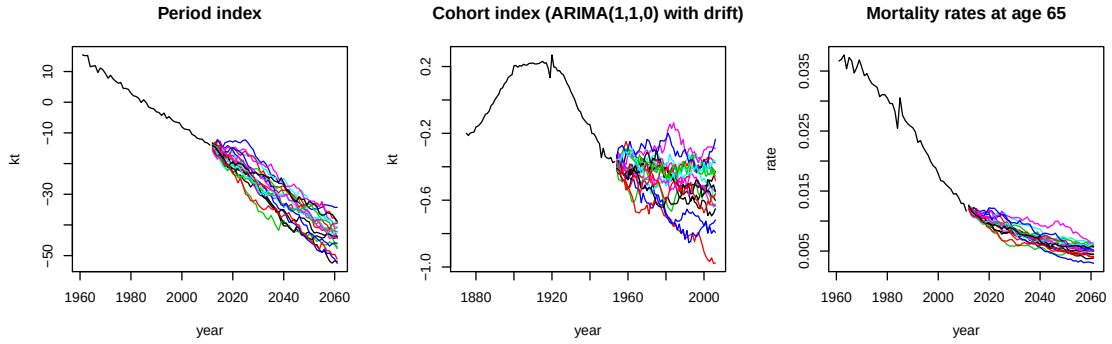


FIGURE 5.8: 20 simulated trajectories of the period index $\kappa_t^{(1)}$, cohort index γ_{t-x} and one-year death rates at age 65 q_{xt} for model RH fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

```
#Plot rates at age 65
qxt <- Dxt / Ext
plot(RHfit$years, qxt["65", ], xlim = range(RHfit$years, RHsim$years),
     ylim = range(qxt["65", ], RHsim$rates["65", , 1:20]), type = "l",
     xlab = "year", ylab = "rate", main = "Mortality rates at age 65")
matlines(RHsim$years, RHsim$rates["65", , 1:20], type = "l", lty = 1)
```

We can also use the output from function `simulate` to produce fan charts depicting the uncertainty associated with a model forecast. According to Cairns et al. (2011), such plots are central to the analysis of the plausibility of the forecast from a model, and can be used as a criterion when deciding upon what is the most appropriate model amongst a group of possible stochastic mortality models. Figure 5.9 shows fan charts depicting 50%, 80% and 95% prediction intervals for mortality rates at ages 65, 75 and 85 for each of the six models fitted to the England and Wales experience. Figure 5.9b, for instance, was produced with the code:

```
library(fanplot)
probs = c(2.5, 10, 25, 50, 75, 90, 97.5)
qxt <- Dxt / Ext
#age 65
plot(CBDfit$years, qxt["65", ], xlim = c(1960, 2061),
     ylim = c(0.0025, 0.2), xlab = "year",
     ylab = "mortality rate (log scale)", pch = 20, log = "y")
```

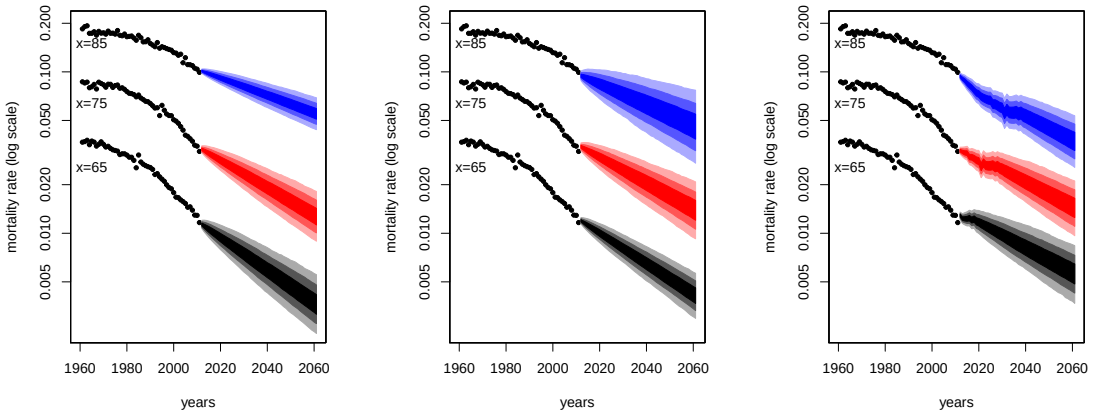
```
fan(t(CBDsim$rates["65", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("black", "white")), ln = NULL)
#age 75
points(CBDfit$years, qxt["75", ], pch = 20)
fan(t(CBDsim$rates["75", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("red", "white")), ln = NULL)
#age 75
points(CBDfit$years, qxt["85", ], pch = 20)
fan(t(CBDsim$rates["85", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("blue", "white")), ln = NULL)
#labels
text(1965, qxt[c("65", "75", "85"), "1990"],
     labels = c("x = 65", "x = 75", "x = 85"))
```

From Figure 5.9 we note the following:

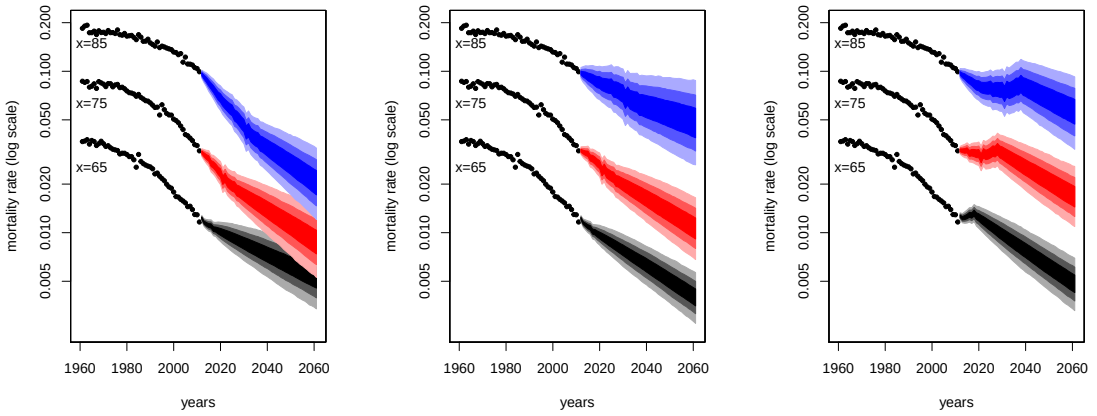
- Whilst for models CBD, RH, M7 and PLAT the fans at age 85 are wider than at age 65 in accordance with historical evidence (see Cairns et al. (2011, Appendix B)), for models LC and APC the fans at age 85 are narrower than at age 65. This suggest that forecasts from models LC and APC are not plausible for the dataset used in this chapter.
- Forecasts for the PLAT model show an implausible increase of mortality rates. This is because the central trend is linked to the estimated cohort effect γ_{t-x} for the PLAT model (see Figure 5.7d) which shows a steep upward trend between 1935 and 1955.
- The central trend and uncertainty levels produced by the each of the models have noticeably differences. This highlights the importance of recognising model risk as a significant issue when modelling mortality (Cairns et al., 2011).

5.8 Parameter uncertainty and bootstrapping

When analysing the uncertainty in mortality projections in an actuarial context it is important to consider all sources of risk. However, the prediction intervals (fan charts)



(A) LC (B) CBD (C) APC



(D) RH (E) M7 (F) PLAT

FIGURE 5.9: Fan charts for mortality rates q_{xt} at ages $x = 65$ (bottom fan), $x = 75$ (middle fan) and $x = 85$ (top fan) from the six models fitted to the England and Wales males population for ages 55-89 and the period 1961-2011. The dots show historical mortality rates for 1961-2011. Shades in the fan represent prediction intervals at the 50%, 80% and 95% level.

obtained in the previous section only account for the uncertainty arising from the error in the forecast of the period and cohort indexes and ignore the uncertainty arising from the estimation of the parameters of the GAPC model.

Due to the analytical intractability of many stochastic mortality models, parameter uncertainty is typically accounted for using the bootstrap procedure. This procedure yields B samples $\alpha_x^b, \beta_x^{(1),b}, \ldots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \ldots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b$, $b = 1, \ldots, B$, of the parameters of the GAPC model which can then be used to produce confidence and prediction intervals of demographic and actuarial quantities. To sample the parameters of GAPC models we consider the semiparametric bootstrap proposed by Brouhns et al. (2005) and the residual bootstrap first considered in Koissi et al. (2006):

- **Semiparametric bootstrap.** Brouhns et al. (2005) propose a semiparametric bootstrap where first B samples of the number of deaths d_{xt}^b, $b = 1, \ldots, B$, are generated by sampling from the Poisson Distribution with mean d_{xt}. Each bootstrapped sample d_{xt}^b, $b = 1, \ldots, B$, is then used to re-estimate the model to obtain B bootstrapped parameter estimates

$$\alpha_x^b, \beta_x^{(1),b}, \ldots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \ldots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b, \quad b = 1, \ldots, B.$$

  Renshaw and Haberman (2008) use the fitted number of death $\hat{d}_{xt}$ (instead of the observed deaths d_{xt}) to perform the sampling from the Poisson distribution. Wang and Lu (2005) consider a similar semiparametric approach using a Binomial distribution of deaths.

- **Residual bootstrap.** Another possibility is to bootstrap the residuals of the model as suggested by Koissi et al. (2006). Under this approach the deviance residuals r_{xt} are resampled with replacement to generate B replications r_{xt}^b, $b = 1, \ldots, B$, which are mapped to the corresponding resampled death counts d_{xt}^b, $b = 1, \ldots, B$, using the inverse formula. The model is then re-fitted using these resampled number of deaths to produce B sets of estimated parameters

$$\alpha_x^b, \beta_x^{(1),b}, \ldots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \ldots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b, \quad b = 1, \ldots, B.$$

  We refer to Koissi et al. (2006) and Renshaw and Haberman (2008) for details on the inverse formula under a Poisson distribution of deaths and to Debón et al. (2010) for the inverse formula under a Binomial framework. Finally, we note that in implementing the residual bootstrap we follow Renshaw and Haberman (2008) and apply the inverse formula to produce samples of the observed number of deaths rather than samples of the fitted number of deaths as originally done by Koissi et al. (2006).

In what follows we illustrate the assessment of the parameter uncertainty using the package **StMoMo**. In doing so we deviate from the England and Wales example we have used so far and use instead New Zealand mortality data. This new example follows closely the work of Li (2014) who uses New Zealand mortality data to compare several

simulation strategies for assessing the risk in mortality projections with a Poisson Lee-Carter model. The main rationale for the change of dataset is that parameter uncertainty is particularly important when analysing the mortality of smaller populations such as smaller countries or pension plans[16]. This new example also serves as a means for illustrating the use of the **StMoMo** package with other datasets.

Mortality data for New Zealand can be extracted from the Human Mortality Database (2014) using function `hmd.mx` of the **demography** package with the code:

```
library(demography)
NZdata <- hmd.mx(country = "NZL_NP", username = username,
                 password = password)
```

We note that the `username` and `password` above are for the Human Mortality Database and should be replaced appropriately. Following Li (2014), we fit a Poisson Lee-Carter model to New Zealand male data for ages 0 to 89 and for the period 1985 to 2008. This can be carried out with the commands:

```
Ext_NZ <- NZdata$pop$male
Dxt_NZ <- NZdata$rate$male * Ext_NZ
LCfit_NZ <- fit(lc(), Dxt = Dxt_NZ, Ext = Ext_NZ, ages = NZdata$age,
            years = NZdata$year, ages.fit = 0:89, years.fit = 1985:2008)
```

In the **StMoMo** package the bootstrap of GAPC stochastic mortality models is implemented with the generic function `bootstrap`. This functions supports both the semiparametric bootstrap and the residual bootstrap. For instance, 5000 semiparametric bootstrap samples of the Lee-Carter model can be obtained with the code[17]:

```
LCboot_NZ <- bootstrap(LCfit_NZ, nBoot = 5000, type = "semiparametric")
```

[16] While the population of England and Wales in 2008 was 54.8 million, the population of New Zealand in 2008 was 4.3 million.

[17] We note that the bootstrap is a computer intensive procedure. In particular, the 5000 semiparametric bootstrap samples of the Lee-Carter model took about two hours to run in a computer with an Intel Core i5-3320m processor running at 2.60 GHz under Windows 7 Home Premium Edition (64 bits) with 8 GB of RAM.

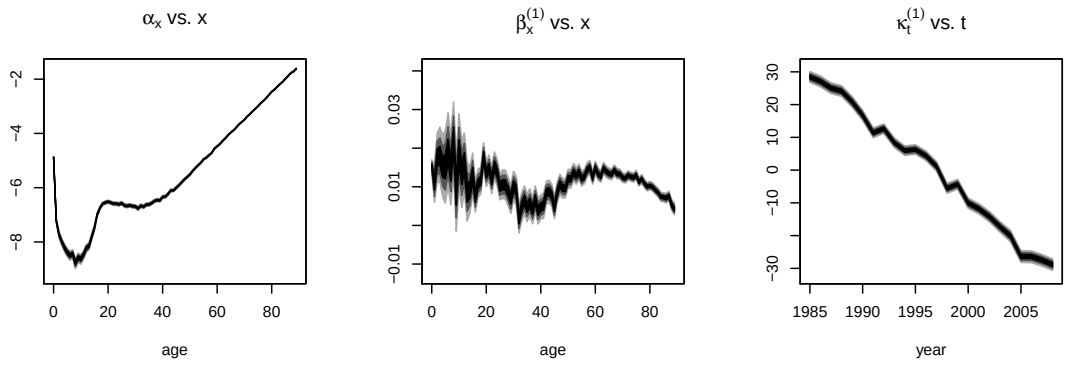


FIGURE 5.10: Bootstrapped parameters for the Poisson Lee-Carter model fitted to the New Zealand male population for ages 0-89 and the period 1985-2008. Shades in the fan represent confidence intervals at the 50%, 80% and 95% level.

The output from function `bootstrap` is an object of the class `"bootStMoMo"` in which the component `bootParameters` contains the `nBoot` replications of the bootstrap parameters. A fan chart depicting the 50%, 80% and 95% confidence intervals of the bootstrapped Lee-Carter model (Figure 5.10) can be obtained with the command:

```
plot(LCboot_NZ)
```

From Figure 5.10 we note that whilst the parameter uncertainty in the static age function α_x and the period index $\kappa_t^{(1)}$ is modest, the uncertainty in the age-modulating parameters $\beta_x^{(1)}$ is more significant.

Once a stochastic mortality model has been bootstrapped we can simulate it forward to obtain simulated trajectories which account for both the forecast error in the period and cohort indexes and the error in the model fitting. In **StMoMo** we can accomplish this using the function `simulate` applied to an object of class `"bootStMoMo"`. Thus, to obtain 5000 simulated trajectories of the Lee-Carter model for the next 24 years taking into account parameter uncertainty we use the instruction:

```
LCsimPU_NZ <- simulate(LCboot_NZ,  h = 24)
```

To highlight the impact of parameter risk on mortality rate projections it is instructive to compare prediction intervals with and without the allowance of parameter uncertainty. 24-year ahead central forecast together with 5000 trajectories of the Lee-Carter model allowing only for forecast error in the random walk with drift and ignoring the model fitting error can be obtained with the code:

```
LCfor_NZ <- forecast(LCfit_NZ, h = 24)
LCsim_NZ <- simulate(LCfit_NZ, nsim = 5000, h = 24)
```

Figure 5.11 depicts 95% prediction intervals for mortality rates at age 40, 60 and 80 with and without allowance for parameter uncertainty. This graph was produced with the code:

```
#Observed, fitted and central forecasts
mxt <- LCfit_NZ$Dxt / LCfit_NZ$Ext
mxtHat <- fitted(LCfit_NZ, type = "rates")
mxtCentral <- LCfor_NZ$rates
#95% Prediction intervals without parameter uncertainty
mxtPred2.5 <- apply(LCsim_NZ$rates, c(1, 2), quantile, probs = 0.025)
mxtPred97.5 <- apply(LCsim_NZ$rates, c(1, 2), quantile, probs = 0.975)
#95% intervals with parameter uncertainty (in sample, and predictions)
mxtHatPU2.5 <- apply(LCsimPU_NZ$fitted, c(1, 2), quantile, probs = 0.025)
mxtHatPU97.5 <- apply(LCsimPU_NZ$fitted, c(1, 2), quantile,
                      probs = 0.975)
mxtPredPU2.5 <- apply(LCsimPU_NZ$rates, c(1, 2), quantile, probs = 0.025)
mxtPredPU97.5 <- apply(LCsimPU_NZ$rates, c(1, 2), quantile,
                      probs = 0.975)
#Plot
x <- c("40", "60", "80")
matplot(LCfit_NZ$years, t(mxt[x, ]),
        xlim = range(LCfit_NZ$years, LCfor_NZ$years),
        ylim = range(mxtHatPU97.5[x, ], mxtPredPU2.5[x, ], mxt[x, ]),
        type = "p", xlab = "years", ylab = "mortality rates (log scale)",
        log = "y", pch = 20, col = "black")
matlines(LCfit_NZ$years, t(mxtHat[x, ]), lty = 1, col = "black")
matlines(LCfit_NZ$years, t(mxtHatPU2.5[x, ]), lty = 5, col = "red")
matlines(LCfit_NZ$years, t(mxtHatPU97.5[x, ]), lty = 5, col = "red")
matlines(LCfor_NZ$years, t(mxtCentral[x, ]), lty = 4, col = "black")
matlines(LCsim_NZ$years, t(mxtPred2.5[x, ]), lty = 3, col = "black")
matlines(LCsim_NZ$years, t(mxtPred97.5[x, ]), lty = 3, col = "black")
```

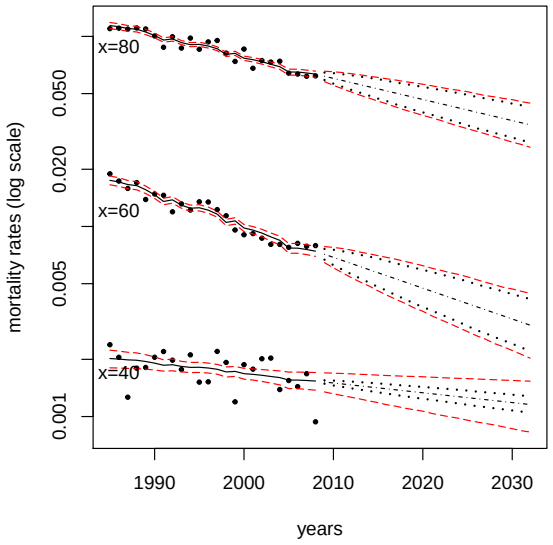


FIGURE 5.11: 95% Prediction intervals for mortality rates q_{xt} at ages $x = 40$ (bottom lines), $x = 60$ (middle lines) and $x = 80$ (top lines) for the Poisson Lee-Carter model fitted to the New Zealand male population for ages 0-89 and the period 1985-2008. Dots show historical mortality rates for 1985-2008 and solid black lines show the corresponding fitted rates. Dashed lines represent central forecast, black dotted lines represent 95% prediction intervals excluding parameter uncertainty and dot-dashed red lines depict 95% confidence and prediction intervals including parameter uncertainty.

```
matlines(LCsimPU_NZ$years, t(mxtPredPU2.5[x, ]), lty = 5, col = "red")
matlines(LCsimPU_NZ$years, t(mxtPredPU97.5[x, ]), lty = 5, col = "red")
text(1986, mxtHatPU2.5[x, "1995"], labels = c("x=40", "x=60", "x=80"))
```

In Figure 5.11 we can clearly see that parameter uncertainty has an important impact on the prediction intervals. This is particularly notable at age 40 where in year 2030, for instance, the width of the prediction interval with parameter uncertainty is around 3 times bigger than that without parameter uncertainty. These results are in line with the results obtained by Li (2014) using the same dataset.[18]

5.9 Conclusion

In this chapter we have introduced the family of Generalised Age-Period-Cohort stochastic mortality models by paralleling the standard framework of generalised linear models. In addition, we have presented the **R** package **StMoMo** which takes advantage of the

[18] We note that our prediction intervals without and with parameter uncertainty correspond to methods (I) and (IVa) in Li (2014), respectively.

unifying framework of the GAPC family to provide tools for fitting a diverse number of stochastic mortality models, assessing their goodness of fit and also performing mortality projections. A key feature of the GAPC family and of **StMoMo** is that they not only encompass models from the Lee-Carter and CBD families, but can also accommodate possible new models. Furthermore, model risk is a prevalent issue when forecasting mortality and we therefore believe that the possibility of easily implementing and comparing a wide range of models makes **StMoMo** a valuable addition to the toolkit for measuring and managing longevity risk.

Our package can be expanded in several directions. The current version of **StMoMo** only allows the use of a log link with a Poisson distribution of deaths or a logit link with a Binomial distribution of deaths. However, we plan to expand the possible combinations of error distribution and link function to include, for instance, Binomial errors with a complementary log-log link as suggested by Currie (2014).

In agreement with the standard in the literature, the **StMoMo** package assumes a multivariate random walk with drift for the dynamics of the period index $\boldsymbol{\kappa}_t$. Nevertheless, the package could be expanded to accommodate more general time series models which might be more appropriate for some datasets or some models. For instance, Renshaw and Haberman (2006) use an ARIMA$(2,1,0)$ for $\kappa_t^{(1)}$ when forecasting mortality for England and Wales males with the Lee-Carter model with cohorts, while Plat (2009) uses general univariate ARIMA(p,d,q) for each $\kappa_t^{(i)}$, $i = 1,\ldots,N$, fitted using seemingly unrelated regressions to allow for correlations between the different period indexes.

In the GAPC family it is assumed that the age-modulating terms $\beta_x^{(i)}$, $i = 0,\ldots,N$, are either non-parametric functions of age which need to be estimated or parametric functions of age with a pre-specified functional form $f^{(i)}(x)$. This latter case could be extended to include the more general case of a pre-specified functional form with a set of parameters that need to be estimated, that is, $\beta_x^{(i)} = f^{(i)}(x;\theta_i)$ with θ_i being some model parameters. This generalisation would allow the implementation of the family of models considered in Hunt and Blake (2014).

Finally, the increasing attention that multiple population mortality models are receiving in the mortality forecasting literature provides an avenue for a wealth of extensions of our package. In particular, we are considering the development of a two-population version

of **StMoMo** that implements the class of relative two-population models considered in Chapter 4.

Acknowledgements

This work stems from research funded by the Institute & Faculty of Actuaries and the Life & Longevity Markets Association. We thank Sveinn Gunnlaugsson and Mario Sison for their assistance in the development and testing of the preliminary code underlying **StMoMo**. We are indebted to Steven Haberman for his encouragement in developing this project and for commenting on a previous version of this chapter. We also thank Anastasios Bardoutsos for his suggestions on the structure of the package and Andrew Hunt for many useful discussions which have helped improved this work. Andrés Villegas thanks Ana Debón for first showing him the great things that can be done with the **R** package **gnm** and selflessly sharing her code and knowledge.

Bibliography

Alai, D. H., Sherris, M., 2014. Rethinking age-period-cohort mortality trend models. Scandinavian Actuarial Journal (3), 208–227.

Aro, H., Pennanen, T., 2011. A user-friendly approach to stochastic mortality modelling. European Actuarial Journal 1, 151–167.

Booth, H., Hyndman, R. J., Tickle, L., 2014. Prospective life tables. In: Charpentier, A. (Ed.), Computational Actuarial Science with R. CRC Press, Boca Raton, FL, pp. 319–344.

Booth, H., Maindonald, J., Smith, L., 2002. Applying Lee-Carter under conditions of variable mortality decline. Population Studies 56 (3), 325–36.

Börger, M., Fleischer, D., Kuksin, N., 2013. Modeling the mortality trend under modern solvency regimes. ASTIN Bulletin 44 (1), 1–38.

Brouhns, N., Denuit, M., Van Keilegom, I., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. Scandinavian Actuarial Journal (3), 212–224.

Brouhns, N., Denuit, M., Vermunt, J., 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics 31 (3), 373–393.

Butt, Z., Haberman, S., Shang, H. L., 2014. ilc: Lee-Carter Mortality Models using Iterative Fitting Algorithms.
URL http://cran.r-project.org/package=ilc

Cairns, A. J., Blake, D., Dowd, K., 2006. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. Journal of Risk and Insurance 73 (4), 687–718.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2011. Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics 48 (3), 355–367.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1–35.

Camarda, C., 2012. MortalitySmooth: An R package for smoothing Poisson counts with P-splines. Journal of Statistical Software 50 (1), 1–24.

Clayton, D., Schifflers, E., 1987. Models for temporal variation in cancer rates. II: Age-period-cohort models. Statistics in Medicine 6 (4), 469–81.

Currie, I. D., 2006. Smoothing and forecasting mortality rates with P-splines.
URL http://www.macs.hw.ac.uk/~iain/research/talks/Mortality.pdf

Currie, I. D., 2014. On fitting generalized linear and non-linear models of mortality. Scandinavian Actuarial Journal.

Debón, A., Martínez-Ruiz, F., Montes, F., 2010. A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. Insurance: Mathematics and Economics 47 (3), 327–336.

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

Haberman, S., Renshaw, A., 2011. A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics 48 (1), 35–55.

Hobcraft, J., Menken, J., Preston, S., 1982. Age, period, and cohort effects in demography: a review. Population Index 48 (1), 4–43.

Human Mortality Database, 2014. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL www.mortality.org

Hunt, A., Blake, D., 2014. A general procedure for constructing mortality models. North American Actuarial Journal 18 (1), 116–138.

Hunt, A., Blake, D., 2015. On the structure and classification of mortality models. Working paper.

Hyndman, R. J., 2014. demography: Forecasting mortality, fertility, migration and population data.
URL http://cran.r-project.org/package=demography

Hyndman, R. J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software 27 (3), 1–22.

Hyndman, R. J., Ullah, S., 2007. Robust forecasting of mortality and fertility rates: A functional data approach. Computational Statistics & Data Analysis 51 (10), 4942–4956.

Koissi, M., Shapiro, A., Hognas, G., 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. Insurance: Mathematics and Economics 38 (1), 1–20.

Lee, R., Miller, T., 2001. Evaluating the performance of the Lee-Carter method for forecasting mortality. Demography 38 (4), 537–49.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Li, J., 2014. A quantitative comparison of simulation strategies for mortality projection. Annals of Actuarial Science 8 (2), 281–297.

Lovász, E., 2011. Analysis of Finnish and Swedish mortality data with stochastic mortality models. European Actuarial Journal 1 (2), 259–289.

McCullagh, P., Nelder, J., 1989. Generalized Linear Models, 2nd Edition. Chapman & Hal, London.

Oeppen, J., Vaupel, J. W., 2002. Broken limits to life expectancy. Science 296 (5570), 1029–1031.

O'Hare, C., Li, Y., 2012. Explaining young mortality. Insurance: Mathematics and Economics 50 (1), 12–25.

Plat, R., 2009. On stochastic mortality modeling. Insurance: Mathematics and Economics 45 (3), 393–404.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL http://www.r-project.org/

Renshaw, A., Haberman, S., 2003. Lee-Carter mortality forecasting with age-specific enhancement. Insurance: Mathematics and Economics 33 (2), 255–272.

Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556–570.

Renshaw, A., 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.

Renshaw, A., Haberman, S., Hatzopoulos, P., 1996. The modelling of recent mortality trends in United Kingdom male assured lives. British Actuarial Journal 2 (2), 449–477.

Turner, H., Firth, D., 2012. Generalized nonlinear models in R: An overview of the gnm package.
URL http://cran.r-project.org/package=gnm

van Berkum, F., Antonio, K., Vellekoop, M., 2014. The impact of multiple structural changes on mortality predictions. Scandinavian Actuarial Journal.

Wang, D., Lu, P., 2005. Modelling and forecasting mortality distributions in England and Wales using the Lee–Carter model. Journal of Applied Statistics 32 (9), 873–885.

Wikipedia, 2014. King Momo, in Wikipedia, The Free Encyclopedia.
URL http://en.wikipedia.org/w/index.php?title=King_Momo&oldid=635124783

Willets, R., 2004. The cohort effect: Insights and explanations. British Actuarial Journal 10 (4), 833–877.

6

Robustness and Convergence in the Lee-Carter Model with Cohort Effects

This chapter has been published in Insurance: Mathematics and Economics. The full reference follows:

– Hunt, A., Villegas, A. M., 2015. Robustness and convergence in the Lee-Carter model with cohort effects. Insurance: Mathematics and Economics 64. 186-202.

The research presented in this chapter has been carried out in collaboration with Andrew Hunt, a fellow PhD Student at Cass Business School. Both of us have contributed equally to the development of the ideas and to the writing of the associated paper.

Robustness and Convergence in the Lee-Carter Model with Cohort Effects

Andrew Hunt, Andrés M. Villegas

Cass Business School, City University London, United Kingdom

Interest in cohort effects in mortality data has increased dramatically in recent years, with much of the research focused on extensions of the Lee-Carter model incorporating cohort parameters. However, some studies find that these models are not robust to changes in the data or fitting algorithm, which limits their suitability for many purposes. It has been suggested that these robustness problems may be the result of an unresolved identifiability issue. In this chapter, after investigating systemically the robustness of cohort extensions of the Lee-Carter model and the convergence of the algorithms used to fit it to data, we demonstrate the existence of such an identifiability issue and propose an additional approximate identifiability constraint which solves many of the problems found.

Keywords: Lee-Carter model; cohort-effects; mortality modelling; Robustness; convergence

6.1 Introduction

For many decades, actuaries and demographers have studied changing mortality rates across the dimensions of age, period and cohort, for instance in Hobcraft et al. (1982) or Wilmoth (1990). These can often be thought of relating to the biological effects of ageing, the influence of contemporaneous factors and the impact of past history on an individual's mortality, as discussed in Hobcraft et al. (1982). However, for just as long, it has been known that this decomposition was problematic, as the relationship cohort = period − age makes it difficult to clearly separate the mutually dependent changes across the three dimensions. Often, the effect of cohort was relegated to a secondary role with the dimensions of age and period being assumed to be dominant, for instance in the classic model of Lee and Carter (1992).

Interest in the cohort dimension has increased dramatically in recent years, as the financial implications of longevity risk have become more apparent. This has been most notable in the UK, largely due to the work of Willets (1999, 2004) and the Continuous Mortality Investigation (2002). These studies revealed the dramatic improvements in mortality rates experienced by the so-called "golden generation", born in the UK between 1920 and 1940. However, the importance of allowing for cohort effects has also been found in other European countries, such as Sweden and Finland in Lovász (2011) and Belgium and the Netherlands in van Berkum et al. (2014). These striking increases in life expectancy have been attributed to a number of factors, such as the introduction of universal healthcare and the reduction in the prevalence in smoking. Although subsequent studies, such as Murphy (2009, 2010), have cast doubt on the true importance of the "golden generation" in improving life expectancies, the idea and salience of a cohort effect entered the consciousness of the actuarial community. As the fastest increases in life expectancy have been concentrated in generations who have now started to retire, the presence of cohort effects had enormous financial consequences, pushing up the costs of providing benefits for many governments, pension schemes and annuity providers.

The need to understand and project cohort effects consequently spurred the development of mortality models capable of analysing such features in the historical data. The most natural method of achieving this was to extend the widely used Lee-Carter model by introducing a set of cohort parameters in a similar method to the model of Hobcraft et al. (1982). This was first done by Renshaw and Haberman (2006) and later developed in Haberman and Renshaw (2009), and this model has subsequently become widely adopted.

However, during tests comparing the Renshaw and Haberman (2006) model with a number of comparators, the model has come in for criticism. Much of this, for instance in Continuous Mortality Investigation (2007), Cairns et al. (2009, 2011) and Currie (2014), finds that the model was slow to converge to an optimum solution and was not robust to changes in the data. Cairns et al. (2009) conjecture that these problems may be the result of an unresolved identifiability issue. The criticisms of Cairns et al. (2009, 2011) in particular were disputed by Haberman and Renshaw (2011), who suggest that the issues observed in these studies were a result of the fitting procedure used to obtain parameter estimates.

In this chapter, we review the differences in the fitting procedures used in Renshaw and Haberman (2006), Haberman and Renshaw (2009, 2011) and that in Cairns et al. (2009, 2011) and discuss the issues with robustness and convergence found. We perform a battery of tests to investigate systematically these phenomena for the model in Renshaw and Haberman (2006) and a widely used simplification of this model, to test the sensitivity of the fitted parameters to features of the fitting algorithm and the range of data used. In doing so, we find many of the same robustness issues observed in Continuous Mortality Investigation (2007) and Cairns et al. (2009, 2011).

Based on this analysis, we demonstrate mathematically the existence of an additional identifiability issue in models extending the Lee-Carter model with cohort effects, as suggested by Cairns et al. (2009), and quantify its impact on the fit to data. This issue can be resolved by imposing an additional constraint on the parameters of the model, with the effect of improving the robustness and convergence rate of the model and ensuring that the parameters fit with our subjective interpretation of their significance.

The structure of this chapter is as follows. Section 6.2 describes the data and notation used in this study, which has been selected for ease of comparability with the previous work in this field. Section 6.3 discusses the mathematical framework of the models used in this chapter and the known issues with the identifiability of the parameters in them. In Section 6.4, we implement the two-stage fitting procedure originally used in Renshaw and Haberman (2006) to fit the model to data and discuss the problems with such an approach. In Section 6.5, we implement an alternative one-stage fitting procedure, used in Cairns et al. (2009) and other studies, and consider the advantages and drawbacks of this technique. Based on these results, in Section 6.6 we isolate the additional approximate identifiability issue which causes many of the problems experienced in the previous sections, and solve this by imposing an additional constraint on the parameters of the model. In Section 6.7, we summarise the results of similar tests performed on other datasets in order to check the generalisability of these findings. Finally, Section 6.8 summarises our findings and offers some conclusions about the practical use of models extending the Lee-Carter approach to allow for cohort effects.

6.2 Data and notation

In order to ease the comparison with previous studies considering the Lee-Carter model with cohort effects, we illustrate our discussions with historical mortality data for the England and Wales male population covering calendar years 1961-2007 and ages 0-89, as compiled by the UK Government Actuary Department (GAD) and as used by Haberman and Renshaw (2009, 2011). In general, we assume that the available data comprise a cross classified mortality experience containing observed numbers of deaths d_{xt} with matching central exposures e_{xt}. We further assume, consistent with Brouhns et al. (2002), that the random number of deaths, D_{xt}, at age x and for period t is a Poisson distributed random variable, with distribution

$$D_{xt} \sim Poisson(e_{xt}\mu_{xt})$$

and, hence, that $\mu_{xt} = \mathbb{E}(D_{xt})/e_{xt}$.

The data cover k ages, $x \in [x_1, x_k]$, and n calendar years, $t \in [t_1, t_n]$. Year of birth or cohort year, defined as $y = t - x$, is therefore in the range $y \in [t_1 - x_k, t_n - x_1]$. In common with Haberman and Renshaw (2009), we exclude from our model fittings the three earliest and most recent cohorts, along with the 1886 cohort, for which Renshaw and Haberman (2006) have noted some issues with the corresponding exposures.

In Section 6.7, we also consider data for other populations including women in England and Wales, both sexes in the USA, men in the Netherlands, and men in Spain. The data for these populations are in a similar format to that for men in England and Wales, with the female England and Wales data compiled by GAD and the data for the USA, the Netherlands and Spain taken from the Human Mortality Database (2014).

6.3 Age/period/cohort mortality models

The first widely used stochastic mortality model was introduced by Lee and Carter (1992) and proposes that mortality rates μ_{xt} at age x and time t obey the relationship

$$\log \mu_{xt} = \alpha_x + \beta_x \kappa_t \tag{6.1}$$

In this model (referred to subsequently as the LC model), α_x gives the shape of the mortality curve across ages, κ_t controls the rate at which mortality changes with time, and β_x modulates this change across the age range.

A key feature of many stochastic mortality models is that their parameters are not uniquely identified. For the LC model in particular, the parameters in Equation (6.1) can be transformed in the following two ways

$$\{\tilde{\alpha}_x, \tilde{\beta}_x, \tilde{\kappa}_t\} = \left\{\alpha_x, \frac{1}{a}\beta_x, a\kappa_t\right\} \tag{6.2}$$

$$\{\tilde{\alpha}_x, \tilde{\beta}_x, \tilde{\kappa}_t\} = \{\alpha_x - b\beta_x, \beta_x, \kappa_t + b\} \tag{6.3}$$

leaving the fitted mortality rates unchanged. The first transformation implies that the "scale" of β_x and κ_t can be chosen freely, whilst the latter implies that the "level" of α_x and κ_t can also be freely chosen. To overcome this lack of identifiability, Lee and Carter (1992) impose additional constraints on the parameters, namely

$$\begin{aligned} \sum_x \beta_x &= 1 \\ \sum_t \kappa_t &= 0 \end{aligned} \tag{6.4}$$

These identifiability constraints are not unique, however. Alternatives, such as $\kappa_{t_1} = 0$ in place of Equation (6.4), as done by Renshaw and Haberman (2003b), can also be used to obtain alternative parameters which give an identical fit to the data.

A consequence of the constraint in Equation (6.4) is that the least squares estimator for α_x is $\frac{1}{n}\sum_t \log\left(\frac{d_{xt}}{e_{xt}}\right)$, leading to the natural interpretation of α_x as the average level of mortality across the period of the data. However, it is important to note that this result is dependent upon the constraint in Equation (6.4) and the use of least squares, and will not be true with other constraints or for other estimation approaches. For instance, in Renshaw and Haberman (2003b), the constraint $\kappa_{t_1} = 0$ gives the interpretation that α_x represents the fitted mortality rates in the first year of the data.

As discussed in Section 6.1, it is often desirable to include cohort effects in a model to allow for systematic differences between cohorts born in different years. In order to achieve this, Renshaw and Haberman (2006) propose to extend the structure of the LC

model by incorporating an additional bi-linear age/cohort interaction to give

$$\log \mu_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t + \beta_x^{(0)}\iota_{t-x} \tag{6.5}$$

where parameters α_x, $\beta_x^{(1)}$, and κ_t have the same interpretation as their equivalents in the LC model, whilst new parameter ι_{t-x} controls the specific features of mortality for individuals from the same cohort with $\beta_x^{(0)}$ modulating its effect across the age range. This model is referred to as the M2 model in Cairns et al. (2009, 2011) and as model M in Renshaw and Haberman (2006), Haberman and Renshaw (2011), and throughout this chapter.

There exist four transformations of the parameters of model M which leave the mortality rates in Equation (6.5) unchanged, namely

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\beta}_x^{(0)}, \tilde{\iota}_y\} = \left\{\alpha_x, a\beta_x^{(1)}, \frac{1}{a}\kappa_t, \beta_x^{(0)}, \iota_y\right\} \tag{6.6}$$

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\beta}_x^{(0)}, \tilde{\iota}_y\} = \left\{\alpha_x, \beta_x^{(1)}, \kappa_t, b\beta_x^{(0)}, \frac{1}{b}\iota_y\right\} \tag{6.7}$$

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\beta}_x^{(0)}, \tilde{\iota}_y\} = \left\{\alpha_x - c\beta_x^{(1)}, \beta_x^{(1)}, \kappa_t + c, \beta_x^{(0)}, \iota_y\right\} \tag{6.8}$$

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\beta}_x^{(0)}, \tilde{\iota}_y\} = \{\alpha_x - d\beta_x^{(0)}, \beta_x^{(1)}, \kappa_t, \beta_x^{(0)}, \iota_y + d\} \tag{6.9}$$

Transformations (6.6) and (6.7) are similar to that in Equation (6.2), while transformations (6.8) and (6.9) are similar to that in Equation (6.3).

Renshaw and Haberman (2006) also discuss several nested models for model M obtained by setting to a constant one or both of the age adjustments associated with the period index κ_t and the cohort effect ι_y. Of particular interest are the sub-models

$$\log \mu_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t + \frac{1}{k}\iota_{t-x}$$

and

$$\log \mu_{xt} = \alpha_x + \frac{1}{k}\kappa_t + \frac{1}{k}\iota_{t-x},$$

obtained by setting $\beta_x^{(0)} = 1/k$ and $\beta_x^{(1)} = \beta_x^{(0)} = 1/k$, respectively. The first sub-model, called model H1 in Renshaw and Haberman (2006), Haberman and Renshaw (2011) and within this chapter, has been suggested by Haberman and Renshaw (2011) as an

alternative simpler structure that resolves some of the reported criticisms of model M. The second sub-model defines the so-called age-period-cohort (APC) model which has a long history in the fields of medicine and sociology (see, e,g,, Clayton and Schifflers (1987) and Hobcraft et al. (1982)), but has not been widely used in the actuarial mortality literature until considered by Currie (2006).[1]

6.4 The two-stage fitting approach

Motivated by *"the central role played by mortality reduction factors in generating mortality projections"*, Renshaw and Haberman (2006) chose to partition model M as

$$\begin{aligned} \mu_{xt} &= \exp(\alpha_x) F(x,t) \\ \log F(x,t) &= \beta_x^{(1)} \kappa_t + \beta_x^{(0)} \iota_{t-x} \end{aligned} \tag{6.10}$$

where $F(x,t)$ are a set of reduction factors, as discussed in Renshaw and Haberman (2003b). This in turn motivates a two-stage approach to fitting the model:

- first, the main age effects, α_x, are set to be $\alpha_x = \frac{1}{n}\sum_t \log\left(\frac{d_{xt}}{e_{xt}}\right)$,
- then the other parameters are found using an iterative approach which minimises the total deviance (or equivalently maximises the log-likelihood), without updating the α_x.

It is important to note that this two-stage approach will generally lead to poorer fits to the available data than a one-stage approach where all the parameters are estimated together, and are therefore not true maximum likelihood estimates. This is because α_x has not been chosen explicitly to maximise the goodness of fit in conjunction with the other parameters in the model, i.e., the other parameter estimates are conditional on our assumption regarding α_x. This may impact the suitability of any tests of the goodness of fit of the model in comparison with others, such as those performed in Haberman and Renshaw (2011), for instance. This explicit hierarchy in the parameters will also bias any estimation of parameter uncertainty in the model, since α_x is assumed to be known

[1] Renshaw and Haberman (2006) originally derived model H1 by setting $\beta_x^{(0)} = 1$ and the APC model by also setting $\beta_x^{(1)} = 1$. To facilitate graphical comparison between the different models, we choose to set $\beta_x^{(1)} = \beta_x^{(0)} = 1/k$ so that $\sum_x \beta_x^{(1)} = \sum_x \beta_x^{(0)} = 1$ and the scales of parameters κ_t and ι_{t-x} in models M, H1 and the APC model are of comparable magnitude.

and therefore all the parameter uncertainty in the model is allocated to the other terms in the model. However, as discussed in Haberman and Renshaw (2009, Section 2.5), the potential loss of goodness of fit may be justified if it gives a model which is more robust to changes in the data, simpler to fit and whose terms are more explainable in terms of their demographic significance.[2]

6.4.1 Original identifiability constraints

In order to obtain a unique set of parameter estimates, Renshaw and Haberman (2006) impose the constraints

$$\sum_x \beta_x^{(1)} = 1 \tag{6.11}$$

$$\sum_x \beta_x^{(0)} = 1 \tag{6.12}$$

$$\kappa_{t_1} = 0 \tag{6.13}$$

The first two of these constraints set the "scales" of the age functions consistent with the formulation of the original LC model, whilst the third sets the "level" for the period function κ_t at zero at the start of the data period, with negative values of κ_t representing falls from this initial baseline level.[3]

Equations (6.11) and (6.12) are imposed using transformations (6.6) and (6.7) with specific values of the constants a and b, as described in Appendix 6.A. However, under a two-stage fitting approach, Equation (6.13) is not necessary in order to uniquely identify the parameters in the model. More specifically, since the static age function, α_x, is fixed in the first stage, transformations (6.8) and (6.9) which involve α_x are no longer valid, eliminating the need for any additional identifiability constraints. The constraint in Equation (6.13) therefore represents an additional, exogenous constraint on the model, which will reduce the goodness of fit to data when imposed.

It is also worth noting that this constraint also appears inconsistent with how the main age effects were set. As discussed above, the estimator $\alpha_x = \frac{1}{n}\sum_t \log\left(\frac{d_{xt}}{e_{xt}}\right)$ was a

[2] Defined in Hunt and Blake (2015) as the interpretation of the components of a model in terms of the underlying biological, medical or socio-economic causes of changes in mortality rates which generate them.

[3] Renshaw and Haberman (2006) suggest $\iota_{t_1-x_1} = 0$ as an alternative constraint to Equation (6.13). However, it appears from the graphs shown in their paper that Equation (6.13) has been used in practice.

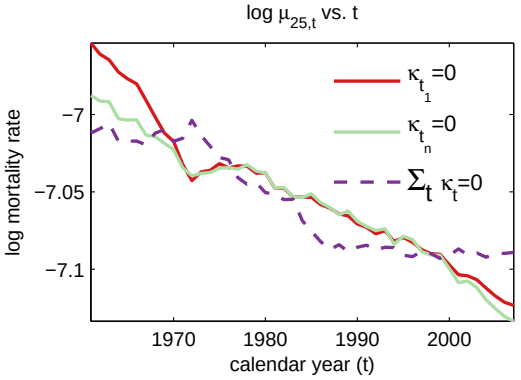


FIGURE 6.1: Fitted log mortality rates at age 25 for model M for different constraints on κ_t fitted using the two-stage approach..

TABLE 6.1: Poisson log-likelihood for model M for different constraints on κ_t fitted using the two-stage approach.

Constraint	Log-Likelihood
$\kappa_{t_1} = 0$	-22,764
$\kappa_{t_n} = 0$	-23,071
$\sum_t \kappa_t = 0$	-23,068

consequence of the constraint $\sum_t \kappa_t = 0$ in the LC model using least squares estimation methods. If we impose $\alpha_x = \frac{1}{n}\sum_t \log\left(\frac{d_{xt}}{e_{xt}}\right)$ on models which nest the LC model, we would therefore expect $\sum_t \kappa_t \approx 0$ to emerge naturally as a consequence.

The imposition of the additional constraint in Equation (6.13) has practical consequences. A true identifiability constraint should not change the fit to data, merely select one out of a range of equivalent parameter sets which give the same fitted mortality rates. However, we can see from Figure 6.1 and Table 6.1 that for alternative (and equally reasonable) choices of a constraint for κ_t we obtain different predicted mortality rates and very different fits to the same data. Not only is the goodness of fit different, but we obtain very different patterns for the parameters, as shown in Figure 6.2. In particular, we note that the apparent impact of the "golden generations" born between 1920 and 1940 and the appropriateness of projecting κ_t in a linear fashion would depend upon the identifiability constraint imposed.

Similar issues have also been found by other users of the two-stage fitting approach. For instance, in an application of model M to a dataset of UK Assured Lives, the Continuous Mortality Investigation (2007) reported that the robustness of the model parameters is affected when different constraints are used. In addition to this lack of robustness of the model to changes in the identifiability constraints, the Continuous Mortality

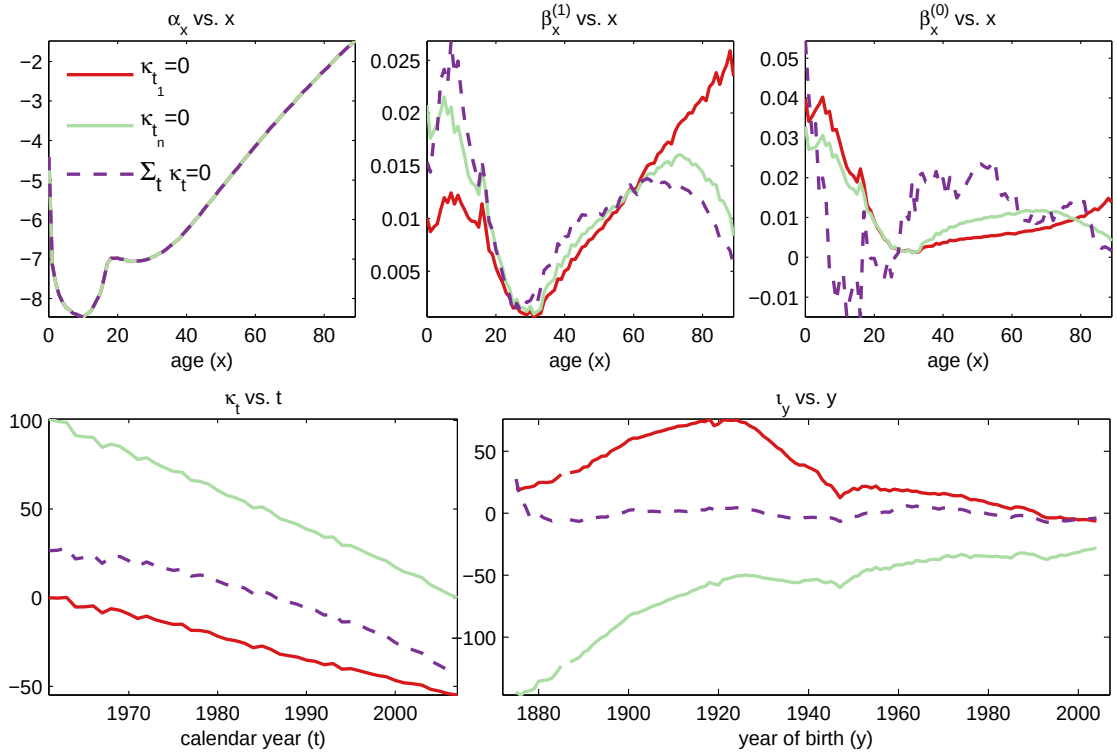


FIGURE 6.2: Parameters for model M for different constraints on κ_t fitted using the two-stage approach.

Investigation (2007) found that they *"were unable to get the model to converge for certain subsets of the data"*, and that *"fitted [parameter] values could differ materially"* when different starting values for the estimation process were used. Similarly, in an application to Spanish mortality data Debón et al. (2010) comment that *"the choice of period or origin cohort influences the convergence of the algorithm."*

Haberman and Renshaw (2011) suggest that some of the problems with model M are resolved using the simpler model H1. However, we have also encountered very similar robustness issues when fitting model H1 using the two-stage procedure of Renshaw and Haberman (2006).

6.4.2 Modifying the two-stage approach

In principle, one obvious solution to the problems of the two-stage method developed in Renshaw and Haberman (2006) would be to fit the model without the unnecessary constraint in Equation (6.13). However, the parameter values obtained using this modified approach are very sensitive to changes in the data used to fit the model. For instance, Figure 6.3 shows that the parameters obtained using data for the periods 1961-2007,

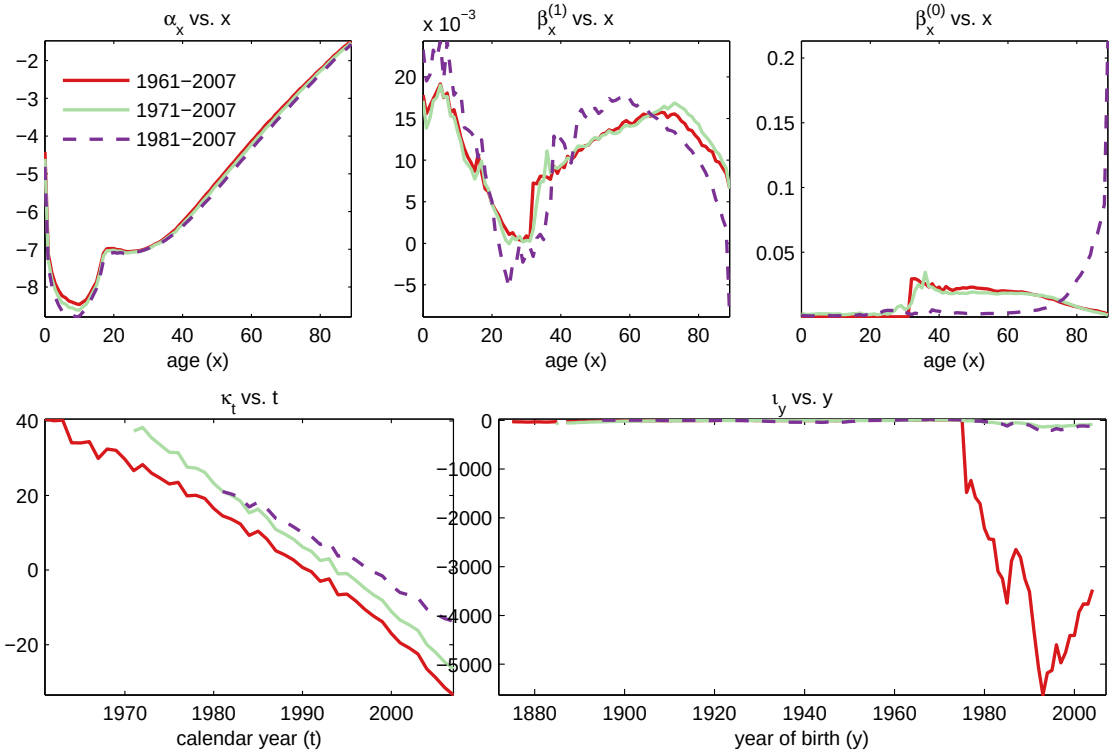


FIGURE 6.3: Parameters for model M for different data periods fitted using the modified two-stage approach.

1971-2007, and 1981-2007 differ significantly. This is specially noticeable for parameters $\beta_x^{(0)}$ and ι_y which define the age/cohort interaction.

The fitted parameters for the period 1961-2007 are particularly interesting. It would be easy to conclude that the model has failed to converge to a viable solution. However, observing the fitted mortality rates rather than the parameters shows a different interpretation. What appears to have happened is that the age function $\beta_x^{(0)}$ has been used to "switch off" the cohort parameters below age 35 for much of the data, i.e., $\beta_x^{(0)} \approx 0$ for $x < 35$. For the most recent years of birth, however, where individuals have yet to reach age 35, the cohort parameters are being used to capture genuine features of data. In order to do so, they must be sufficiently large to give reasonable sized cohort effects when multiplied by the very small values of $\beta_x^{(0)}$. The combined age/cohort interaction term is shown in Figure 6.4.

The model can do this due to the absence of any constraints on ι_y, so there is no mathematical reason why cohort parameters for different years of birth should be of the same order of magnitude, making it very difficult for the model to assign structure in the data as either an age/period or and age/cohort interaction. Doing this however,

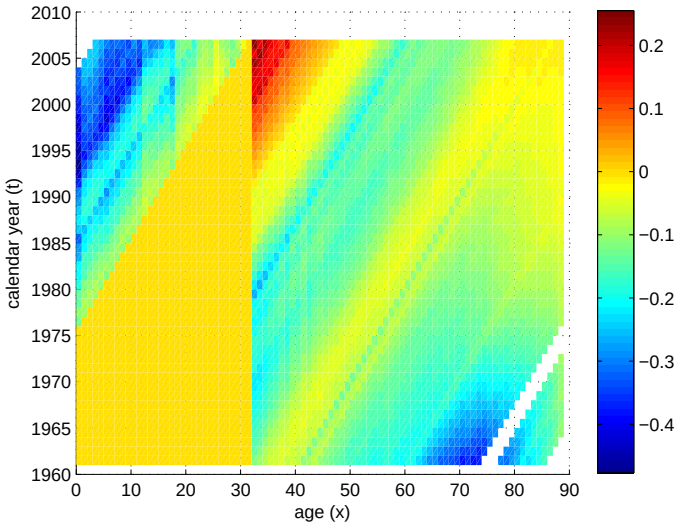


FIGURE 6.4: Combined age/cohort interaction term $\beta_x^{(0)}\iota_{t-x}$ for model M fitted to data for the period 1961-2007 using the modified two-stage approach.

complicates the demographic interpretation of the cohort parameters as they can only be seen in the context of the combined age/cohort term. It will also lead to serious problems when projecting the model, as the extremely large cohort parameters for recent years of birth will be "switched on" at age 35 resulting in them dominating the evolution of mortality rates. For these reasons, such a pattern of cohort parameters is extremely undesirable, even if it gives a plausible fit to the historic data.

We also find that using the simplified model H1 does not solve these robustness issues. Due to this, we do not consider the modified two-stage approach to be a viable fitting procedure for models M and H1. Furthermore, the motives for using a two-stage fitting approach stated by Haberman and Renshaw (2009), namely, obtaining age-effects with interpretable demographic significance and avoiding slow convergence rates for the fitting algorithm, are satisfied in a more theoretically sound way by the approach proposed in Section 6.6 of this chapter.

6.5 The one-stage approach

In their tests of various mortality models, Cairns et al. (2009) adopt an alternative one-stage method to fitting model M. This does not use the reduction factor partition in Equation (6.10) and a two-stage fitting approach, and instead attempts to fit all the

parameters in Equation (6.5) simultaneously using a Poisson likelihood maximisation method. Using this approach should give improved fits to data compared with the two-stage approach, as all the parameters in the model are explicitly chosen to maximise the fit, rather than the main age effects being fixed a priori. It therefore avoids the issues with tests of the goodness of fit and parameter uncertainty that are present in the two-stage approach.

Cairns et al. (2009) also recognise that the model requires additional constraints to be imposed upon the parameters in order to identify them uniquely. They choose to adopt

$$\sum_{t=t_1}^{t_n} \kappa_t = 0 \tag{6.14}$$

$$\sum_{t=t_1}^{t_n} \sum_{x=x_1}^{x_k} \iota_{t-x} = 0 \tag{6.15}$$

in order to set the levels of κ_t and ι_y, along with the constraints in Equations (6.11) and (6.12).

These four constraints are required due to the four parameter transformations in Equations (6.6)-(6.9) which leave the fitted mortality rates in Equation (6.5) unchanged.[4] We note that, as opposed to the two-stage approach in Section 6.4 where constraints (6.11) and (6.12) sufficed, in a one-stage approach the constraints in Equations (6.14) and (6.15) are also required as transformations (6.8) and (6.9) explicitly involve making adjustments to α_x, which is allowable now that the main age effects, α_x, are not fixed in advance.

In our implementation, we use a variation of Equation (6.15) and impose

$$\sum_{y=t_1-x_k}^{t_n-x_1} \iota_y = 0 \tag{6.16}$$

Using this constraint gives an identical fit to the data, but gives the interpretation that the average level of the cohort parameters fitted by the model should be zero.[5] This constraint also generalises naturally to more complicated constraints for other models.

[4] These transformations are then used to impose the constraints during the fitting process, as described in Appendix 6.A.

[5] Specifically, we note that Equation (6.15) implies that $\sum_{y=t_1-x_k}^{t_n-x_1} n_y \iota_y = 0$, where n_y is the number of observations of each cohort, and so represents a weighted average compared to the arithmetic average in Equation (6.16). However, in practice, the difference between the parameter values fitted using these two different constraints is very small.

However, the one-stage approach has been found to lack robustness to changes in the data and take a long time to converge to the solution. For example, Cairns et al. (2009) found that *"parameter values converge very slowly to their maximum likelihood estimates"* and *"the parameter estimates jump to a qualitatively quite different solution when we use less data"*. In Dowd et al. (2010), it was found that *"the forecasts are clearly unstable, ... these projections reflect estimates of the cohort state variables that are sometimes very unstable and highly implausible as we move from one sample to the next"*. Similarly, in Cairns et al. (2011), the authors note that *"the fitting algorithm is generally slow to converge, ... Occasionally, we see that the parameter values jump to a set of values that are qualitatively quite different from the previous year's estimates"*. In Haberman and Renshaw (2009) it was found that with the one-stage approach *"the already slow rate of convergence is made appreciably slower"*. In Currie (2014), it is reported that depending on the choice of identifiability constraints *"convergence could be very fast, very slow or even fail completely"*. Lovász (2011) also finds a lack of robustness with model M, in common with some other models in their study.

In order to investigate systematically the findings of these other studies, in this section we perform tests for convergence and robustness in the M and H1 models. In particular, we test for

- the sensitivity of the fitted parameters to the initial parameter estimates used by the fitting algorithm in Section 6.5.1;
- the sensitivity of the fitted parameters with different convergence criteria for the fitting algorithm in Section 6.5.2;
- the robustness of the fitted parameters to changes in the range of years in the data in Section 6.5.3; and
- the robustness of the fitted parameters to changes in the range of ages in the data in Section 6.5.4.

6.5.1 Sensitivity to initial parameter estimates

We fit the parameters of the model using the maximum likelihood procedure described in Appendix 6.A. This iterative fitting algorithm requires us to have an initial estimate

for the parameters in order to start the iterations. Throughout this chapter we use $\hat{\alpha}_x = 1/n \sum_t \log(d_{xt}/e_{xt})$, $\hat{\beta}_x^{(1)} = 1/k$, $\hat{\kappa}_t = 0$, $\hat{\beta}_x^{(0)} = 1/k$, and $\hat{\iota}_y = 0$ as initial parameter estimates. However, for any reasonable choice of these initial parameter estimates, our choice should not affect the final parameter values found, as these should be the unique set which maximises the log-likelihood.

In order to test this in models M and H1, we rely on the unidentifiability of the parameters of the APC model to generate alternative sets of initial parameters. The APC model is known to be unidentified under the following transformations

$$\begin{aligned}
\{\tilde{\alpha}_x, \tilde{\kappa}_t, \tilde{\iota}_y\} &= \left\{\alpha_x - \frac{c}{k}, \kappa_t + c, \iota_y\right\} \\
\{\tilde{\alpha}_x, \tilde{\kappa}_t, \tilde{\iota}_y\} &= \left\{\alpha_x - \frac{d}{k}, \kappa_t, \iota_y + d\right\} \\
\{\tilde{\alpha}_x, \tilde{\kappa}_t, \tilde{\iota}_y\} &= \left\{\alpha_x + \frac{e}{k}(x - \bar{x}), \kappa_t - e(t - \bar{t}), \iota_y + e(y - \bar{y})\right\}
\end{aligned} \tag{6.17}$$

The first two of these are similar to the transformations in Equations (6.8) and (6.9) for model M, and so are used to impose the same constraints in Equations (6.14) and (6.15) on the APC model as used in Cairns et al. (2009).

Equation (6.17) is unlike any of the transformations described previously. Using it allows us to tilt the parameters α_x, κ_t and ι_y around the midpoints of their ranges as we vary the constant, e. We use this transformation to obtain different sets of parameters from the APC model with different degrees of titling to use as our initial parameter estimates for fitting model M and model H1. These different APC parameter estimates are depicted in Figure 6.5 and the corresponding log-likelihoods are shown in Table 6.2. It is important to note that the different initial parameters are equivalent, in the sense that they produce exactly the same fitted mortality rates and the same goodness of fit when using the APC model. The APC model is a simplification of models M and H1 and so the parameter estimates from it are likely to be qualitatively similar to those fitted to the more complex models.

Using these initial parameter estimates to fit model M,[6] we find the parameters shown in Figure 6.6. The set of parameters labelled *"tilt 3"* is closest to what we would typically consider reasonable from the model. As we deviate from this we see that

[6] With a convergence criteria of 10^{-4} here and for the remainder of this chapter unless stated otherwise.

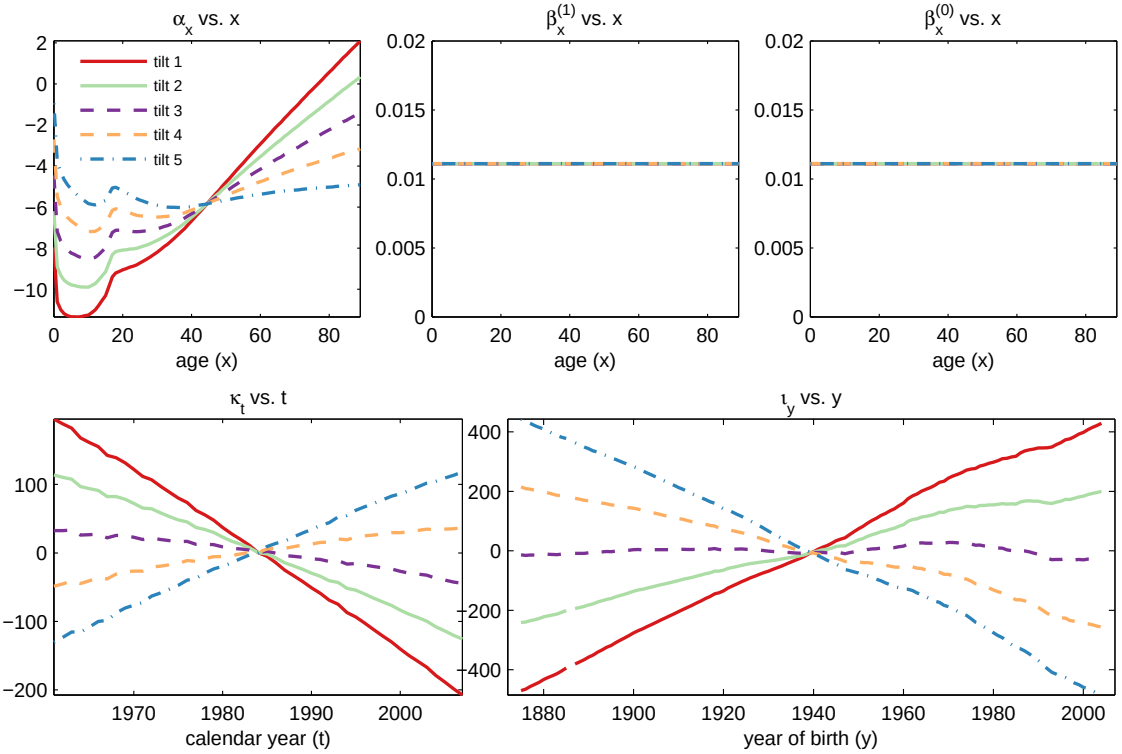


FIGURE 6.5: Parameters for the APC model with different degrees of tilting used as initial parameters estimates.

TABLE 6.2: Poisson log-likelihood for the different models under alternative initial parameter estimates using the one-stage approach.

	Log-likelihood		
Initial parameters	APC Model	Model M	Model H1
tilt 1	-27,997	-21,987	-22,397
tilt 2	-27,997	-21,985	-22,395
tilt 3	-27,997	-21,992	-22,400
tilt 4	-27,997	-22,025	-22,441
tilt 5	-27,997	-22,008	-22,426

- κ_t and ι_y both exhibit similar tilts to the initial parameter estimates. In addition, some sets of κ_t parameters show significant curvature, which may present a problem when projecting the model.

- α_x shows a similar tilt to the initial parameter estimate. It also becomes significantly less smooth, and displays some of the same features as $\beta_x^{(0)}$ (e.g., the spikes around ages 18 and 29).

- $\beta_x^{(1)}$ also starts showing many of the same features as $\beta_x^{(0)}$ in a far more prominent fashion than α_x.

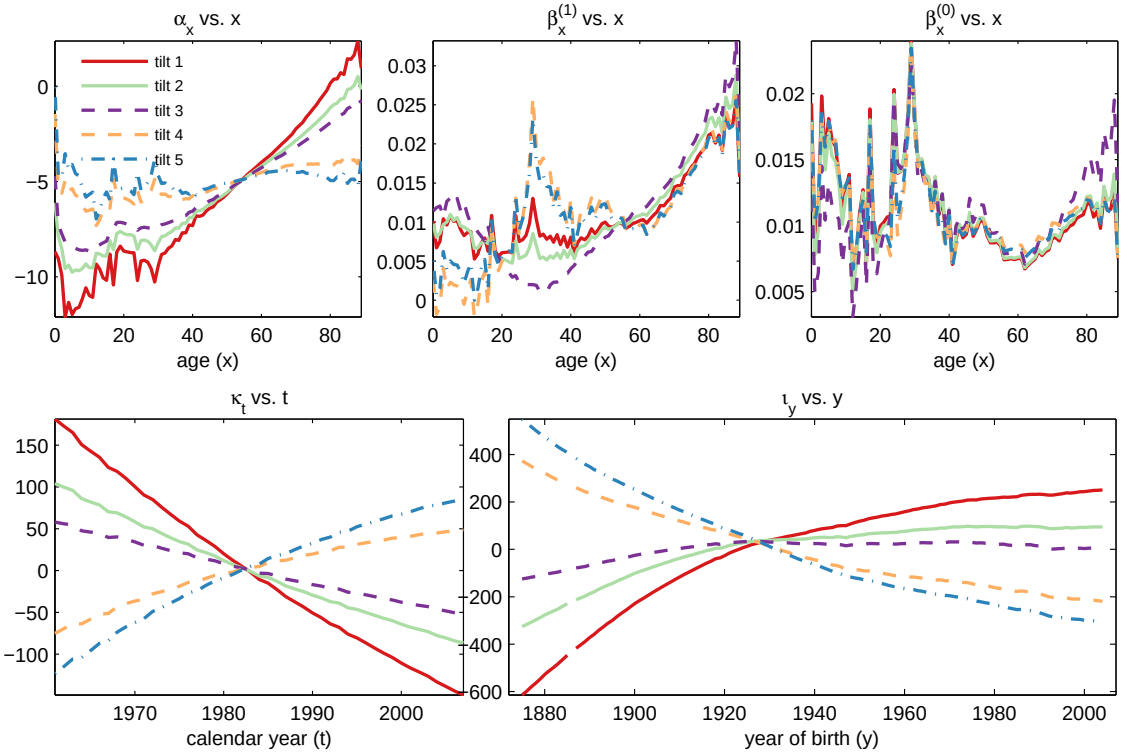


FIGURE 6.6: Parameters for model M with different initial parameters estimates fitted using the one-stage approach.

- $\beta_x^{(0)}$ appears to be robust to changes in the initial parameter estimates using this procedure (although in all cases, our initial parameter estimates were $\beta_x^{(0)} = \frac{1}{k}$).

Table 6.2 shows that using alternative initial parameter estimates in the fitting algorithm gives final parameters with different log-likelihoods. This may indicate that the algorithm has not found a global maximum of the likelihood function, but has instead found the local maximum of the function closest to the initial parameter estimates. Alternatively, it may be that the fitting algorithm has stopped without finding the maximum of the log-likelihood, since the likelihood function is flat in a region around its maximum value.

Figure 6.7 shows the results of a similar analysis for model H1. In this we see many of the same features as observed for model M, with the notable exception of the changes in the smoothness of α_x and $\beta_x^{(1)}$ as the initial parameters are tilted.

These results on the sensitivity to the initial parameters estimates are strongly indicative that models M and H1 can be transformed in a similar manner to the transformation of the APC model in Equation (6.17). This is discussed in Section 6.6.

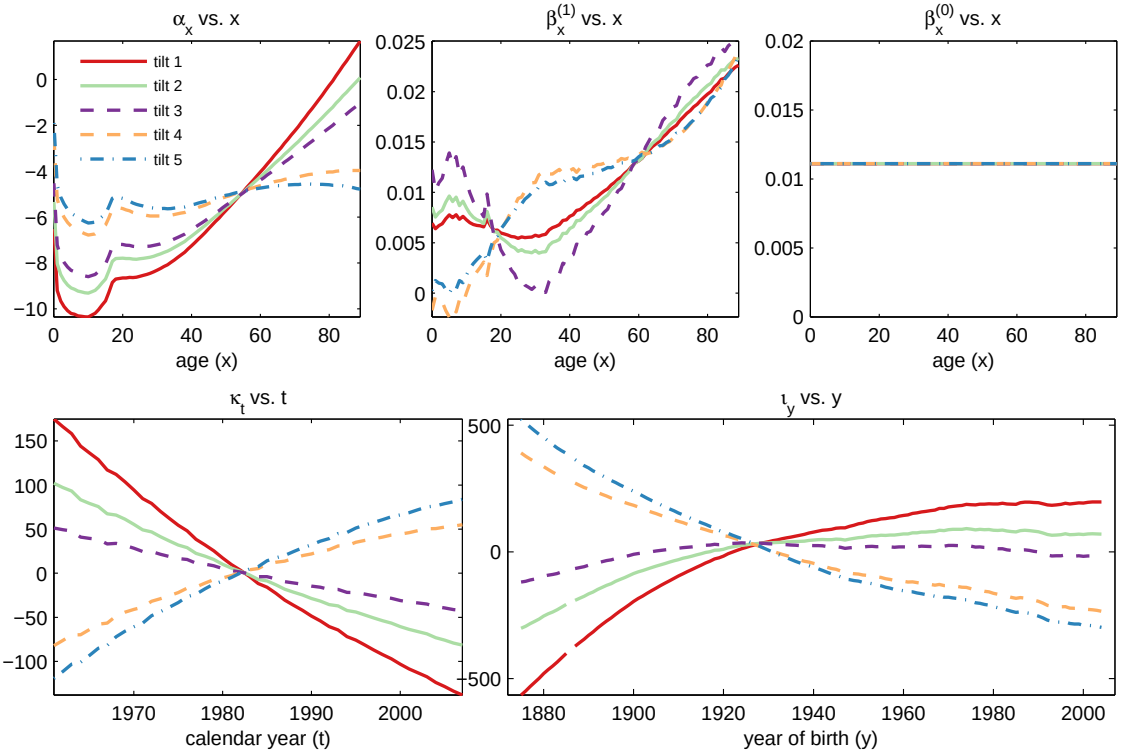


FIGURE 6.7: Parameters for model H1 with different initial parameters estimates fitted using the one-stage approach.

6.5.2 Sensitivity to convergence criteria

The iterative fitting algorithm described in Appendix 6.A stops once the changes in the fitted log-likelihood become smaller than a pre-set tolerance value. Using a smaller value for this convergence criterion should result in more accurately estimated parameters but require the algorithm to run for longer. While the parameter estimates found using a smaller convergence criterion should be more accurate, they should not, however, show significantly different patterns to those observed for higher criteria, but merely be more refined estimates.

Figure 6.8 shows the fitted parameters for model M with different convergence criteria (i.e., the algorithm ceases when the change in log-likelihood in each iteration becomes smaller than 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}). Table 6.3 shows the corresponding fitted log-likelihoods and computing times for models M and H1 using these convergence criteria. As can be seen, allowing for a smaller convergence criterion improves the fitted log-likelihood marginally. However, this is at the expense of far longer times needed to

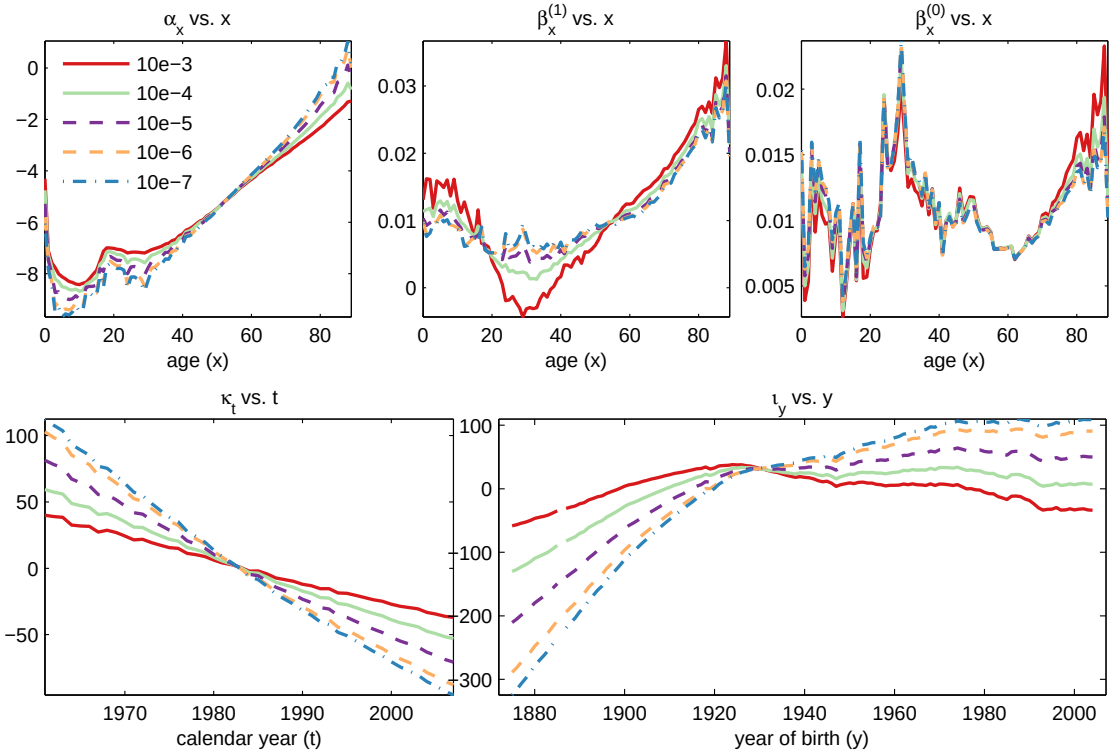


FIGURE 6.8: Parameters for model M for different tolerance levels fitted using the one-stage approach.

fit the model.[7] We note that long fitting times limit the practicality of techniques such as those proposed in Brouhns et al. (2005) and Koissi et al. (2006) for measuring parameter uncertainty, as these involve refitting the model multiple times to randomly generated data. We also see from Figure 6.8 that the estimated parameter values for model M change by more than would be expected for such small improvements in the fit to the data.[8] In particular, we note that many of the issues observed with the parameters are similar to those described in Section 6.5.1 and Figure 6.6. However, it is also interesting to note that $\beta_x^{(0)}$, in spite of its erratic behaviour at younger ages, does not change significantly when we change the convergence criteria. We also observe broadly similar behaviour for model H1 when the convergence criteria of the fitting algorithm is changed.

The fact that the significant changes in the fitted parameters observed are not matched by equivalent increases in the log-likelihood suggests that the likelihood functions of models M and H1 are approximately flat in certain dimensions. This fact, which was

[7]For these model fittings, we implemented in `MATLAB` the iterative fitting algorithm described in Appendix 6.A using a computer with an Intel Core i5-3320m processor running at 2.60 GHz under Windows 7 Home Premium Edition (64 bits) with 8 GB of RAM.

[8]Similar results are reported by Currie (2014, Figure 3 and Table 9) using USA data, who finds that the deviance of Model M hardly changes after thousands of iterations but the coefficients show no sign of convergence.

TABLE 6.3: Poisson log-likelihood and computing time in seconds for models M and H1 for different tolerance levels using the one-stage approach.

	Model M		Model H1	
Tolerance	Log-Likelihood	Computing time (s)	Log-Likelihood	Computing time (s)
10^{-3}	-22,006	22	-22,410	13
10^{-4}	-21,991	115	-22,400	67
10^{-5}	-21,986	437	-22,396	234
10^{-6}	-21,984	1,393	-22,395	643
10^{-7}	-21,984	2,377	-22,395	1,228

previously noted in Cairns et al. (2009, 2011), Haberman and Renshaw (2011) and Currie (2014), is explored further in Section 6.6.

6.5.3 Robustness to changes of the period range

The range of years in the data used to fit mortality models is often a subjective choice. In particular, as Booth et al. (2002) point out, the model user often has a choice between only looking at more recent data or using a longer period of data extending further back in time, which might be less relevant to current mortality but provides better estimates of the time series for the κ_t. We would therefore like the parameters estimated by the model to be robust to this choice.

Figure 6.9 shows the parameters fitted for model M when the start date of the data is increased by five-year increments (i.e., the data starts in 1961, 1966, 1971, 1976 and 1981). As can be seen, the model is mostly stable except for the last (and smallest) data range, where a very different pattern for α_x and ι_y is observed. It is interesting to note that these changes do not cancel each other out directly: the cohort parameters for the most recent years of birth change most, whilst the biggest changes in α_x occur at high ages which these cohorts have yet to reach.

Figure 6.10 shows the parameters fitted for the same data for model H1. Here we see a similar phenomenon, with the cohort parameters for the most recent years of birth changing significantly when only data since 1981 is used. However, for model H1, it is the values of $\beta_x^{(1)}$ which change rather than α_x in model M.

We have also tested the stability of the models with respect to changing the end year of the data, and obtain broadly consistent results.

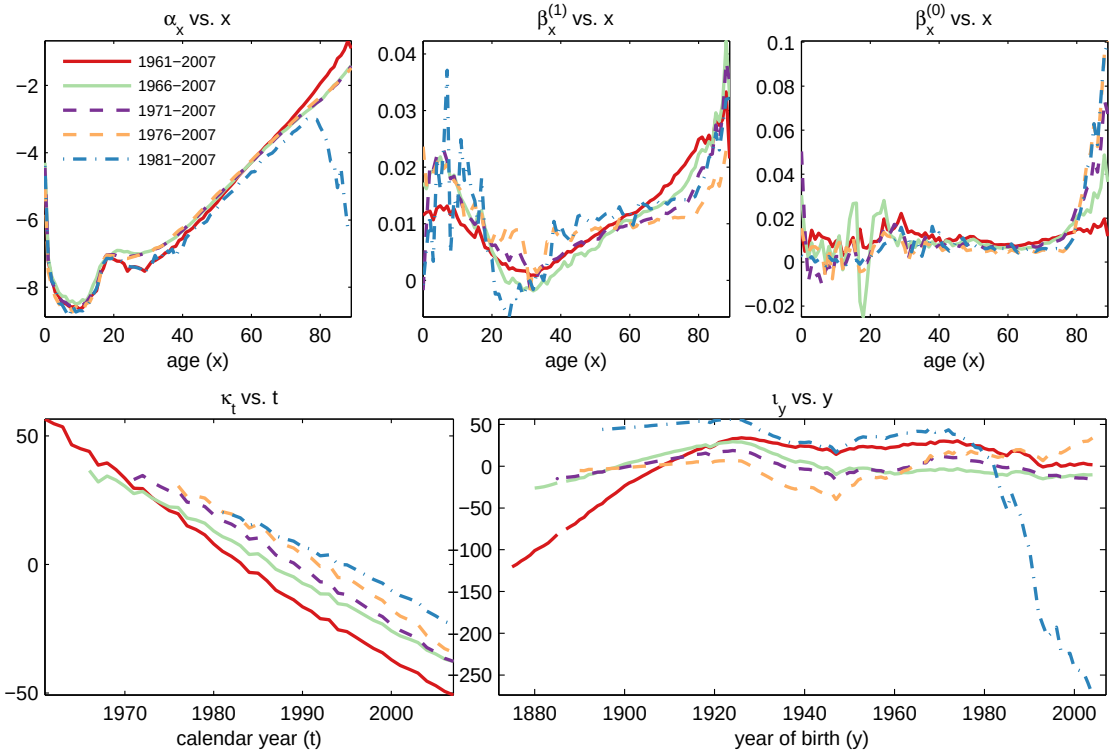


FIGURE 6.9: Parameters for model M for different start years fitted using the one-stage approach.

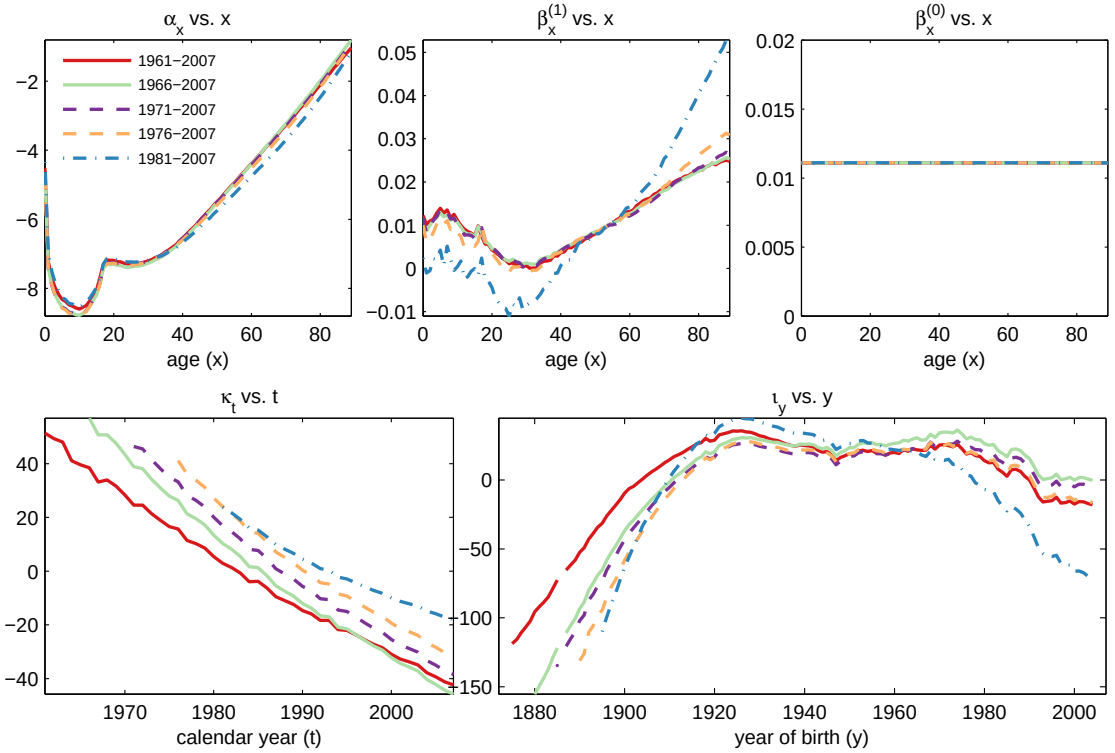


FIGURE 6.10: Parameters for model H1 for different start years fitted using the one-stage approach.

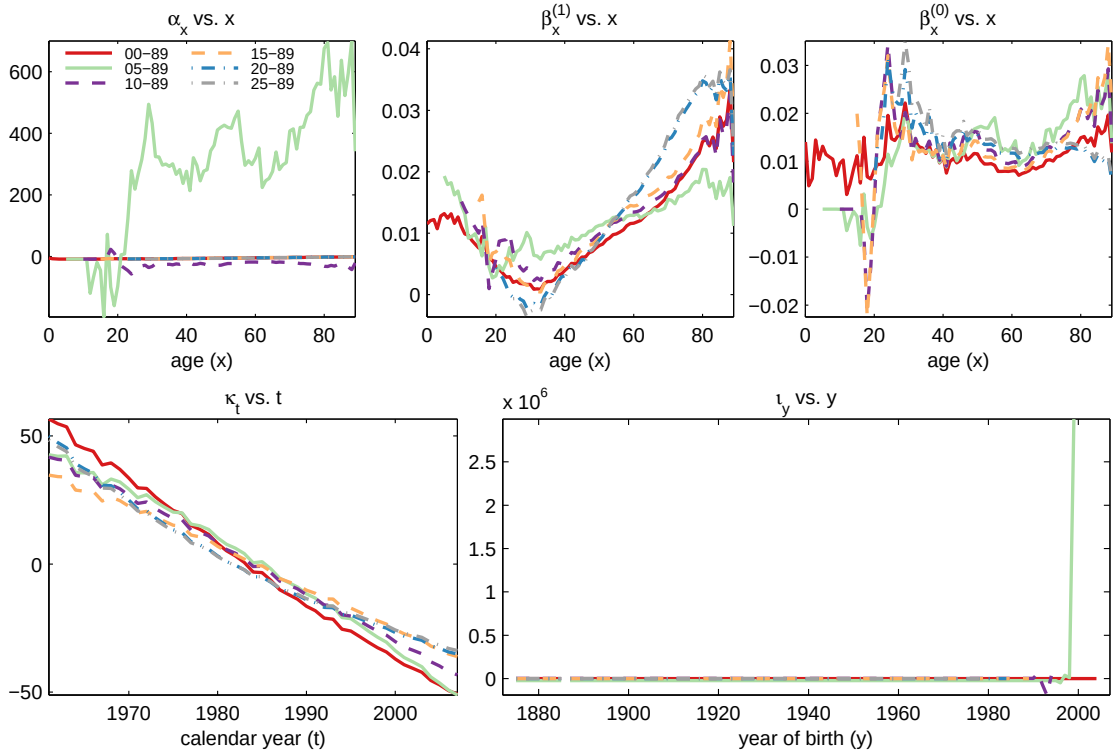


FIGURE 6.11: Parameters for model M for different start ages fitted using the one-stage approach.

6.5.4 Robustness to changes of the age range

The choice of age range to fit the models to is also somewhat subjective, as we may prefer to omit mortality data at very young ages due to the very different causes of death operating for these individuals. Therefore, we also test for the models' robustness when the age range of the data is increased in five-year increments from age zero (i.e., the data starts at ages 0, 5, 10, 15, 20 and 25).

Figure 6.11 shows the estimated parameters from model M fitted to different age ranges. The model failed to converge for data starting at age 5, which dramatically illustrates the lack of robustness of the model to different age ranges. The parameters estimated by the model also show significant dependence to the age range of the data included, with the pattern of the age functions $\beta_x^{(1)}$ and $\beta_x^{(0)}$ changing for different age ranges.[9]

Figure 6.12 shows the parameters estimated for model H1 using the same procedure. Here we observe that the model is considerably more stable, with the fitted parameters changing very little as the age range of the data is curtailed.

[9]Some of this variation however is attributable to the scaling of the parameters changing in order to maintain the constraints $\sum_x \beta_x^{(1)} = 1$ and $\sum_x \beta_x^{(0)} = 1$.

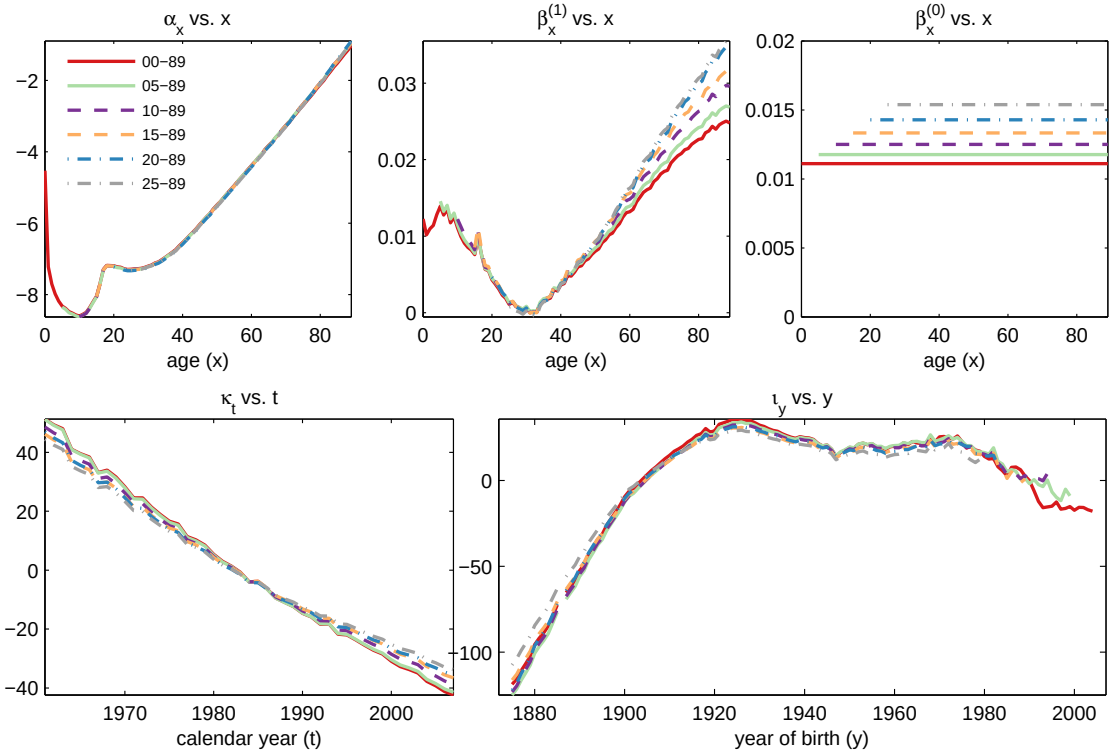


FIGURE 6.12: Parameters for model H1 for different start ages fitted using the one-stage approach.

6.5.5 Summary

In summary, we observe many of the phenomena reported in the previous studies by Cairns et al. (2009, 2011), Dowd et al. (2010) and Currie (2014) when using the one-stage fitting procedure and agree with their findings. Models M and H1 show considerable sensitivity to both the convergence criteria and the initial parameter estimates used by the algorithm used to fit the data, which is highly undesirable in a model. In addition, model M shows considerable lack of robustness to changes in the range of the data, although this is mitigated significantly by simplifying the model by fixing $\beta_x^{(0)} = 1/k$ and using model H1.

6.6 An approximate identifiability constraint

In Sections 6.5.1 and 6.5.2, we saw that the one-stage fitting approach suffers from a number of convergence and robustness issues. These meant that models M and H1 could show large changes in the patterns of the fitted parameters for relatively small changes in the fit to data. Typically, these parameter changes were of the form:

- a tilting of the cohort parameters, ι_y, around the midpoint of the range;
- a tilting of the period parameters, κ_t, around the midpoint of the range;
- some tilting of the static age function, α_x, with it becoming considerably less smooth as the tilt in ι_y increases in model M;
- a lack of robustness in the age function, $\beta_x^{(1)}$, with it picking up some of the features of $\beta_x^{(0)}$ in some cases (for instance, Figures 6.6 and 6.8); and
- the age function, $\beta_x^{(0)}$, remaining largely unchanged.

The issues with the robustness of the models to changes in the data ranges also possess many of these features. However, these changes do not entirely explain the failure of the models to converge in certain circumstances (for example, in Figure 6.11) or the abrupt changes in the form of the fitted parameters in others.

In Cairns et al. (2009), the authors comment that the issues with the robustness of model M may be due to an approximate identification issue which means that *"the likelihood function will be close to flat in certain dimensions"* making estimation slow, giving a number of different parameter values which could be selected by the algorithm and making the ones chosen dependent on the initial conditions specified. Such an identification issue would be caused by a transformation of the parameters which left the fit to data approximately unchanged. We agree with the Cairns et al. (2009) conjecture and, based on the parameter changes noted above and in Sections 6.5.1 and 6.5.2 and on the exact identifiability issue present in the classic APC model, have found the identification issue in question.

6.6.1 An approximate transformation and its impact

We start by rewriting κ_t as

$$\kappa_t = K(t - \bar{t}) + \delta_t \tag{6.18}$$

where $\bar{t} = 0.5(t_1 + t_n)$. $K(t - \bar{t})$ represents the linear trend in the period parameters, whilst δ_t represent the deviations of κ_t from this trend. Equation (6.18) is consistent with the constraint in Equation (6.14) if $\sum_t \delta_t = 0$, which would be the case if the value

of K is found by linear regression. For most datasets, κ_t is approximately linear.[10] In these cases, we would expect δ_t to be small relative to κ_t: however, it is important to note that κ_t can always be written in the form of Equation (6.18). Using this, model M has the form

$$\log \mu_{xt} = \alpha_x + \beta_x^{(1)}(K(t-\bar{t}) + \delta_t) + \beta_x^{(0)}\iota_{t-x}$$

This revised form of the model then permits transformations of the form

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\beta}_x^{(0)}, \tilde{\iota}_y\} = \Big\{\alpha_x + e\beta_x^{(0)}(x-\bar{x}), \frac{K}{K-e}\beta_x^{(1)} - \frac{e}{K-e}\beta_x^{(0)}, \frac{K-e}{K}\kappa_t, \beta_x^{(0)}, \iota_y + e(y-\bar{y})\Big\} \quad (6.19)$$

with similar definitions for $\bar{x}$ and $\bar{y}$ as for $\bar{t}$ above. This transformation preserves the other constraints in the model, i.e., $\sum_{x=x_1}^{x_k} \tilde{\beta}_x^{(1)} = 1$, $\sum_{t=t_1}^{t_n} \tilde{\kappa}_t = 0$, $\sum_{x=x_1}^{x_k} \tilde{\beta}_x^{(0)} = 1$, and $\sum_{y=t_1-x_k}^{t_n-x_1} \tilde{\iota}_y = 0$ by construction.

We see that this transformation affects the parameters in the manner described at the start of Section 6.6. Specifically, variations of the constant e will result in:

- a tilt of ι_y around its midpoint;
- a change in the gradient (i.e., tilt) of κ_t;
- a tilt of α_x around the midpoint and a change in its smoothness, depending upon the smoothness of $\beta_x^{(0)}$;
- the revised $\beta_x^{(1)}$ being a weighted average of the old $\beta_x^{(1)}$ and $\beta_x^{(0)}$; and
- no change in $\beta_x^{(0)}$.

However, the transformation in Equation (6.19) is approximate rather than exact, i.e. the fitted mortality rates in Equation (6.5) are only approximately the same. This is because the transformed fitted mortality rates, $\tilde{\mu}_{xt}$ are related to the original mortality rates fitted by model M by

$$\log \tilde{\mu}_{xt} = \log \mu_{xt} - \epsilon_{xt}$$

[10] See, for instance, Tuljapurkar et al. (2000) who go so far as to call this the *"universal pattern of mortality decline"* in the context of the LC model.

where $\epsilon_{xt} = \frac{e}{K}\beta_x^{(0)}\delta_t$.

In order to quantify what impact this has on our fit to data, we can Taylor expand the form of the Poisson log-likelihood shown in Equation (6.22) in Appendix 6.A for the death counts using the transformed mortality rates, $\tilde{d}_{xt}$, around the original fitted death counts, $\hat{d}_{xt}$, to give

$$\mathcal{L}(d_{xt}, \tilde{d}_{xt}) = \mathcal{L}(d_{xt}, \hat{d}_{xt}) - \sum_x \sum_t \omega_{xt} \left(d_{xt} - \hat{d}_{xt}\right) \epsilon_{xt} - \frac{1}{2} \sum_x \sum_t \omega_{xt} \hat{d}_{xt} \epsilon_{xt}^2 + o(\epsilon^{*2}) \tag{6.20}$$

where $\epsilon^* = \max_{xt} |\epsilon_{xt}|$.[11] The use of maximum likelihood techniques to fit the model ensures that $d_{xt} \approx \hat{d}_{xt}$ (i.e., the expected death counts are close to those observed) and consequently that locally, the likelihood is approximately unchanged to second order. This effectively defines a dimension where the log-likelihood is close to flat. However, the impact of this on the fitted mortality rates and the log-likelihood, will depend upon the magnitudes of the ϵ_{xt}. In particular, we find that the transformation will leave the fitted mortality rates approximately unchanged if either of the following two cases hold:

(a) e is small, i.e., we are not tilting the cohort parameters significantly from their original values; or

(b) δ_t is small, i.e., κ_t is approximately linear.

This theoretical result is born out in the results of the tests performed in Sections 6.6.3 and 6.7.

Since we can always make e smaller, Case (a) (e is small) will occur independently of the linearity of κ_t. Hence, around any set of parameters there will always be a small neighbourhood defined by varying e where the log-likelihood is approximately flat. This explains the behaviour of the fitted parameters and the slow convergence of the iterative maximum likelihood estimation algorithm observed in Section 6.5.2. As the convergence criterion was made smaller (and the number of fitting iteration increased), the algorithm could move slowly along the parameter space defined by small variations of e. This resulted in the significant changes in the parameter estimates shown in Figure 6.8

[11] If the model is fitted using a binomial setting with the logit link function, as in Haberman and Renshaw (2011), a similar expansion for the log-likelihood can be obtained.

which were not matched by equivalent increases in the log-likelihood (recall Table 6.3). Furthermore, this case explains the findings of Currie (2014) where fitting algorithms with stopping criteria based on the log-likelihood converge, while those with stopping criteria based on changes in the parameters (e.g., the score vector) sometimes fail to converge. This also means that the fitting algorithm performs numerous additional iterations, which substantially increase the time taken to converge but do not result in significant improvements in the fit to data. This explains the long fitting times shown in Table 6.3 and the slow convergence reported by Currie (2014, Figure 3 and Table 9) when fitting model M to data for the US population.

In contrast, the validity of Case (b) (δ_t is small), will depend on how well κ_t can be approximated by a straight line. For example, we find that for the fitted values of κ_t shown in Figure 6.6 using different initial values, the regression of κ_t on $t - \bar{t}$ in order to find K in Equation (6.18) have R^2 values of over 98%. This indicates that δ_t is comparatively small for this dataset and, hence, that the fitted mortality rates will be close to invariant under the transformation in Equation (6.19). For some datasets, however, κ_t will not be approximately linear and, hence, the transformation in Equation (6.19) can result in substantial changes in the fitted mortality rates and the log-likelihood as ι_y is tilted. An example of this is found in Section 6.7 using data for the USA.

Consequently, we see that the transformation in Equation (6.19) can always be applied approximately. In fact, it will give identical fitted mortality rates for the alternative model

$$\log \mu_{xt} = \alpha_x + \beta_x^{(1)} K(t - \bar{t}) + \beta_x^{(0)} \iota_{t-x}$$

This model is similar to the models considered in Renshaw and Haberman (2003a) and Callot et al. (2014), albeit with a cohort term.

Furthermore, the transformation in Equation (6.19) does not depend on the form of $\beta_x^{(0)}$. In particular, making the simplifying assumption that $\beta_x^{(0)} = 1/k$, as is done in model H1, does not change the form of the transformation. Therefore, for model H1 we have a similar approximate transformation of the form

$$\{\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \tilde{\kappa}_t, \tilde{\iota}_y\} = \left\{\alpha_x + \frac{e}{k}(x - \bar{x}), \frac{K}{K-e}\beta_x^{(1)} - \frac{e}{k(K-e)}, \frac{K-e}{K}\kappa_t, \iota_y + e(y - \bar{y})\right\}$$

The form of this transformation, however, explains why the smoothness of α_x does not change for model H1 when it tilts, unlike for model M where a multiple of $\beta_x^{(0)}$ is added to it.

6.6.2 An additional constraint

Having identified the transformation in Equation (6.19), we can use this to impose an additional constraint on the parameters of the model. The analysis above shows that, if κ_t is well approximated by a straight line, this constraint will have minimal impact on the fit to data obtained by the model. In our implementation, we choose to impose the additional constraint

$$\sum_{y=t_1-x_k}^{t_n-x_1} (y - \bar{y})\iota_y = 0, \tag{6.21}$$

which is the natural extension of the constraint on the level of ι_y in Equation (6.16) and ensures that the cohort function does not possess a linear trend over the range of the data. This imposes the demographic significance in Section 6 of Hunt and Blake (2015) that the cohort parameters should be close to trendless and that systematic changes in time should be controlled by the period functions. In addition, this constraint coupled with constraints (6.4) and (6.16) will allow us to interpret α_x as the average level of mortality across the period of the data. However, the choice of this constraint is arbitrary and depends upon our subjective interpretation of the meaning of the parameters, and other identifiability constraints could be imposed with equal justification.[12] The steps required for incorporating the constraint in Equation (6.21) in the fitting of model M are described in Appendix 6.A.

6.6.3 Testing the models with the additional constraint

As applying the transformation in Equation (6.19) will change the static age function, α_x, it can only be used when the models are fitted using the one-stage procedure discussed in Section 6.5. Since the additional constraint restrains the freedom of model the parameters to vary in the dimension where the likelihood function is approximately

[12] An alternative constraint could be to impose $\min_x \beta_x^{(1)} = 0$, which would ensure that mortality rates across all ages are positively correlated with the period function κ_t. This may be desirable for some purposes, as discussed in Haberman and Renshaw (2009).

flat, we expect that models M and H1 will be considerably more robust, more stable and faster to converge than observed previously without the additional constraint. Moreover, since κ_t is linear to close approximation in the figures in Section 6.5, the analysis above indicates that the transformation in Equation (6.19) will leave the fitted mortality rates approximately unchanged and adding the constraint in Equation (6.21) will have a minimal impact on the model log-likelihood. To test this, we apply the same tests to the models with the approximate identifiability constraint as were performed in Section 6.5.

With the additional constraint, both models M and H1 show almost no sensitivity to the initial parameter estimates used to fit them using the same procedure as in Section 6.5.1. This is understandable, given the similarity between the approximate transformation in Equation (6.19) for model M and that in Equation (6.17) in the APC model.

We also find that models M and H1 are considerably faster to converge and less sensitive to changes in the convergence criteria. For instance, Figure 6.13 shows the fitted parameters from model M with the approximate constraint for different convergence criteria in the same manner as performed in Section 6.5.2. We can see that there is no appreciable difference between the parameter sets, and only very slight improvements in the fitted log-likelihood for smaller convergence criteria. Model H1 shows the same lack of sensitivity to the convergence criteria. Comparing Tables 6.3 and 6.4, we also note that, as predicted by the Taylor expansion in Equation (6.20), the additional identifiability constraint has only marginally worsened the goodness of fit of the models. By contrast, the convergence of the model has been dramatically speeded up, which is desirable for many purposes such as testing for parameter uncertainty using the bootstrapping techniques proposed in Brouhns et al. (2005) and Koissi et al. (2006) which involve repeatedly refitting the model. For instance, while without the additional identifiability constraint model H1 took about four minutes to converge for a tolerance level of 10^{-5}, with the additional constraint this time was reduced to 13 seconds.

Nevertheless, introducing the approximate identifiability constraint does not appear to solve the lack of robustness shown by model M to changes in the range of the data used to fit the model. For example, Figure 6.14 shows the estimated parameters for model M with the approximate identifiability constraint when the start year of the data is

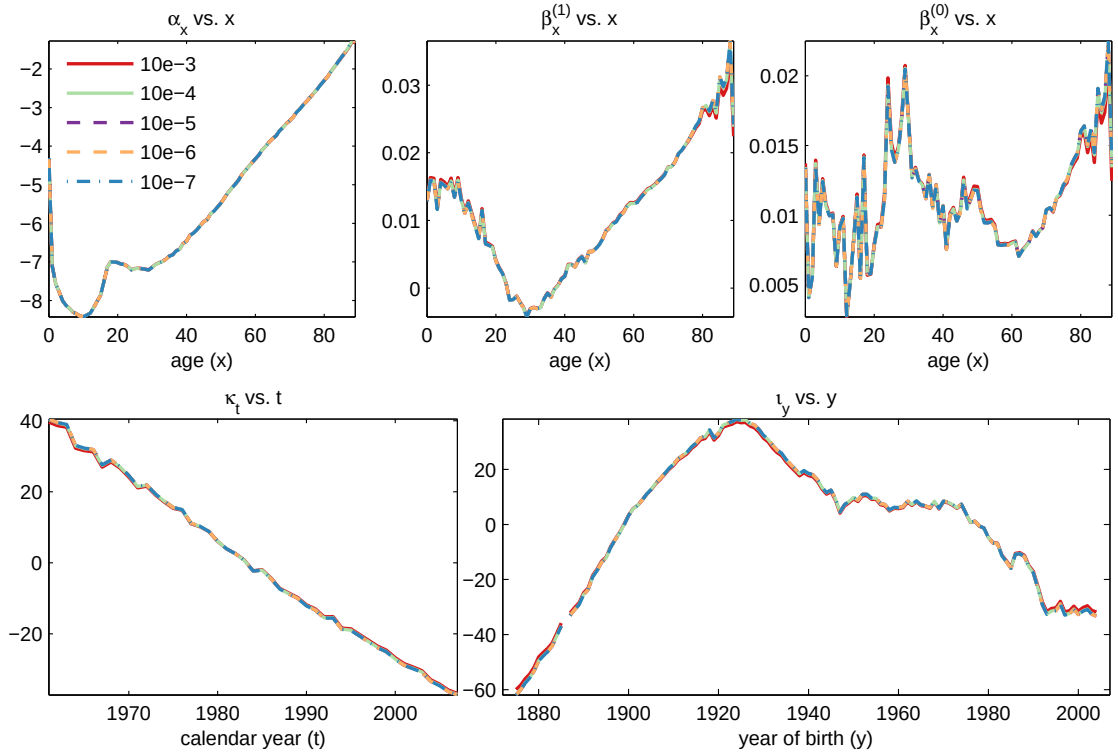


FIGURE 6.13: Parameters for model M for different tolerance levels fitted using the one-stage approach with the approximate identifiability constraint.

TABLE 6.4: Poisson log-likelihood and computing time in seconds for models M and H1 for different tolerance levels obtained using the one-stage approach with the approximate identifiability constraint.

	Model M		Model H1	
Tolerance	Log-Likelihood	Computing time (s)	Log-Likelihood	Computing time (s)
10^{-3}	-22,007	10	-22,408	5
10^{-4}	-22,005	20	-22,407	8
10^{-5}	-22,005	31	-22,407	13
10^{-6}	-22,005	43	-22,407	17
10^{-7}	-22,005	53	-22,407	21

changed in the same fashion as in Section 6.5.3. This shows very similar instability as before, shown in Figure 6.9.

In a similar fashion, we find that using the approximate identifiability constraint does not change the lack of robustness of model H1 to using only data from 1981 onwards to fit the model, although the fitted parameters are still robust to the other changes as found in Sections 6.5.3 and 6.5.4.

This is not entirely surprising, as the parameter changes exhibited in Sections 6.5.3 and 6.5.4 were not entirely consistent with the form of the transformation in Equation (6.19),

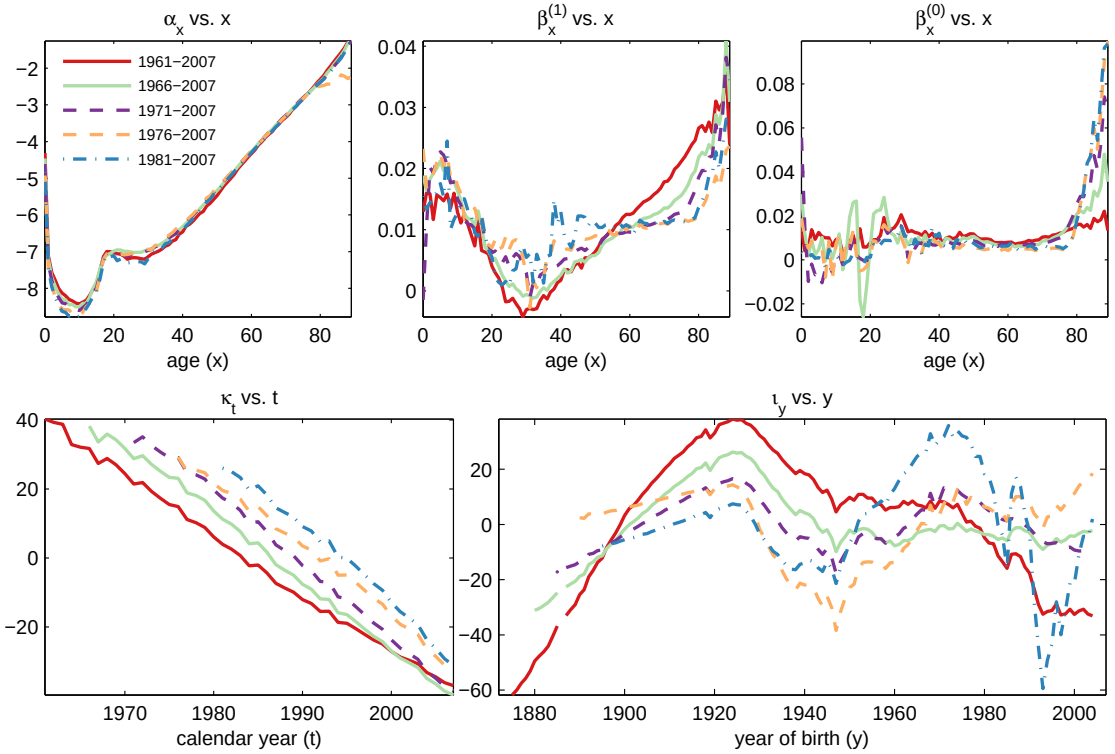


FIGURE 6.14: Parameters for model M for different start years fitted using the one-stage approach with the approximate identifiability constraint.

unlike those in Sections 6.5.1 and 6.5.2. This indicates that it may not be possible to assign structure in the data to an age/period or age/cohort interaction uniquely in these models. Alternatively, it may be that the assumption that the shape of the age functions $\beta_x^{(1)}$ and $\beta_x^{(0)}$ is not constant, as assumed, over the full range of the data but changes as the underlying drivers of improving mortality change. This phenomenon has been observed for the LC model by Carter and Prskawetz (2001), Booth et al. (2002) and others. However, it is noticeable that model H1 is considerably more robust to changes in the data ranges due to the simplification of the age/cohort term.

6.7 Other datasets

The results presented so far in this study are based on the analysis of data for men in England and Wales for ages 0 to 89. This has been consistent with the primary data used in other similar studies such as Renshaw and Haberman (2006), Haberman and Renshaw (2009, 2011) and Cairns et al. (2009, 2011). As a robustness check on the

reasonableness of our results, we have performed a similar analysis for a number of other datasets,[13] including

- female data for England and Wales for ages 0 to 89,
- male and female data for England and Wales for ages 60 to 89,
- male and female data for the USA for ages 0 to 89,
- male data for the Netherlands for ages 0 to 89, and
- male data for Spain for ages 0 to 89.

These alternative datasets have been chosen because they are either direct complements to the primary dataset used in this study (i.e., the alternative England and Wales datasets), or because they have been used in other studies which have encountered robustness and stability issues with the models (data for the USA in Cairns et al. (2009) and Currie (2014), the Netherlands in van Berkum et al. (2014) and Spain in Debón et al. (2010)). Furthermore, van Berkum et al. (2014) shows that there is considerable non-linearity to the first period function for mortality models fitted to data for the Netherlands, which makes this a useful test of the approximate transformation described in Section 6.6.

For these alternative datasets we find broadly similar results as those found for male data for England and Wales in the rest of this study. In particular, we note that, for all the datasets used, applying the additional identifiability constraint helps resolve the lack of robustness the models have to the initial parameter estimates and the convergence criteria of the fitting algorithm. Table 6.5 shows the impact on the goodness of fit of applying the additional identifiability constraint for different datasets, which is relatively small in most cases. In particular, we note that in the case of using model M for women aged 0-89, the imposition of the approximate constraint actually increases the goodness of fit of the model, which we believe is due to the greater efficiency of the fitting algorithm. We thus discuss only briefly the main highlights of these results, but further details are available on request from the authors.

For the England and Wales female data for the full age range, we find very similar results to those presented for men, i.e., that the additional identifiability constraint significantly

[13] All data for the period 1961 to 2007.

TABLE 6.5: Poisson log-likelihood for models M and H1 without and with the additional identifiability constraint for different datasets.

Dataset			Log-likelihood			
			Model M		Model H1	
Country	Sex	Age Range	Without Constraint	With Constraint	Without Constraint	With Constraint
E&W	Male	0-89	-21,992	-22,005	-22,400	-22,407
E&W	Male	60-89	-8,847	-8,847	-8,999	-9,003
E&W	Female	0-89	-20,499	-20,497	-20,797	-20,799
E&W	Female	60-89	-8,448	-8,448	-8,642	-8,649
USA	Male	0-89	-38,096	-39,510	-41,467	-43,706
USA	Female	0-89	-28,691	-29,085	-30,323	-30,489
NL	Male	0-89	-17,514	-17,522	-17,698	-17,721
Spain	Male	0-89	-23,370	-23,410	-25,517	-26,530

improves the robustness of the model to different choices of the initial parameter values and the tolerance level used by the fitting algorithm, but model M still lacks robustness to changes in the age and period range of the data. However, as seen in the men dataset, most of these robustness issues are resolved by using model H1 with the same notable exception of the lack of robustness when using only data from 1981 onwards (recall Figure 6.10).

The results for the more limited range of data in England and Wales are in general consistent with the results obtained for the full age range. However, we highlight that for both men and women, model H1 with the additional identifiability constraint is now robust to changes in the period range of the data, including even when using data for the period 1981 to 2007. It therefore seems likely that these issues are related to the behaviour of mortality rates at low ages.

For the USA data, we note the following features.

- The values of κ_t fitted for model M are less linear (R^2 values for the regression in Equation (6.18) of approximately 95%) and, hence, the using transformation in Equation (6.19) results in larger changes in the fitted mortality rates and log-likelihoods compared with the England and Wales data.

- The fitted parameters for model M without the additional identifiability constraint for both men and women lack demographic significance, in the sense that it is difficult to associate any sensible interpretation to the parameters. In particular,

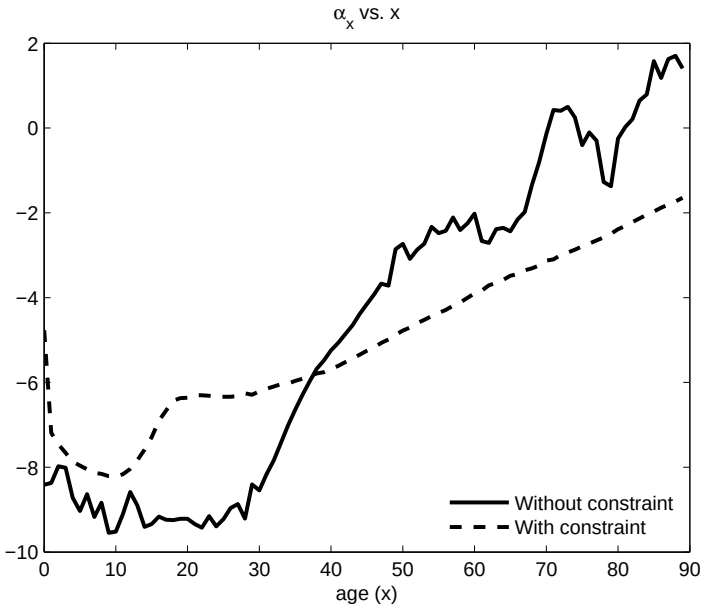


FIGURE 6.15: α_x parameters for model M for for USA male data for ages 0-89 and period 1961-2007 fitted using the one-stage approach without and with the approximate identifiability constraint.

we find that α_x lacks smoothness for both sexes (as shown in Figure 6.15 for men) and does not show the familiar pattern of mortality across ages. This is especially important, since the desire for demographically significant α_x was cited as a large motivation of the two-stage fitting approach in Haberman and Renshaw (2009). However, these issues are resolved by using the approximate identifiability constraint. These results may be in part attributed to the weaker cohort structure of USA mortality, as suggested by Tuljapurkar (2013), which means that the model has difficulty uniquely estimating the cohort parameters.

- Similar to the England and Wales data, model M exhibits a lack of robustness to changes in the range of the data being used even with the additional identifiability constraint. However, using model H1 in preference substantially improves this.

We find that, for the Dutch and Spanish datasets, the models perform have very similar issues to England and Wales. However, we do not find the same non-linear behaviour for κ_t for models M and H1 for the Netherlands that was found in van Berkum et al. (2014), which means that the approximate invariant transformation gives smaller reductions in the goodness of fit than were expected.

In summary, we find that the other datasets tested show many of the same issues with robustness and stability when using models M and H1 as the England and Wales male data. As with this dataset, using the additional identifiability constraint partially resolves these issues, especially in conjunction with using the simplified model H1, which makes the model substantially more stable to changes in the data range. These results are, therefore, strongly indicative that the issues and solutions are general features of models M and H1 and are not specific just to the data used in Section 6.5

6.8 Conclusions

In this chapter, we have investigated the robustness of models M and H1 using both the two-stage fitting approach pioneered in Renshaw and Haberman (2006) and the one-stage approach in Cairns et al. (2009). We broadly agree with the findings of both of these studies and agree with the conclusions of Haberman and Renshaw (2011) that the difference in robustness found are explained by the different fitting procedures.

However, we find in Section 6.4 that the two-stage implementation procedure proposed in Renshaw and Haberman (2006) and used subsequently in Haberman and Renshaw (2009, 2011) introduces an additional constraint which is not theoretically justified and worsens the fit to data considerably. Using alternative constraints, which are equally or sometimes more justifiable, gives parameter estimates which are substantially different. This is not desirable and should be avoided. However, simply omitting the additional constraint whilst retaining the two-stage fitting procedure in Section 6.4.2 gives a model which is highly unstable, even more than when using the one-stage procedure. Coupled with the fact that fixing the static age function, α_x, a priori worsens the fit to data in all eventualities, we therefore recommend the use of a one-stage procedure when fitting models M or H1 to data.

Despite being more justifiable theoretically, the results of Section 6.5 show that the one-stage procedure has considerable convergence and robustness problems. These make the results of the model very sensitive to both the specific details of the fitting algorithm used (such as the initial parameter estimates and the tolerance levels for the algorithm to converge) and the data the model is fitted to. This is clearly undesirable. The maximum likelihood algorithms used to fit the models also take a relatively long time to converge to

the optimum solution, which limits the practicality of the models in many circumstances. However, we find that model H1, which simplifies the age/cohort interaction in model M, is both more robust to changes in the data and faster to converge, which may make it more suitable for many purposes.

In Cairns et al. (2009), it was conjectured that many of the robustness problems with model M (and, by extension, H1) were due to an additional, approximate identifiability issue present in the model. In Section 6.6, we find the transformation of the parameters in the model which generates this approximate identifiability model and determine its impact on the fit to data. This transformation is then used to impose an additional constraint on the cohort function, ι_y, in the model. In Section 6.6.3, we then show that including this additional constraint resolves many of the robustness issues discussed in Section 6.5 using the one-stage procedure.

However, the model is still unstable to changes in the data even with the additional constraint. This may be due to it being difficult to allocate structure in the data between the age/period and age/cohort interactions, or due to the shape of the age functions $\beta_x^{(1)}$ and $\beta_x^{(0)}$ changing over the period of the data, as discussed in Carter and Prskawetz (2001) and Booth et al. (2002) in the context of the LC model. If so, this may explain why model H1 is generally more robust to changes in the data range than model M.

We have also found broadly similar results for other populations in Section 6.7, indicating that our findings are general features of the models rather than specific to the data for men in England and Wales. We also find that model H1 with the additional constraint is more appropriate for populations with a weak cohort effect and a period index with bigger deviations from linearity, such as in the USA.

In conclusion, we recommend using a one-stage procedure with an additional constraint of the form discussed in Section 6.6.2 to fit model H1. This has the advantage of the greater theoretical justification and improved fit to data of the one-stage procedure of Cairns et al. (2009) whilst obtaining the faster convergence speeds, robustness and interpretability of the parameters of the two-stage procedure of Renshaw and Haberman (2006). The use of model H1 as opposed to model M also improves the robustness of the model with respect to changes in the underlying data and gives a simpler subjective interpretation of the demographic significance of the age/cohort interaction. We therefore

believe that this approach combines the best of all worlds and helps resolve many of the outstanding issues regarding extensions of the LC model incorporating cohort effects.

Acknowledgements

The authors would like to thank Ana Debón, Steven Haberman and the anonymous referee for their comments which have helped improved this manuscript.

Appendix 6.A One-stage fitting procedure

To fit model M we consider that death counts are independent Poisson responses $D_{xt} \sim Poisson(e_{xt}\mu_{xt})$ and derive parameter estimates by maximising the log-likelihood

$$\mathcal{L}(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \left\{ d_{xt} \log \hat{d}_{xt} - \hat{d}_{xt} - \log d_{xt}! \right\} \tag{6.22}$$

where $\hat{d}_{xt} = e_{xt} \exp\left(\alpha_x + \beta_x^{(1)}\kappa_t + \beta_x^{(0)}\iota_{t-x}\right)$ and ω_{xt} are 0-1 weights indicating empty or omitted data cells.

In order to maximise the log-likelihood function we employ a straightforward extension of the Newton-Raphson iterative procedure used by Brouhns et al. (2002) in the estimation of the Lee-Carter model. In iteration $h+1$, a typical parameter θ is updated using the updating relationship

$$\hat{\theta}^{\{h+1\}} = \hat{\theta}^{\{h\}} - \frac{\partial \mathcal{L}/\partial\theta}{\partial^2 \mathcal{L}/\partial\theta^2}$$

Starting with appropriate initial parameters, we apply iteratively the updating relationships until the log-likelihood of the model converges to a desired tolerance. Table 6.6 shows the updating relationships for model M and Algorithm 1 presents the detailed parameter estimation procedure.

We note that if the additional approximate identifiability constraint of Section 6.6.2 is being used, then Step 7 in Algorithm 1 imposes constraint (6.21) by applying transformation (6.19) with particular values of constants K and e. Specifically, K is found by

TABLE 6.6: Updating relationships for the maximum likelihood estimation of the parameters of model M.

Parameter	Updating formula
α_x	$\hat{\alpha}_x^{\{h+1\}} = \hat{\alpha}_x^{\{h\}} + \dfrac{\sum_t \omega_{xt}(d_{xt} - \hat{d}_{xt})}{\sum_t \omega_{xt}\hat{d}_{xt}}$
$\beta_x^{(1)}$	$\hat{\beta}_x^{(1)\{h+1\}} = \hat{\beta}_x^{(1)\{h\}} + \dfrac{\sum_t \omega_{xt}\hat{\kappa}_t(d_{xt} - \hat{d}_{xt})}{\sum_t \omega_{xt}\hat{\kappa}_t^2\hat{d}_{xt}}$
κ_t	$\hat{\kappa}_t^{\{h+1\}} = \hat{\kappa}_t^{\{h\}} + \dfrac{\sum_x \omega_{xt}\beta_x^{(1)}(d_{xt} - \hat{d}_{xt})}{\sum_x \omega_{xt}\left(\beta_x^{(1)}\right)^2\hat{d}_{xt}}$
$\beta_x^{(0)}$	$\hat{\beta}_x^{(0)\{h+1\}} = \hat{\beta}_x^{(0)\{h\}} + \dfrac{\sum_t \omega_{xt}\hat{\iota}_{t-x}(d_{xt} - \hat{d}_{xt})}{\sum_t \omega_{xt}\hat{\iota}_{t-x}^2\hat{d}_{xt}}$
ι_y	$\hat{\iota}_y^{\{h+1\}} = \hat{\iota}_t^{\{h\}} + \dfrac{\sum_{\substack{x,t \\ y=t-x}} \omega_{xt}\beta_x^{(0)}(d_{xt} - \hat{d}_{xt})}{\sum_{\substack{x,t \\ y=t-x}} \omega_{xt}\left(\beta_x^{(0)}\right)^2\hat{d}_{xt}}$

Algorithm 1 Maximum likelihood parameter estimation procedure for model M

1: **set** initial values $\hat{\alpha}_x$, $\hat{\beta}_x^{(1)}$, $\hat{\kappa}_t$, $\hat{\beta}_x^{(0)}$, $\hat{\iota}_y$; **compute** $\hat{d}_{xt}$; and **compute** $\mathcal{L}(d_{xt}, \hat{d}_{xt})$
2: **update** $\hat{\alpha}_x$ and **compute** $\hat{d}_{xt}$
3: **update** $\hat{\kappa}_t$; **set** $c := \frac{1}{n}\sum_t \hat{\kappa}_t$, $\hat{\kappa}_t := \hat{\kappa}_t - c$, $\hat{\alpha}_x := \hat{\alpha}_x + c\hat{\beta}_x^{(1)}$ to ensure $\sum_t \hat{\kappa}_t = 0$; and **compute** $\hat{d}_{xt}$
4: **update** $\hat{\beta}_x^{(1)}$; **set** $a := 1\big/\sum_x \hat{\beta}_x^{(1)}$, $\hat{\beta}_x^{(1)} := a\hat{\beta}_x^{(1)}$, $\hat{\kappa}_t := \frac{1}{a}\hat{\kappa}_t$ to ensure $\sum_x \hat{\beta}_x^{(1)} = 1$; and **compute** $\hat{d}_{xt}$
5: **update** $\hat{\iota}_y$; **set** $d := \frac{1}{n+k-1}\sum_y \hat{\iota}_y$, $\hat{\iota}_y := \hat{\iota}_y - d$, $\hat{\alpha}_x := \hat{\alpha}_x + d\hat{\beta}_x^{(0)}$ to ensure $\sum_y \hat{\iota}_y = 0$; and **compute** $\hat{d}_{xt}$
6: **update** $\hat{\beta}_x^{(0)}$; **set** $b := 1\big/\sum_x \hat{\beta}_x^{(0)}$, $\hat{\beta}_x^{(0)} := b\hat{\beta}_x^{(0)}$, $\hat{\iota}_y := \frac{1}{b}\hat{\iota}_y$ to ensure $\sum_x \hat{\beta}_x^{(0)} = 1$; and **compute** $\hat{d}_{xt}$
7: **if** constraint in Equation (6.21) is imposed using Equation (6.19); **set** $K := \sum_t (t-\bar{t})\hat{\kappa}_t \big/ \sum_t (t-\bar{t})^2$, $e := -\sum_y (y-\bar{y})\hat{\iota}_y \big/ \sum_y (y-\bar{y})^2$, $\hat{\alpha}_x := \hat{\alpha}_x + e\hat{\beta}_x^{(0)}(x-\bar{x})$, $\hat{\beta}_x^{(1)} := \frac{K}{K-e}\hat{\beta}_x^{(1)} - \frac{e}{K-e}\hat{\beta}_x^{(0)}$, $\hat{\kappa}_t := \frac{K-e}{K}\hat{\kappa}_t$, $\hat{\iota}_y := \hat{\iota}_y + e(y-\bar{y})$ to ensure $\sum_y (y-\bar{y})\hat{\iota}_y = 0$; and **compute** $\hat{d}_{xt}$
8: **compute** $\mathcal{L}(d_{xt}, \hat{d}_{xt})$
9: **if** $\mathcal{L}(d_{xt}, \hat{d}_{xt})$ has not converged **go to** 2
10: **return** $\hat{\alpha}_x$, $\hat{\beta}_x^{(1)}$, $\hat{\kappa}_t$, $\hat{\beta}_x^{(0)}$, $\hat{\iota}_y$

regressing κ_t on $(t-\bar{t})$, so that $\kappa_t = K(t-\bar{t}) + \delta_t$, $\delta_t \sim N(0, \sigma_\kappa^2)$ i.i.d.; and e is found by regressing ι_y on $-(y-\bar{y})$, so that $\iota_y = -e(y-\bar{y}) + \xi_y$, $\xi_y \sim N(0, \sigma_\iota^2)$ i.i.d.

Bibliography

Booth, H., Maindonald, J., Smith, L., 2002. Applying Lee-Carter under conditions of variable mortality decline. Population Studies 56 (3), 325–36.

Brouhns, N., Denuit, M., Van Keilegom, I., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. Scandinavian Actuarial Journal (3), 212–224.

Brouhns, N., Denuit, M., Vermunt, J., 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics 31 (3), 373–393.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2011. Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics 48 (3), 355–367.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1–35.

Callot, L., Hales, S., Lamb, M., 2014. Deterministic and stochastic trends in the Lee-Carter mortality model. Tech. rep., Aarhus University.

Carter, L. R., Prskawetz, A., 2001. Examining structural shifts in mortality using the Lee-Carter method. Tech. rep., Max Planck Institute for Demographic Research.

Clayton, D., Schifflers, E., 1987. Models for temporal variation in cancer rates. II: Age-period-cohort models. Statistics in Medicine 6 (4), 469–81.

Continuous Mortality Investigation, 2002. Working Paper 1 – An interim basis for adjusting the “92" series mortality projections for cohort effects.
URL http://www.actuaries.org.uk/research-and-resources/pages/cmi-working-paper-1

Continuous Mortality Investigation, 2007. Working Paper 25 – Stochastic projection methodologies: Lee–Carter model features, example results and implications.

Currie, I. D., 2006. Smoothing and forecasting mortality rates with P-splines.
URL http://www.macs.hw.ac.uk/~iain/research/talks/Mortality.pdf

Currie, I. D., 2014. On fitting generalized linear and non-linear models of mortality. Scandinavian Actuarial Journal.

Debón, A., Martínez-Ruiz, F., Montes, F., 2010. A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. Insurance: Mathematics and Economics 47 (3), 327–336.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2010. Backtesting stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts. North American Actuarial Journal 14 (3), 281–298.

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

Haberman, S., Renshaw, A., 2011. A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics 48 (1), 35–55.

Hobcraft, J., Menken, J., Preston, S., 1982. Age, period, and cohort effects in demography: a review. Population Index 48 (1), 4–43.

Human Mortality Database, 2014. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL www.mortality.org

Hunt, A., Blake, D., 2015. On the structure and classification of mortality models. Working paper.

Koissi, M., Shapiro, A., Hognas, G., 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. Insurance: Mathematics and Economics 38 (1), 1–20.

Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87 (419), 659–671.

Lovász, E., 2011. Analysis of Finnish and Swedish mortality data with stochastic mortality models. European Actuarial Journal 1 (2), 259–289.

Murphy, M. J., 2009. The 'golden generations' in historical context. British Actuarial Journal S1 (15), 151–184.
URL http://www.actuaries.org.uk/research-and-resources/pages/baj-current-issue

Murphy, M. J., 2010. Reexamining the dominance of birth cohort effects on mortality. Population and Development Review 36 (2), 365–390.

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

Renshaw, A., Haberman, S., 2003b. On the forecasting of mortality reduction factors. Insurance: Mathematics and Economics 32 (3), 379–401.

Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556–570.

Tuljapurkar, S., 2013. How and why mortality change varies between sexes and countries. Tech. rep., Stanford University, California.
URL http://www.cias-cufe.org/Longevity9/downloads/Plenary_Session_I_TuljaPresentationFINAL.pdf

Tuljapurkar, S., Li, N., Boe, C., 2000. A universal pattern of mortality decline in the G7 countries. Nature 405 (6788), 789–792.

van Berkum, F., Antonio, K., Vellekoop, M., 2014. The impact of multiple structural changes on mortality predictions. Scandinavian Actuarial Journal.

Willets, R., 1999. Mortality in the next millennium. Staple Inn Actuarial Society.

Willets, R., 2004. The cohort effect: Insights and explanations. British Actuarial Journal 10 (4), 833–877.

Wilmoth, J. R., 1990. Variation in vital rates by age, period, and cohort. Sociological Methodology 20, 295–335.

7

General Conclusions

The central topic of this thesis has been the understating of the trends and financial consequences of socio-economic differences in mortality. We have developed modelling techniques for the quantification of socio-economic mortality differentials in aggregate and cause-specific mortality, which we have applied in the study of the relationship between mortality and deprivation in the English population. Furthermore, stemming from the use of stochastic mortality models in the investigation of mortality differentials, we have also made contributions to the assessment of longevity basis risk and to the general use of stochastic mortality models.

In Chapter 2 we have developed an extension of the Lee-Carter model especially designed to take into consideration the characteristics of socio-economic subpopulations as well as some of the typical issues of mortality data disaggregated by socio-economic circumstances. This newly introduced model allows a clear separation of mortality level differentials and mortality improvement differentials among socio-economic subpopulations, and produces subpopulation-specific mortality forecasts which are consistent with mortality forecasts for the national population. Using this model, we have found that people living in more deprived areas of England not only have significantly lower life expectancies than those living in less deprived areas, but have also experienced slower mortality improvement rates.

In order to shed light on the drivers behind socio-economic mortality differentials, in Chapter 3 we have investigated how mortality trends for the leading causes of death differ between socio-economic groups. To do this, we have introduced an extension of

the Lee-Carter model to allow for the consideration of coding changes in cause-specific mortality data, which we then have embedded into a multiple population setting to enable the quantification of socio-economic differences in cause-specific mortality. The application of this approach to English data has revealed that for the leading causes of death people living in more deprived areas have significantly higher mortality rates than those living in less deprived areas, with circulatory, digestive, and respiratory diseases showing the strongest socio-economic gradient. Furthermore, with the exception of neoplasms, we have found that for the leading causes of death there is a widening of the relative mortality gap between more and less deprived areas of England.

The persistency and general widening of mortality inequalities has implications for the managing of longevity risk in pension funds and annuity portfolios. In particular, the basis risk emerging from the mis-match between the socio-economic characteristics of annuitants and those of the national population is one of the key obstacles preventing many pension schemes and insurers from considering index based longevity transactions as a legitimate alternative for the management of their longevity risk. In order to contribute to the development of the market of standardised longevity transactions, in Chapter 4 we have systematically evaluated the suitability of several multi-population stochastic mortality models for assessing basis risks and provided guidelines on how to use these models in practical situations. Our investigations show that two-populations models can offer market participants an appropriate approach to assess the amount of demographic basis risk involved in an index-based longevity hedge. Moreover, our results suggests that standardised hedges have the potential to provide an effective and flexible solution to mitigate longevity risk.

As a by-product of our research on socio-economic differences and basis risk assessment, we have also developed new modelling tools that allow a more effective and widespread use of stochastic mortality models. Specifically, in Chapter 5 we have introduced a new **R** package for modelling and forecasting mortality. This package provides a general framework for implementing many of the stochastic mortality models proposed in the demographic and actuarial literature and includes tools for fitting, analysing goodness of fit, evaluating parameter uncertainty, and performing mortality projections. We believe that this new **R** package is a valuable addition to the actuarial toolkit for measuring and managing longevity risk.

Finally, in Chapter 6 we have systematically investigated the well-known robustness and stability issues of cohort-extensions of the Lee-Carter model. As a result of this investigation, we have proposed a new fitting approach which solves many of the prevalent problems with this type of models.

There are a number of areas where the work of this thesis can be extended. The socio-economic differences in life expectancy and mortality that we have evidenced in Chapters 2 and 3 are unacceptable in a fair society. Therefore, with the aim of providing policy advice for reducing mortality inequalities, it is worth working in developing methods for identifying which causes of death are most influential in shaping socio-economic difference in mortality.

The forecasts of cause-specific mortality rates carried in Chapter 3 were made using independent univariate ARIMA models for each cause, implicitly assuming independence among the the different causes of death. This is clearly unrealistic as, for instance, the same risk factors can affect several causes at the same time. It is thus worth exploring the use of multivariate time series models such as Vector Autoregression (VAR) and Vector Error Correction Models (VECMs) so to explicitly allow in the forecasts for the possible dependence among the different causes of deaths.

In accordance with the standard practice in the multiple population mortality literature, the investigations on the assessment of basis risk carried out in Chapter 4 relied on the assumption that in the long-run the index reference portfolio and the pension fund/annuity book being hedged share the same mortality improvements. This so-called non-divergence assumption was captured through the use of autoregressive time series to model the mortality spread between the reference and book populations. However, in Chapters 2 and 3 we saw that over the past three decades the most affluent areas of England have consistently experienced faster mortality improvements than less affluent areas, suggesting that the non-divergence assumption might not be appropriate in all cases. Therefore, we believe that there is a need for further research considering alternative times series models beyond the autoregressive model and evaluating the implication for basis risk assessment of a departure from the non-divergence assumption.

In Chapter 4 we found that the accurate calibration of two-population models for basis risk assessment requires that the pension fund/annuity book has an annual exposure of over 20,000-25,000 lives and at least 10-12 years of reliable data. However, in practice

a large proportion of pension scheme books and life company portfolios will not meet these data requirements. Thus, there is a need for research looking at the assessment of longevity basis risk when the available mortality data is scarce.

Finally, the market of longevity risk transfers would benefit from computational tools which make two-population mortality models easily accessible to practitioners and facilitate the assessment of longevity basis risk. We thus encourage the development of a multiple population extension of the **R** package **StMoMo** introduced in Chapter 5.