Instituto Superior de Estatística e Gestão da informação

Universidade Nova de Lisboa

Lisboa, Portugal

Modelling and Forecasting Mortality Patterns

A thesis submitted

in partial fulfilment of the requirements

for the degree of

Doutor em Estatística e Gestão de Informação

by

Edviges Isabel Felizardo Coelho

Lisboa, October 2012

Instituto Superior de Estatística e Gestão da informação

Universidade Nova de Lisboa

Lisboa, Portugal

Modelling and Forecasting Mortality Patterns

A thesis submitted

in partial fulfilment of the requirements

for the degree of

Doutor em Estatística e Gestão de Informação

by

Edviges Isabel Felizardo Coelho

Advisor: Professor Luis Catela Nunes

Lisboa, October 2012

ACKNOWLEDGMENTS

First and foremost I would like to express my gratitude to Professor Luis Catela Nunes. I am grateful for the guidance and advice, as well as the time and effort to answer my questions and to diligently comment on my drafts. The work presented in this thesis would not be possible without his continuous support and encouragement.

Also, I would like to take this opportunity to thanks to those who helped me in one way or another on conducting this research. In particular, I would like to thank Professor Jorge Bravo for teaching me how to implement the Poisson-Lee-Carter method and for providing his SAS code for Poisson estimation, and to Professor David Harvey for providing his GAUSS code for structural change testing. I am particular grateful to Ana Patricia Martins and Daniel Fernandes for all the help with R programming, and for always being available to answer my questions.

I thank my colleagues for the support.

I owe a special thanks to my friends for all the patience and support.

Finally, a very special thanks to my parents, whose love and encouragement supported and motivated me throughout this work.

Edviges Isabel Felizardo Coelho

October, 2012

ABSTRACT

This work considers the problem of modelling and forecasting mortality.

In the context of the Lee-Carter model, where the overall temporal pattern of mortality is well described by a time-varying index moderated by age-specific effects, the study proposes a sequential testing procedure valid in the presence of a structural change in the index of mortality, which allows the identification of the appropriate model used in forecasting future mortality patterns. Testing procedures proposed are applied to Portugal, and other eighteen developed countries, using data for the period 1950-2007. Structural changes in the overall time trend are found in Portugal, as well as in other developed countries, occurring mainly in male mortality. The impacts of a neglected structural change in the overall temporal behaviour of mortality are illustrated for the case of male Portuguese population.

In order to better understand male and female mortality patterns in Portugal, several descriptive measures and visualization techniques are used and various extensions of the Lee-Carter model are applied. Changing rhythms of decline over ages and time are found. The possibility of a cohort effect in male mortality is also suggested.

Finally, cohort specific influences in male Portuguese mortality patterns are studied using log-linear additive age-period-cohort models, as well as models allowing for age interaction with period and cohort effects. The evidence is favourable to the existence of cohort effects influencing male population patterns, with some generations experiencing stabilizing or even poor mortality conditions than preceding and subsequent cohorts.

Keywords: Age-period-cohort model, cohort effects, Lee-Carter model, mortality forecasting, structural change, unit root.

RESUMO

Este trabalho considera o problema da modelização e previsão da mortalidade.

No contexto do modelo de Lee-Carter, em que a tendência temporal global da mortalidade é bem descrita por um índice que varia ao longo do tempo, moderado por efeitos específicos por idade, o estudo propõe um procedimento sequencial de teste que é válido na presença de uma alteração estrutural no índice de mortalidade, permitindo a identificação do modelo mais apropriado para previsão do comportamento futuro da mortalidade. O procedimento de teste proposto é aplicado a Portugal, e a mais dezoito países desenvolvidos, utilizando dados para o período 1950-2007. Alterações estruturais na tendência de declínio da mortalidade são encontradas em Portugal e em quase todos os outros países considerados no estudo, ocorrendo principalmente na mortalidade masculina. Os impactos de negligenciar uma alteração estrutural na tendência da mortalidade são ilustrados para o caso da população masculina portuguesa.

A fim de compreender melhor os padrões de mortalidade de homens e mulheres em Portugal, são usadas várias medidas descritivas e técnicas de visualização, bem como, desenvolvimentos recentes ao modelo de Lee-Carter. São evidentes alterações nas taxas de declínio da mortalidade ao longo do tempo e por idades, bem como, uma forte possibilidade de um efeito geracional na mortalidade masculina.

Finalmente, é estudada a influência de efeitos de coorte no comportamento da mortalidade masculina em Portugal através da utilização de modelos log-lineares idade-período-coorte, considerando também interações idade-período e idade-coorte. A evidência empírica é favorável à existência de efeitos de geração que influenciam o padrão de mortalidade masculina, em que algumas coortes experimentam uma estabilização ou mesmo deterioração das condições de mortalidade comparativamente a coortes anteriores e subsequentes.

Palavras-chave: Alteração estrutural, efeitos de coorte, modelos idade-período-coorte; modelo de Lee-Carter, previsão da mortalidade, raiz unitária.

CONTENTS

1. INTRODUCTION

Mortality in Portugal, following the trend in other developed countries, has changed tremendously over the twentieth and first decade of the twenty-first centuries. During this period, with the exception of the influenza epidemic crisis in 1918, one saw a general decline in the level of mortality, a dramatic reduction in infant mortality, the increase of survival at more advanced ages, and extraordinary gains in the life expectancy of the population. Over the last 60 years, life expectancy at birth in Portugal increased more than 20 years.

Although positive for the individual as well as for the society as a whole, in the actual social and demographic context, these achievements and the perspective of further longevity gains pose a challenge to public sector health care provision, to public retirement system and to private life annuities business, among many others. As such, the analysis and monitoring of the mortality patterns, as well as modelling and forecasting future mortality patterns are of fundamental importance.

The major problem for age-specific modelling and forecasting mortality is the high dimensionality of the data, especially if single ages and calendar years are used. In order to deal with the dimensionality problem, it is necessary to use models allowing a more parsimonious representation of the data.

The Lee-Carter methodology (Lee and Carter, 1992) constitutes a landmark in the evolution of mortality stochastic modelling. Several methodological developments have been proposed to improve the accuracy of the model, among which the choice of an optimal fitting period (Booth, Maindonald and Smith, 2002), adding an age-specific enhancement factor (Renshaw and Haberman, 2003), a log-bilinear Poisson specification (Brounhs, Denuit and Vermunt, 2002), and the introduction of an age-cohort term in the log-bilinear model (Renshaw and Haberman, 2006). These methodological developments motivated the present work, either in what concerns the research on a possible structural change in the temporal mortality pattern as in exploring other specific patterns in overall mortality in Portugal, in our knowledge not yet, such as cohort effects. The present work was also inspired on the classical log-linear additive

age-period-cohort model, largely used in health and epidemiological studies, as well as on recent research concerning the intrinsic identification problem of this type of models.

The purpose of this thesis is to further explore mortality patterns in Portugal, over the period 1950 – 2007, investigate the possibility of a structural break in the overall temporal trend of mortality, identify different rhythms of decline of age-specific rates, detect and model the eventual presence of cohort effects in mortality patterns. These contributions are important in establishing more accurate models on which to base future forecasts.

The thesis consists of three papers organized in chronological order. The first paper addresses the structural change in time-varying mortality index and proposes a sequential testing procedure to test for structural change and unit-roots prior to the identification step in the Box-Jenkins time-series methodology. The second paper, using descriptive measures and visualization techniques, explores the mortality patterns of male and female populations. Different empirical models are considered for the whole 1950 – 2007 period and for several sub-periods, with the objective of obtaining consistency with the main mortality patterns that were found. The third paper considers age-period-cohort models, exploring different model specifications, as well as measures to evaluate the goodness of fit of the various models, and concludes for the evidence of cohort influences in male mortality.

The main contributions of this work are as follows. We expect to demonstrate how the Lee-Carter model might be improved by the developments in the theory of unit roots and structural changes in order to provide better mortality and longevity forecasts. Also, we expect to contribute to a more accurate knowledge of mortality patterns in Portugal, specially the influence of cohort effects, moderated by age, which should be considered in mortality modelling and forecasting.

A more detailed description of each paper follows.

1.1. Forecasting Mortality in the Event of a Structural Change

The paper considers the problem of forecasting future mortality patterns in the event of a structural change in the context of the Lee-Carter model (1992), a pioneering contribution for stochastic mortality forecasting methods and that, as well as some of its extensions, is widely used by demographers and actuaries to model and forecast mortality. A core assumption of the Lee-Carter model is that the evolution of age-specific death rates is driven by a common time-varying index, which represents the improvements over time of the general level of mortality. This mortality index is then forecast to obtain future mortality patterns.

Several recent studies have documented changes in the pace of decline of the general level of mortality (Booth, et al., 2002; Renshaw and Haberman, 2003; Booth, Hydman, Tickle, and Jong, 2006; Shang, Hydman and Booth, 2010), providing evidence of changing trends in the mortality declining rate. The eventual presence of a structural break in the general level of mortality poses a methodological problem to modelling the mortality index, since the standard procedures to check the stationarity of a time series are not valid and may induce the researcher to select a sub-optimal model for the index of mortality and consequently obtain biased forecasts. The paper demonstrates how recent advances in statistical testing for structural changes can be used to arrive at a properly specified model for the general level of mortality. Specifically, the results of tests for a change in the trend of the mortality index and for the presence of a unit-root are used to identify appropriate forecasting models. The testing methodology proposed also allows, in the case of the presence of a structural change, to identify the date when the change occurred. The paper also shows the importance of considering structural changes when making predictions, illustrating, for the case of Portuguese male mortality, the impact in mortality and life expectancy forecasts of accounting for such structural changes.

It concludes that there is significant evidence supporting the presence of a structural change in the evolution of male mortality, which is associated with a more accentuated decline in the overall mortality rate in recent years for almost every country considered. In contrast, evidence of structural changes in female mortality is found only for a few countries. It is interesting to note that in the case of the European countries

these structural changes have all taken place in the second half of the sample, that is, after the mid-1970s. These findings are in line with changes documented in other studies.

The main contribution is the definition of a sequential testing procedure, which is valid in the presence of a structural change, and, if a change is present, allows determining the date when it occurred, contributing to the identification of the appropriate model used to forecasting.

A succinct summary of each section of the paper follows.

After a brief introduction, which states the problem and explains the contribution intended by the researchers, the paper introduces the classical Lee-Carter method, proposed by Ronald Lee and Lawrence Carter in 1992, and the Poisson regression extension proposed by Brouhns, Denuit and Vermunt (2002), that is followed in the empirical application.

Lee and Carter (1992) propose a demographic model for describing the temporal evolution of age-specific death rates as a function of a single time-varying index. The model is estimated and the resulting time-varying parameter is then forecast using Box-Jenkins time-series methods. Forecasts for the age-specific death rates are obtained using the estimated effects and the forecasts of the time-varying index. To estimate the demographic model for a given matrix of death rates by age and calendar year, Lee and Carter (1992) propose a least squares solution using a singular value decomposition, assuming errors with zero mean and constant variance. Following several critics on the homoskedastic nature of the errors, Brounhs et al. (2002) propose an alternative model that keeps the functional form of the Lee-Carter model but replaces the least-squares approach with a Poisson regression for the number of deaths, and a maximum likelihood interactive estimation process.

The main steps of the Box-Jenkins methodology used to model the mortality index time-series are also briefly reviewed. In particular, the first step in obtaining an appropriate ARIMA model for the mortality index using the Box-Jenkins methodology. It consists of determining if some transformation of the series, such as differencing or de-trending, is necessary to induce stationarity before identifying and estimating the forecasting model. Different transformations chosen to induce stationarity generate different point forecasts and different probability intervals.

In the third section, the paper discusses the implications of the presence of a structural change in the trend of the mortality index, and a sequential testing approach to cope with this possibility when forecasting the mortality index, based on the papers of Harvey, Leybourne, and Taylor (2009) and Harris, Harvey, Leybourne, and Taylor (2009), is proposed.

The importance of testing for a structural change is that any neglected or wrongly placed structural change in a time-series occurring in the estimation phase may result in forecasts with a tendency to deviate from the future realizations of the series, resulting in potentially large forecast errors. However, the approach to test for and estimate a model with a structural change is highly dependent on the stationarity proprieties of the time-series process.

In the presence of a structural change in the trend, standard procedures to detect the presence of a unit-root and the need to first-difference the data, such as the analysis of the empirical autocorrelation functions and the standard Dickey-Fuller unit root tests, are not valid, having the tendency to wrongly suggest a unit-root when the series is, in fact, stationary around a broken trend. Modified unit-root tests valid in the presence of a structural change, when the change date is not known and must be estimated were proposed by Zivot and Andrews (2002) and Perron and Rodriguez (2003). However, when no change is actually present, several problems arise, such as a spurious change date or erroneously find a structural change when there is none due to the presence of a unit root.

A sequential testing procedure is proposed to deal with this dilemma. In a first step, since it is not possible to know beforehand if there is a unit-root or not in the mortality index, we test for the presence of a structural change in the trend. To do that we apply the test for a structural change in the trend proposed by Harvey, et al. (2009) that is valid regardless of whether the series is difference-stationary or trend-stationary. Secondly, as in Harris, et al. (2009), the result of this structural change test is used to decide on the appropriate unit-root test to use: with or without allowing for a structural change. Thirdly, if according to the Harvey et al. (2009) test (1) the hypothesis of a structural change is rejected, one should use standard unit root procedures (ADF-GLS test not allowing for a break); (2) the hypothesis of a structural change is not rejected, one should use the Harris et al. (2009) unit root test allowing for a break (ADF – GLS

test allowing for a break). Fourth, according to the results of the appropriate unit-root test, in the case where in the first step a structural change was detected, the date when the change occurred is estimate using first-differences if a unit root is found or using levels otherwise. At last, based on the conclusions of the previous steps, the appropriate ARIMA model is fitted to the mortality index and forecasts obtained.

In the fourth section, the paper illustrates the sequential testing methodology proposed by an application to post-1950 mortality data for a set of 18 developed countries: Austria, Belgium, Canada, Denmark, England and Wales, Finland, France, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Switzerland, Sweden, United States of America, and West Germany.

Fitting the Poisson Lee-Carter to male and female populations in each country, the estimated mortality indices confirm the downward trend in mortality over time observed in all countries, and their visual inspection suggest that the rate of decline after the mid-1970s, especially for male mortality, might have become more accentuated in a number of countries. Such visual inspection, however, may suggest a spurious break if a unit-root in present in the time-series.

The results of the first step of the sequential testing approach to identify the correct forecasting model in the potential presence of a structural change show significant evidence of a structural change in the trend slope of male mortality indices in all countries, except Finland and Spain. The situation is quite different for females, where the evidence is favourable to a structural change in the trend slope only in Austria, Ireland, Italy, Japan and Sweden. Evidence of a unit-root is found in all countries, except Denmark for males and Japan and Sweden for females, which implies that the mortality indices are stationarity in first-differences, that is, the model should be fitted to the first-differences of the mortality indices, except for males from Denmark and females from Japan and Sweden. The authors note that in the case of the European countries these structural changes have all taken place in the second half of the sample, that is, after the mid-1970s.

The authors consider the male Portuguese mortality case to illustrate how the results of the proposed sequential tests methodology can be used to arrive at an appropriate forecasting model, as well as the impact of the results of the unit-root and structural change tests obtained previously in terms of mortality and life expectancy

forecasts considering models estimated with different specifications of the structural change and unit-root.

According to the results of the sequential testing approach the mortality index for the Portuguese males presents a structural change in 1996 and first-differences should be applied to induce stationarity. The appropriate model is then an ARIMA (0,1,1) model allowing for a structural change in 1996. Using this model we obtain forecasts of the mortality index for the period 2008 – 2050 and the corresponding 95% confidence bands. The confidence bands widen over time as a direct consequence of the unit-root non stationary model.

The impact of allowing for a structural change in the mortality index is illustrated by comparison of the above model with three other models estimated with different specifications of the structural change and the unit-root: (1) ARIMA (0,1,0) with drift or random walk with drift model, (2) the ARIMA (1,1,0) not allowing for a break, and (3) the ARMA (1,0) allowing for a break in 1973.

The random walk with drift model, which is the model most often used to fit the mortality index and that was found to be optimal in the original work of Lee and Carter (1992), does not fit well to the male Portuguese mortality index. The ARIMA (1,1,0) is the model without a structural change that better fit the estimated mortality index. By construction using this model the forecasted rate of decline in mortality is basically given by the average rate of decline during the whole estimation period, therefore the more accentuated rate of decline in mortality since the end of 20th century until the end of the estimation sample is not translated into the future projections. Comparatively with the optimal model, the projected decline in mortality is much less using this model and its confidence bands are wider as a consequence of the poorer fit. The ARMA (1,0) model allowing for a break in 1973 corresponds to the trend-stationary model that better fit the mortality index. The forecasted decline in mortality is also not as great as in the optimal forecasting model, and the confidence bands, as expected, do not grow as the forecasting horizon grows, since this model assumes stationarity around a broken trend. Thus, for longer horizons, this model will produce forecast intervals with the least amplitude, but obviously underestimating future uncertainty, given the results of the sequential testing procedure that led to the optimal forecasting model.

For each of the estimated models, mortality index forecasts are used also to produce forecasts for the age-specific death rates and forecasts for the life expectancy at birth and at 65 years were calculated. The optimal model with a unit-root and a change in 1996 gives the highest projected gains in life expectancy during the next four decades. The confidence bands for the mortality forecasts and the amplitude of the confidence bands for the 2050 life expectancy forecasts are substantially smaller for the optimal model than for the unit-root models without allowing for a structural change, but larger when compared with the trend-stationary model with a break.

The paper concludes with a discussion of the results comparatively with several research studies that show evidence regarding changes in the rate of decline of the general level of mortality. The findings are in line with these results and the paper concludes that the testing procedures can be used to select the appropriate forecasting model in the context of Lee-Carter model. Future research paths point to allowing more than one latent mortality index in the model to capture dissimilar evolutions of mortality at different ages, as well as the possibility of considering more than one structural change.

1.2. Patterns of Mortality Decline in Portugal since 1950

This paper considers mortality improvement patterns in Portugal, over the period 1950 – 2007. The aim is to show how a detailed descriptive analysis of the patterns of mortality improvement can be used to benefit the formulation and validation of a parsimonious model for estimating mortality dynamics over age and time. To identify the major patterns in the evolution of mortality over time for different ages, detailed descriptive analysis and graphical visualizations of mortality data are explored, as well as the analysis of the fitting results from the Poisson Lee-Carter model applied to the whole 1950-2007 period and to different sub-periods, and results from an extended model adding a second order age-period term.

The paper was motivated by previous research findings (Coelho and Nunes, 2011), in particular the evidence of a clear pattern in deviance residuals from the Poisson Lee-Carter model applied to Portuguese male mortality, suggesting that there

are patterns in male mortality that have not been fully captured by the model. Deviance residuals plotted against age show larger residuals at infancy and around the accident hump ages, suggesting that there are age-period interactions not captured by the model. The plot of deviance residuals versus year of birth shows some sort of ripple effect, which suggest that not only period effects may differ between ages but also cohort effects may be influencing the mortality pattern.

The descriptive analysis of age-specific mortality rates, as well as composite longevity indicators such the life expectancy at birth, show that the rate of mortality decline at infant and younger ages has changed dramatically, and improvements at other ages have also had several distinct decline patterns over the period 1950 - 2007. In the case of male population, there are even periods of some deterioration in mortality conditions for some age groups. These differences in sensitivity of age-specific mortality to a temporal trend are not properly captured by the Lee-Carter model, and suggest that additional parameters should be added to the demographic model to allow for greater flexibility.

The findings concern the relevance of descriptive analysis for the specification of the model that most adequately captures mortality patterns. The Lee-Carter model with a second age-period interaction term is found to be the most appropriate model, however with less significant improvement for females than for males, and still, in the case of male mortality, some patterns are not fully captured. The possibility of a cohort effect is left for further research.

A succinct summary of each section of the paper follows.

After a brief introduction, the paper describes the improvement in observed mortality patterns over age and time, for males and females separately, based on calculation and graphical analysis of indicators of longevity, age-specific mortality indicators, and rates of mortality improvement over age groups and time intervals. The most relevant conclusions of the analysis are the remarkable increase in life expectancy at birth, that surpassed twenty years during the period 1950 – 2007; the huge decline in infant mortality; the decline of death rates at almost all ages but at different paces; lower or even deterioration of male mortality conditions for some age groups that move upward with age; recent contributions for increases life expectancy at birth are specially becoming from those aged 60 to 79 years; the concentration of deaths around the mode

of the distribution of deaths and the consequent rectangularization and expansion of the survival curve.

The third section of the paper briefly reviews the classical Lee-Carter method (Lee and Carter, 1992), and its Poisson extension proposed by Brouhns, et al. (2002), which assumes a Poisson distribution for the number of deaths and proposes a maximum likelihood interative estimation process. The latter is used in the empirical modelling. The addition of a second bilinear age-period term, which allows for a greater flexibility in capturing the mortality patterns, is also considered, following the work of Booth et al. (2002) and Renshaw and Haberman (2003). Model performance measures, based on residuals, are discussed, such as the deviance residuals and the pseudo R^2, proposed by Cameron and Windmeijer (1997) for Poisson regression models.

Section 4 presents and discusses the estimation results of the Poisson Lee-Carter model, applied for males and females separately, to the whole 1950 – 2007 period. In general the model captures the overall trends in mortality, the exceptions being those ages with a less regular behaviour where some difficulties of adjustment are evident, as in the case of men between 20 to 35 years old. Systematic pattern in the deviance residuals are visible, in particular for men, presenting a structure hardly consistent with the assumption of independence of the residuals, showing a diagonal clustering of residuals which suggest the presence of cohort effects.

The results of fitting the Poisson Lee-Carter model to four subsequent sub-periods, to address the problem of age-specific changing patterns, are presented in section 5. The results confirm that the rates of mortality improvement at each age have changed over time.

In section 6, the paper presents the results of the Lee-Carter model with a second age-period interaction term applied to the entire period 1950 – 2007. Overall, adding a second age-period interaction term improves the fit of the model, in particular for men by capturing the nonlinearity in the log-rates in particular at young ages. The impact of the second bilinear term for women is less significant. The analysis of the residuals for males, however, continues to show some clustering patterns.

The paper concludes with the discussion of the consistency of the results of the descriptive analysis and the estimated results from each of the above models. The relevancy of descriptive analysis for the specification of the model that most adequately

captures mortality patterns is acknowledged. The possibility of a cohort effect in the male mortality is left for further research.

1.3. Cohort Effects in Mortality Modelling

This paper studies the possibility of a cohort effect in male Portuguese mortality, using age-period-cohort models, considering additive period and cohort effects as well as age-period and age-cohort interactions. The purpose of age-period-cohort models is to estimate the respective contributions of the effects of age, period and cohort on the evolution of the overall mortality. All three types of effects are distinct, but operate simultaneously, given the relation among them, which poses the problem of identification at the core of age-period-cohort modelling.

It is useful to define the three types of effects in age-period-cohort analysis. Age represents the intrinsic mortality risks related to biological and physiological factors, which are associated to the aging process, and which affect equally all individuals of the same age, regardless of their generational experience and observation period. Period effects are environmental conditions common to a given time period that change the level of mortality risk at all ages. Cohort effects or generational effects represent the influence of past conditions on current mortality risks; these are attributable to the life course experience of a cohort.

The study was motivated by a previous exploratory analysis of the patterns of mortality decline in Portugal over the period 1950 – 2007 (Coelho and Nunes, 2011), which showed that the rhythm at which mortality has been improving differs between ages and over periods of time. In particular, an apparently odd pattern of mortality improvement was detected in the male population, suggesting that over time a specific group of male individuals might have experienced a non-declining or even increasing mortality. The deviance residuals, from fitting the Lee-Carter model to male Portuguese mortality, also suggest that particular age-time influences are not fully captured by the model. Specifically the plot of the deviance residuals versus year of birth show a ripple effect, a pattern similar to that found in England and Wale due to cohort effects in mortality. Furthermore, to our knowledge, all studies of the mortality evolution in

Portugal have focused on the impacts of age and period, and relatively little is known about possible cohort effects.

The paper, using the Bayesian Information Criterion (BIC), concludes that an age, period and cohort model with period and cohort effects moderated by age - the Renshaw-Haberman model - is the most suitable model to describe Portuguese male mortality patterns from 1950 to 2007. These findings are consistent with previous descriptive findings. The paper also concludes about the validity of BIC in the selection of the most parsimonious and adequate model through a simulation procedure.

A concise summary of each section of the paper follows.

Following a brief introduction, the paper presents the theoretical framework of the Age-Period-Cohort (APC) modelling, based on data arrayed by age and period, with equal interval widths for age and period, which identify cohorts by the diagonal cells of the table.

First, it contemplates the additive APC model, the identification problem of the APC model, the constrained-based approach proposed by K. Mason, W. Mason, Winsborough and Poole (1973) to estimate and solve the identification problem, as well as more recent developments, in particular the approach of Kuang, Nilsen and Nilsen (2008). The introduction of age-period and age-cohort interactions in the model is presented next. This is illustrated by the Lee-Carter model (Lee and Carter, 1992; Brouhns, et al., 2002), the Renshaw-Haberman model (Renshaw and Haberman, 2006), and the Lee-Carter model with additive cohort effects.

Age-Period-Cohort analysis was initially based on the visual inspection of death rates plotted by year of birth. Researchers, such as Derrick in 1927 (see Hobcraft et al., 1982), Kermack, McKendrick and McKinlay in 1934 (see Murphy, 2010; Hobcraft et al., 1982), and Frost in 1939 (Case, 1956), demonstrated graphically that age-specific mortality rates seem to be more regular for cohorts than for periods, and as such more useful to project future trends. A formal approach came several years later. Although the statement of the APC model is straightforward, its estimation poses a statistical problem, known as the identification problem, due to the linear relationship between age, period and cohort. For the case of mortality, the APC model states that some monotonic transformation of the observed age-specific mortality rates can be modelled as a linear additive function of age, period and cohort effects.

The seminal work of K. Mason, W. Mason, Winsborough and Poole (1973), address the identification problem in the age-period-cohort model and, using a multiple classification framework, propose a constrained-based approach, based on *a priori* and theoretical knowledge, to estimate the age, period and cohort model parameters. Several other approaches to solve the identification problem of the APC model have followed, such as estimable functions of the parameters invariant to the particular constrain applied, among which second differences of age, period and cohort effects (Clayton and Schifflers, 1987). Using second order differences of the three effects as measures of curvature, Kuang et al. (2008) propose an alternative approach to the identification problem in the age-period-cohort model, which consists in transforming the design matrix so that curvatures (the second differences) of age, period and cohort are estimated directly along with the two linear components. Still the original parameters are obtained imposing identification restrictions on the parameters.

Either explicitly or not, attempts to solve the identifiability problem, and obtain estimates of the absolute values of the age, period and cohort effects, are ultimately based on some type of constraint. According to Clayton and Schifflers (1987) and Carstensen (2007) it is not possible to uniquely estimate the absolute effects of age, period and cohort, unless the researcher is willing to make some assumptions on the relative importance of the effects, since it is not achievable to surmount the problem of having two variables as well as their sum in the same linear model.

The introduction of age-period and/or age-cohort interactions may be a way of solving the identification problem, but also carry specific identification requirements that need to be attended. The Lee-Carter model (Lee and Carter, 1992) includes an age-period interaction, but no cohort effect. The Renshaw-Haberman model (Renshaw and Haberman, 2006) generalizes the Lee-Carter model adding an additional age-cohort interaction. The authors also generalize the Lee-Carter model to include an additive cohort effect.

The data and a visual analysis of the age-specific mortality rates for Portuguese males are presented in section 3 of the paper. Data for the number of deaths and the corresponding exposure to risk, by single ages and years of calendar, for the period 1950 – 2007, are obtained from the Human Mortality Database. The evidence from the

graphical analysis is that besides aging changing effects, period and cohort effects may both be present in the data.

In section 4 the estimation results of the following models are presented: the age-cohort (AC) model, the age-period (AP) model, the age-period-cohort (APC) model, the Lee-Carter (LC) model, the Lee-Carter with additive cohort effects (LCC), and the Renshaw-Haberman (RH) model. The fitting quality of each model is checked through the graphical analysis of the deviance residuals plotted against age, calendar year and year of birth, and by the values of deviance and BIC.

The Renshaw-Haberman model, with period and cohort effects both moderated by age, is found to be the most suitable model to describe Portuguese male mortality patterns from 1950 to 2007. This result is consistent with our expectations from the exploratory analysis of male mortality pattern in Portugal over 1950 – 2007, showing different rates of improvement for different ages, and the fact that some birth cohorts showed particular poor mortality improvements over adulthood when compared to other generations. The evidence of the graphical analysis, as well as the results from the estimated models, allude to the existence of cohort effects influencing Portuguese male mortality data.

In section 5 the paper proceeds with a simulation study. The simulation allows us to determine if BIC is able to select the appropriate model for different data generating processes. Through the simulation study it is also possible to understand what happens when the fitted model does not correspond to the true model, that is, what are the consequences of an incorrect specification of the model. Simulation settings are chosen such that they mimic real world situations, in particular Portugal and England and Wales. England and Wales have well documented cohort effects (Willets, 1999, 2003). The simulation concludes that BIC is in fact choosing the most adequate model to the data, since it selects the model corresponding to the data generating process. In what concerns errors in the specification it is found that if only additive cohort effects are present in the data, allowing for age interaction terms will result in an over parameterized model. On the other hand, if the cohort effect is omitted and a model with an age-period interaction term is fitted to such data, the estimated age-period interaction terms will capture part of the cohort effect in the data. If only age-period interaction terms and additive cohort effects are present in the data, allowing for a cohort effect and

the respective age interaction terms will result in an over parameterized model. Conversely, the omission of an interaction term will result in spurious cohort effects.

References

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under Conditions of Variable Mortality Decline. *Population Studies*, 56, 325– 336.

Booth, H, Hyndman, R.J., Tickle, L., and Jong, P. (2006). Lee-Carter Mortality Forecasting: a Multi-country Comparison of Variants and Extensions. *Demographic Research*, 15, 289 – 310.

Brouhns, N., Denuit, M., and Vermunt, J.K. (2002). A Poisson Log-linear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics*, 31, 373-393.

Case, R.A.M. (1956). Cohort analysis of mortality rates as an historical or narrative technique. *British Journal of Preventive and Social Medicine* 10, 159 – 171.

Cameron, A. C., and Windmeijer, F. A. G. (1997). An R-squared Measure of Goodness of Fit for Some Common Nonlinear Regression Models. *Journal of Econometrics*, 77, 329 -342.

Carstensen, B. (2007). Age-period-cohort models for the Lexis diagram. *Statistics in Medicine 2007*; 26: 3018-3045.

Clayton, D. and Schifflers, E. (1987). Models for temporal variation in cancer rates II: Age-Period-Cohort Models. *Statistics in Medicine*, Vol. 6, 469-481.

Coelho, E. and Nunes, L. C. (2011). Forecasting mortality in the event of a structural change. *Journal of the Royal Statistical Society A* 174, Part 3, pp 713-736.

Harris, D., Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009). Testing for a Unit-root in the Presence of a Possible Break in Trend. *Econometric Theory*, 25, 1545-1588.

Harvey, D.I, Leybourne, S.J. and Taylor, A.M.R. (2009). Simple, Robust and Powerful Tests of the Changing Trend Hypothesis. *Econometric Theory*, 25, 995-1029.

Hobcraft, J., Menken, J. and Preston, S. (1982). Age, period, and cohort effects in demography: a review. *Population Index*, Vol. 48, n.° 1, pp. 4-43.

Kuang, D., Nielsen, B. and Nielsen, J.P. (2008a). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika*, 95, 4, pp.979 – 986.

Lee, R. and Carter, L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87(419), 659-675.

Mason, K. O., Mason, W. M., Winsborough, H. H. and Poole, W. K. (1973). Some Methodological Issues in Cohort Analysis of Archival Data. *American Sociological Review*, Vol. 38, n. °2, pp. 242-258.

Murphy, M. (2010). Reexamining the dominance of birth cohorts effects on mortality. *Population and Development Review* 36 (2): 365 -390.

Perron, P., and Rodríguez, G. (2003). GLS Detrending, Efficient Unit-root Tests and Structural Change. *Journal of Econometrics*, 115, 1-27.

Renshaw, A., and Haberman, S., (2003). Lee–Carter mortality forecasting with age specific enhancement. *Insurance: Mathematics and Economics* 33 (2), 255–272.

Renshaw, A. E., and Haberman, S., (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 58: 556–570.

Shang, H. L., R. J. Hyndman, and H. Booth (2010). A comparison of ten principal component methods for forecasting mortality rates. Working paper 08/10. Department of Econometrics & Business Statistics, Monash University.

Willets, R.C. (1999). Mortality in the next millennium. Staple Inn Actuarial Society Meeting, 7 December 1999. Available at: http://www.sias.org.uk/.

Willets, R.C. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, 10, 833-877.

Zivot, E. and Andrews, D.W.K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, 10(3), 251-270.

2. FORECASTING MORTALITY IN THE EVENT OF A STRUCTURAL CHANGE*

Abstract

In recent decades, life expectancy in developed countries has risen to historically unprecedented levels driven by unforeseen declines in mortality rates. The prospects of further reductions are of fundamental importance in various areas. In this context, we consider the problem of forecasting future mortality and life expectancy in the event of a structural change. We show how recent advances in statistical testing for structural changes can be used to arrive at properly specified models for the general level of mortality in the context of the Lee-Carter model. Specifically, the results of tests for a change in the trend of the mortality index and for the presence of a unit-root are used to identify appropriate forecasting models. The proposed methodology is applied to post-1950 mortality data for a set of developed countries in order to test the usual assumption made in the literature of a sustained decline in mortality at roughly constant rates in this period. Structural changes in the rate of decline in overall mortality are found for almost every country, especially in male populations. We also illustrate how accounting for such a change can lead to a major impact in mortality and life expectancy forecasts over the next decades.

Keywords: Lee-Carter model, life expectancy, mortality forecasting, structural change, unit root.

2.1. Introduction

Over the past decades, life expectancy in developed countries has risen to historically unprecedented levels. The prospects of further future reductions in mortality rates are of fundamental importance in various areas such as demography, actuarial studies, public health, social insurance planning, and economic policy. Over the last years, significant progress has been made in mortality forecasting (for a recent review see Booth, 2006). The most popular approaches to long-term forecasting are based on the Lee and Carter (1992) model. It describes the time-series movement of age-specific mortality as a function of a latent level of mortality, also known as the overall mortality index, which can be forecasted using simple time-series methods. Some of the advantages of this approach are its simplicity and the robustness of the forecasts when age-specific log-mortality rates decline linearly over time. The method was initially used to forecast mortality in the U.S., but since then it has been applied to many other countries, including Australia (Booth, Maindonald and Smith, 2002), Austria (Carter and Prskawetz, 2001), Belgium (Brounhs, Denuit and Vermunt, 2002), Brazil (Fígoli, 1998), England and Wales (Renshaw and Haberman, 2003a, 2003b), Japan (Wilmoth, 1996), Spain (Guillen and Vidiella-i-Anguera, 2005; Debón, Montes, and Puig, 2008), Sweden (Lundstrom and Qvist, 2004), the Nordic countries (Koissi, Shapiro and Högnäs, 2006), and the G7 countries (Tuljapurkar, Li and Boe, 2000).

The original Lee-Carter model has also received a number of criticisms (see the discussion in Lee and Miller, 2001) and several extensions have been proposed in the literature (see Booth and Tickle, 2008). One major issue concerns the stability of the model over time. Since the method is usually applied to long time-series there is a risk that important structural changes may have occurred in the past. And any neglected structural change in the estimation period may result in forecasts that have a tendency to deviate from the future realizations of the mortality index, leading to potentially large long-term forecast errors. Moreover, inaccurate modelling of past trends also compromises the estimation of the past variability of mortality, leading to further biases regarding the estimation of forecast uncertainty.

In fact, historically, mortality in the USA has not always declined in a linear way as depicted in Lee and Carter (1992) for the period 1900-1989. These authors also re-

estimated their random walk with drift model for the mortality index for several shorter and more recent periods and concluded that there was some instability. As noted by Lee (2000), if one had used their method to extrapolate the mortality index backward in time, one would arrive at too high mortality rates for the beginning of the 19th century. Other studies also document that there has been a systematic overestimation of the projected mortality rates in many countries (Stoto, 1983, Koissi et al., 2006). In multi-country comparisons of several versions of the Lee-Carter method, Booth, Hyndman, Tickle, and De Jong (2006) and Shang, Hyndman, and Booth (2010) find significant differences in the forecasting performance when alternative fitting periods are used, providing evidence of changing trends in the mortality decline rate.

This paper considers the problem of forecasting future mortality and life expectancy in the event of a structural change. We show how recent advances in statistical testing for structural changes can be used to arrive at a properly specified forecasting model for the general level of mortality in the context of the Lee-Carter model. Specifically, the results of tests for a change in the trend of the mortality index and for the presence of a unit-root are used to identify the appropriate model to be estimated and used to carry out the projections. In particular, it allows one to detect if a change is present or not, to estimate the date where the change occurs, and to determine if the appropriate forecasting model should be based on the levels or the first-differences of the mortality index series. The last distinction is important as the two alternatives lead to substantial differences in the amplitudes of the forecast confidence intervals.

The proposed methodology is applied to post-1950 time-series mortality data for 18 developed countries from Western Europe and North America, and for Japan, for which reliable data is available. According to several studies, during this period, the majority of the developed countries have experienced a sustained decline in mortality at roughly constant rates (see Tuljapurkar et al., 2000, and Lee and Miller, 2001). However, recent studies have questioned the linearity of the decline in mortality, and provide evidence of changes in the pattern of decline in this period (some examples are Booth et al., 2002, for Australia; Renshaw and Haberman, 2003a, for England and Wales; and Booth et al., 2006, for several developed countries). By applying our proposed tests, we provide statistical evidence regarding this question. Indeed,

according to our results, structural changes in the rate of decline in the overall mortality rate are found for almost every country considered and especially in the male populations. We consider as an example the case of Portuguese male mortality and show that accounting for a structural change leads to a major impact in mortality and life expectancy forecasts over the next decades.

The article is organized as follows. Section 2 presents a brief review of the Lee-Carter method and its extensions. The proposed approach to forecast mortality in the presence of structural change is described in Section 3. The empirical application is presented in Section 4. Finally, Section 5 discusses the results obtained and presents some concluding remarks.

2.2.   The Lee-Carter Method

The Lee-Carter method combines a demographic model of age-specific death rates with statistical time-series forecasting methods. In this section, we give a brief description of the implementation of these two components of the method. We also present the Poisson regression extension that we follow in our empirical application.

2.2.1.   The demographic model

The core assumption behind the Lee-Carter (1992) model is that the evolution of death rates for any age x is driven by a common time-varying component, denoted by k_t, which is also referred to as the overall mortality index. More precisely, the following relation is assumed:

$$\ln m_x(t) = \alpha_x + \beta_x k_t + \varepsilon_x(t) \tag{1}$$

where $m_x(t)$ is the central death rate for age x and calendar year t, α_x are the age-specific parameters that affect the overall level of $\ln m_x(t)$ over time, β_x are the age-specific parameters that characterize the sensitivity of $\ln m_x(t)$ to changes in the mortality index k_t, and $\varepsilon_x(t)$ represent error terms capturing particular age-specific

historical influences not explained by the model. These errors are assumed to have zero mean and constant variance. Lee and Carter estimate the age-specific death rates by the following ratio:

$$\hat{m}_x(t) = \frac{D_{x,t}}{E_{x,t}} \tag{2}$$

where $D_{x,t}$ denotes the number of deaths recorded at age x during year t from a corresponding exposure-to-risk $E_{x,t}$.

Lee and Carter (1992) propose a two-stage procedure to estimate the model given by equation (1). First, a least-squares solution to estimate α_x, β_x, and k_t is found. Since the model is clearly underdetermined, some normalization constraints are imposed to obtain a unique solution: the sum of the β_x over all ages equals one and the sum of the k_t over time equals zero. As a consequence, the α_x will be equal to the average values of $\ln \hat{m}_x(t)$ over time. The solution is found using the singular value decomposition. Since the model is written in terms of log mortality, the observed total number of deaths in each year will not equal the sum of the fitted deaths by age. To ensure this equality, the k_t are estimated a second time, taking the α_x and β_x estimates from the first step, such that for each year t, given the actual age distribution for the population, the implied number of deaths equals the observed number of deaths.

The homoskedasticity assumption on the error term $\varepsilon_x(t)$, implied by the singular value decomposition estimation of the parameters α_x, β_x, and k_t, has been considered fairly unrealistic (Wilmoth, 1993; Alho, 2000) since the logarithm of the age-specific death rate is much more volatile at old ages, where the number of deaths and of those exposed to risk are relatively small. To circumvent this problem, Brouhns, Denuit and Vermunt (2002) propose an alternative model that keeps the Lee-Carter log-bilinear functional form, but replaces the least-squares approach with a Poisson regression for the number of deaths. Specifically, they consider the following model describing the distribution of $D_{x,t}$:

$$D_{x,t} \sim Poisson(E_{x,t}\mu_x(t)) \quad \text{with} \quad \mu_x(t) = \exp(\alpha_x + \beta_x k_t), \tag{3}$$

where the parameters α_x, β_x, and k_t retain the meaning originally attributed by the Lee-Carter model, and where $\mu_x(t)$ denotes the force of mortality. Note that if the force of mortality is assumed to be constant within bands of age and calendar years, but allowed to vary from one band to the other, that is

$$\mu_{x+y}(t+\tau) = \mu_x(t) \quad \text{for} \quad 0 \le y, \tau < 1, \tag{4}$$

it follows that the force of mortality is identical with the central death rate, $\mu_x(t) = m_x(t)$.

The parameters of the Poisson model can be estimated by maximum likelihood still subject to the model identification constraints that the sum of the β_x over all ages equals unity and the sum of the k_t over time equals zero. The second stage in the estimation of the k_t, required in the classical Lee-Carter model, is no longer necessary with this Poisson model since the sum of the estimated number of deaths by age will automatically equal the total observed deaths in each year. Note that, although we use the same notation for k_t in equations (1) and (3), the estimates obtained from the Lee-Carter and the Poisson regression will in general be different. Examples of the application of the Poisson model are Brouhns et al. (2002) for Belgium, Renshaw and Haberman (2003b) for England and Wales, and Koissi et al. (2006) for the Nordic countries.

Given the estimated α_x and β_x in equation (1) or (3), the problem of forecasting age-specific mortality rates and, consequently, life expectancies, is reduced to forecasting the mortality index k_t. This is considered in the following subsection.

2.2.2. Modelling the index of mortality

Lee and Carter (1992) propose using the Box-Jenkins methodology to arrive at an appropriate ARIMA model to forecast the mortality index. In that methodology, the first step always consists of determining if some transformation of the series is necessary to induce stationarity before identifying and estimating the forecasting model. Many studies have arrived at an ARIMA(p, 1, q) model, that is, a stationary

ARMA(p,q) model fitted to the first-differences of the mortality index. For instance, an ARIMA(0,1,0) is adopted in the Lee and Carter (1992) study of U.S. mortality and in Tuljapurkar et al. (2000) for the G7 countries, and an ARIMA(1,1,0) is selected in Renshaw and Haberman (2003a) for England and Wales. Series whose first-differences are stationary are also called difference-stationary, and are described by a unit-root in their autoregressive representation. The usual approach to check this is by analysing the behaviour of the empirical autocorrelation functions. However, it is also possible to use formal statistical tests. The most popular are variants of the original Dickey and Fuller (1979) tests for the presence of a unit-root.

For illustration, consider an ARIMA (0,1,0) model, that is, a random walk with drift:

$$k_t - k_{t-1} = d + u_t, \qquad (5)$$

where the drift d gives the mean annual change in k_t, and u_t are i.i.d. errors. The drift d is estimated as the time-average of the series of first-differences $\Delta k_t = k_t - k_{t-1}$. Given any initial value for k_t, this model implies that:

$$k_{t+s} = k_t + ds + u_{t+1} + \cdots + u_{t+s}, \qquad (6)$$

that is, future values of the mortality index equal the sum of a deterministic linear trend component with slope d, given by the term $k_t + ds$ in equation (6), plus an error component given by the sum of several shocks u. Since part of the uncertainty when forecasting k_{t+s} comes from the cumulative sum of the shocks, $u_{t+1} + \cdots + u_{t+s}$, it follows that the forecast uncertainty grows with the forecast horizon s. It is straightforward to show that this result also holds for any other ARIMA(p,1,q) model.

If for a particular mortality index series the unit-root/difference-stationary hypothesis is rejected, the alternative hypothesis that should be considered would be that of stationarity around a linear trend capturing the linear decline in mortality. In such a case, the mortality index would be described by the following trend-stationary model:

$$k_t = d_0 + d_1 t + u_t, \qquad (7)$$

with u_t being a stationary ARMA process. In this equation, the d_1 parameter represents the slope of the linear trend. Examples of applications where a deterministic time trend model is assumed are Sithole, Haberman and Verral (2000) and Denuit and Goderniaux (2005). Renshaw and Haberman (2003a) also consider a deterministic time trend in the

context of a generalized linear modelling version of the Lee-Carter model. Since model (7) is specified and estimated in the levels of k_t, not the first-differences, it will in general produce forecasts that are different from the previous unit-root model. Also, given that in this model the deviations from the linear trend, given by u_t, are stationary, the forecast uncertainty will no longer have a tendency to increase with the forecast horizon.

Unfortunately, the usual procedures based on the analysis of the autocorrelation function or on unit-root tests to decide about the appropriate class of forecasting model to be used, are only valid if the linear trend, and in particular the trend slope, d in model (6) or d_1 in model (7), has remained constant over time. In the next section we discuss in more depth the implications of the presence of a structural change in the trend of the mortality index and propose a suitable approach to cope with this possibility when forecasting the mortality index.

2.3.   Allowing for a Structural Change

A large literature on structural change models has emerged in the last years (see the surveys by Perron 2006, 2008). An important result is that, when producing long-term forecasts, any neglected or wrongly placed structural change occurring in the estimation phase may result in forecasts with a tendency to deviate from the future realizations of the series, resulting in potentially large forecast errors. It is also known that the correct approach to test for and estimate a model with a structural change is highly dependent on the stationarity properties of the time-series process.

In particular, if the mortality index series k_t is known to be trend-stationary then i) the tests for the presence of a structural change in the trend and ii) the estimation of the forecasting model should both be carried out using regressions based on the levels of the series. An example is the join-point regression approach that has been often used in the analysis of cancer mortality and survival rates (see Kim, Fay, Feuer, and Midthune, 2000, and Yu, Huang, Tiwari, Feuer, and Johnson, 2009) and is implemented in the Joinpoint software distributed by the U.S. National Cancer Institute (available at http://srab.cancer.gov/joinpoint). The approach can account for autocorrelated data,

although this option is not available in the latest version of the program (version 3.4.3). However, the method performs badly in the presence of strong positive autocorrelation (see Kim et al., 2000). Moreover, it is not valid in the presence of a unit-root, as is the case of any ARIMA(p,1,q) process, since it will lead to spurious breaks (see Nunes et al., 1996).

On the other hand, if it is known that the k_t series is non-stationary with a unit-root, first-differences should be used instead. In general, the results obtained by following these two alternatives can be quite different in terms of a presence or not of a structural change and, in case one is detected, in terms of its dating.

In the literature of mortality forecasting, there are also some examples of approaches to account for the possible occurrence of a structural change. Booth et al. (2002) find some evidence that the decline in the mortality index k_t is not linear in their application of the Lee-Carter model to Australia, and try to avoid producing biased forecasts by selecting a best fitting period that insures linearity of the trend component assuming a unit-root. The criterion used is based on the ratio of a measure of the lack of fit of the model imposing a constant coefficients ARIMA model to the overall lack of fit of the Lee-Carter model. Lack of fit of the models are measured by mean deviance statistics. The chosen fitting period is that for which this ratio is substantially smaller than for periods starting in previous years. Denuit and Goderniaux (2005) also propose a similar solution, but assuming stationarity of the error term. The optimal starting year is determined by maximizing the R^2 statistic for a linear regression as in equation (7). Therefore, both methods make an *a priori* assumption about the existence, or not, of a unit-root. Moreover, for both methods, the decision about the existence of a structural change and of its location, rely on non-objective criteria based on visual inspection of the time plots of model fit ratios or the R^2. Another disadvantage is that since a specific fitting period is chosen, all information before the selected starting year is discarded. Therefore, these methods do not use all the available information available to estimate the variability of past mortality which is essential for instance in producing reliable forecast confidence intervals.

As mentioned above, the traditional approaches to detect the presence of a unit-root and the need to first-difference the data are based on examining the empirical autocorrelation functions or on performing standard unit-root tests. However, as shown

by Perron (1989), these approaches are not valid if a structural change is present, as they have a tendency to wrongly suggest a unit-root when the series is, in fact, stationary around a broken trend. As a solution, Perron (1997) proposes a modified unit-root test that is valid in the presence of a structural change at a known date. Zivot and Andrews (1992), Perron and Rodríguez (2003), and other authors have extended the test to the case where the change date is not known and must be estimated. However, when no change is actually present, several problems arise. As shown in Harris, Harvey, Leybourne, and Taylor (2009) (henceforth HHLT) there will be severe efficiency losses. Moreover, the estimated change date suggested by these tests may be spurious since, as shown by Nunes, Kuan, and Newbold (1995), the presence of a unit-root may lead one to erroneously find a structural change when there is none. In fact, when using the Zivot and Andrews (1992) unit-root test, the estimated change date is not even consistent for the true date if a change does occur. An example where such an inappropriate approach is followed is Chan, Li, and Cheung (2008) in a study of mortality in Canada, England and Wales, and the United States. Another issue is that the correct critical values to implement these tests become dependent on whether a structural change is present or not.

We propose a solution to these dilemmas in the context of mortality forecasting by following a simple and consistent sequential approach based on recent statistical results in the literature of structural change. In a first step, since it is not possible to know beforehand if there is a unit-root or not, we apply a test for a structural change in the trend proposed by Harvey, Leybourne, and Taylor (2009) (henceforth HLT) that is valid regardless of whether the series is difference-stationary or trend-stationary. Secondly, as in HHLT, the result of this structural change test is used to decide on the appropriate unit-root test to use: with or without allowing for a structural change. Thirdly, if according to the HLT test a change is present, the break date is estimated taking into account the result of the HHLT test, that is, a) using the first-differences of the series if a unit-root is found or b) using the levels of the series otherwise. Finally, based on the conclusions of the previous steps, the appropriate ARIMA model is fitted to the k_t series and used to produce the forecasts. In the following subsections we describe these structural change and unit-root tests in more detail.

2.3.1. Testing for a change in the trend

As mentioned above, we use the HLT test for the existence of a structural change in the trend of the mortality index k_t when it is not known *a priori* if the series has a unit-root or not. The date of the break, if present, is estimated from the available sample. The test is relatively simple to implement since it only requires estimating the following two linear regression models by least squares:

$$k_t = \alpha + \beta t + \gamma DT_t(\tau) + u_t, \quad t = 1, \ldots, T, \tag{8}$$

and

$$\Delta k_t = \beta + \gamma DU_t(\tau) + u_t, \quad t = 2, \ldots, T, \tag{9}$$

where the change dummy variables are defined as $DT_t(\tau) = t\text{-}T_B$ if $t > T_B$ and $DT_t(\tau) = 0$ if $t \leq T_B$, $DU_t(\tau) = 1$ if $t > T_B$ and $DU_t(\tau) = 0$ if $t \leq T_B$, with $T_B = [\tau T]$ denoting the possible trend change date, and τ the associated change date fraction with $\tau \in (0, 1)$. Equation (8) corresponds to the trend-stationary case with stationary shocks u_t and k_t fluctuating around a linear trend with slope β subject to a change of magnitude γ occurring at a given date T_B. Equation (9) corresponds to the case where the shocks to k_t are non-stationary with a unit-root, so that the first-differences, $\Delta k_t = k_t - k_{t-1}$, fluctuate around the mean, given by the drift parameter β which is also subject to a change of magnitude γ occurring at date T_B. In this last equation, the shocks u_t to Δk_t are also stationary. In both equations, the null hypothesis of no change in the trend slope corresponds to H_0: $\gamma = 0$. The model described here corresponds to Model A in HLT. It is straightforward to allow for a simultaneous change in the slope and the level of the trend (Model B in HLT) by including an extra dummy variable in each regression. However, if the mortality index does not show any abrupt changes in its level, it is not necessary to consider such a model.

If k_t is known to be trend-stationary, a valid test for the presence of a structural change in the trend slope can be based on the t-statistic for testing H_0 in equation (8), which we denote as $t_0(\tau)$. To allow for serial correlation in the error term, an auto-correlation robust t-statistic must be used. Since the change date τ is not known *a priori*,

but may be inferred from the data itself, the test statistic is computed by the maximum of the sequence of $t_0(\tau)$ statistics for all possible change fractions τ as:

$$t_0^* = \sup\nolimits_{\tau \in \Lambda} | t_0(\tau) | \tag{10}$$

where the supremum is taken over a set $\Lambda = [\tau_L, \tau_U]$, with $0 < \tau_L < \tau_U < 1$. HLT suggest setting the *trimming* parameters τ_L and τ_U equal to 0.1 and 0.9 respectively. Allowing for structural changes too close to the beginning or the end of the sample may lead to erroneous conclusions on the existence of a structural change in the trend since there may not be enough observations to support it. In our application presented in Section 4, this corresponds to the first five and last five observations in the sample. Li and Chan (2005) in a study of outliers also mention that the identified type and the estimation of an outlier may not be reliable when it lies within the last five observations of the series.

On the other hand, if it is known that k_t is difference-stationary, one should use the corresponding t-statistic from equation (9), which we denote as $t_1(\tau)$. As in the previous case, given that the change date is not known *a priori*, the test statistic is given by

$$t_1^* = \sup\nolimits_{\tau \in \Lambda} | t_1(\tau) |. \tag{11}$$

The change fractions at which the test statistics t_0^* and t_1^* attain their maximum will be denoted as τ_0^* and τ_1^*, respectively.

Finally, the HLT test, which is valid for both unit-root and trend-stationary processes, is computed as a weighted average of the t_0^* and t_1^* tests:

$$t_\lambda^* = \lambda\, t_0^* + m\,(1-\lambda)\, t_1^* \tag{12}$$

with the weight λ given by

$$\lambda = \exp[\, -\,(\, 500\, S_0^*\, S_1^*\,)^2\,] \tag{13}$$

and S_0^* and S_1^* denoting the KPSS (see Kwiatkowski, Phillips, Schmidt, and Shin, 1992) stationary test statistics calculated from the OLS residuals of equations (8) and (9) when evaluated at the τ_0^* and τ_1^* change dates respectively. As shown in HLT, λ converges asymptotically to 1 if the series is trend-stationary or to 0 if it is difference-stationary at sufficiently fast rates so that the correct test statistic, t_0^* or t_1^* respectively, is selected. As shown by HLT, the 5% asymptotic critical value of t_λ^* equals 2.563 provided the constant m is set equal to 0.853.

2.3.2. Unit-root test

As mentioned above, standard unit-root tests have several shortcomings when it is not known *a priori* if a structural change has occurred or not. However, as shown in HHLT, the HLT t_λ^* structural change test in (12) can be used as a pre-test. If it rejects the null hypothesis of no change, an optimal test for a unit-root in the presence of a structural change should be used. If it does not reject, then an optimal unit-root test without allowing for a structural change should be used. We now describe these unit-root tests in more detail.

For the case where a structural change is detected with the HLT test, the estimated change fraction using the first-differenced model, τ_1^*, is used to estimate model (8) by GLS as in Perron and Rodriguez (2003). For $\tau = \tau_1^*$, equation (8) can be rewritten as

$$k_t = X_t(\tau_1^*)\,\theta_0 + u_t, \quad t = 1, \ldots, T, \tag{14}$$

where $X_t(\tau_1^*) = (1, t, DT_t(\tau_1^*))$ and $\theta_0 = (\alpha, \beta, \gamma)'$. The GLS estimation is implemented by estimating with OLS the following quasi-difference transformation of equation (14):

$$k_{c,t} = X_{c,t}(\tau_1^*)\,\theta_0 + u_{c,t}, \quad t = 1, \ldots, T, \tag{15}$$

where $k_{c,t} = k_t - \rho\,k_{t-1}$ and $X_{c,t}(\tau_1^*) = X_t(\tau_1^*) - \rho X_{t-1}(\tau_1^*)$ for $t = 2, \ldots, T$, $k_{c,t} = k_1$ and $X_{c,t}(\tau_1^*) = X_1(\tau_1^*)$ for $t = 1$, and $\rho = 1 - c/T$ where c is the quasi-difference parameter. The value of c can be chosen according to a local power criterion as explained in HHLT. Let $\hat{\theta}_c$ denote the OLS estimator of θ_0 in equation (15) and $\tilde{u}_t = k_t - X_t(\tau_1^*)\hat{\theta}_c$ the residuals. The unit-root test is finally obtained by estimating the following ADF regression:

$$\Delta\tilde{u}_t = \phi\tilde{u}_{t-1} + \sum_{j=1}^{p}\delta_j\Delta\tilde{u}_{t-j} + e_{p,t}, \, t = p+2,\ldots,T. \tag{16}$$

The number of lags p can be chosen by a modified Akaike criterion (see HHLT for details). Values of c and critical values for several significance levels and estimated change fractions needed for the implementation of this test can be found in HHLT and Carrion-i-Silvestre, Kim, and Perron (2008).

When no structural change is found by the HLT test, a unit-root test without allowing for a change should be used such as the ADF-GLS optimal test proposed by

Elliot, Rothenberg, and Stock (1996). This test is computed as above but with the following two modifications: a) the structural change dummy variable is excluded from the $X_t(\tau_1^{*})$ vector and b) the optimal value of c in this case equals 13.5.

2.4. Empirical Application

In this section, the methodology proposed above is illustrated by an application to the following 18 developed countries: Austria, Belgium, Canada, Denmark, England and Wales, Finland, France, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Switzerland, Sweden, United States of America, and West Germany. These countries experienced large increases in life expectancy during the twentieth century, interrupted only by the global influenza epidemic in 1918-19 and the two World Wars (Bongaarts, 2006). However, the decline in mortality rates was not simultaneous in all countries, and rather diverse situations prevailed until the Second World War (Monnier, 2004). The increases in life expectancy were more pronounced during the first half of the century, affecting in particular early ages in life (Wilmoth, 2000). The pace of improvement in life expectancy slowed down around the middle of the century, when the predominant types of mortality reduction shifted from curing infectious diseases that heavily affect the young to degenerative diseases that largely affect the elderly (see White, 2002, and Vallin and Meslé, 2004). Several studies also provide evidence that the age pattern of mortality decline changed over time (e.g. Kannisto, Lauritsen, and Thatcher 1994; Horiuchi and Wilmoth, 1998; Wilmoth, 1998; Carter and Prskawetz, 2001; Lee and Miller, 2001).

Lee and Miller (2001) found evidence in the United States, Sweden, Japan, France and Canada, that the current pattern of mortality change began around the mid-twentieth century and suggest fitting the Lee-Carter model on data since 1950 to reduce structural shifts. Tuljapurkar et al. (2000), in their study of the G7 countries, also adopt 1950 as the initial year of analysis. However, as we discussed above, a number of studies have questioned the linearity of the decline in mortality since 1950 and provide evidence of changes in the pattern of decline in recent years. We investigate this question by applying our proposed methodology based on statistical testing procedures.

We use time-series data obtained from the Human Mortality Database (HMD), sponsored by the University of California, Berkeley, and the Max Planck Institute for Demographic Research (http://www.mortality.org), for the number of deaths, $D_{x,t}$, and exposure-to-risk, $E_{x,t}$, by single year of age, sex, and calendar year from 1950 (for West Germany data are available only since 1956) until the most recent available year (2005 for Austria, Canada, and U.S.A., 2006 for Belgium, England & Wales, France, Ireland, Italy, Netherlands, Spain, and West Germany, and 2007 for Denmark, Finland, Japan, Norway, Portugal, Switzerland, and Sweden). The $E_{x,t}$ are based on annual population estimates with a small correction that reflects the timing of deaths during the interval (see HMD, 2007).

We first apply the Brouhns et al. (2002) Poisson regression extension of the Lee-Carter method described in Section 2 to obtain estimates of the model parameters and the mortality index series for the male and female populations in each country. Next, we apply the procedure proposed in Section 3 to test for the presence of structural changes and unit-roots in the estimated mortality indices. We note that the proposed procedure does not affect the estimation of the k_t. Finally, we illustrate how these results can be used to arrive at appropriate forecasting models by considering the case of male mortality in Portugal. We compare the resulting mortality and life expectancy forecasts with those obtained using alternative models.

2.4.1. Estimation of the index of mortality

The estimated mortality indices k_t for the male and female populations in each country are plotted in Figures 1.a-c and confirm the downward trend in mortality over time observed in all countries. A more detailed visual inspection of the graphs also seems to suggest that after the mid-1970s, the rate of decline, especially for male mortality, might have become more accentuated in a number of countries. Examples of perceptible structural changes can be found in the mid-1970s for Belgium, Canada, and West Germany, mid-1980s for Italy, Norway, and Sweden, mid-1990s for Denmark, and around 2000 for Ireland and Netherlands.
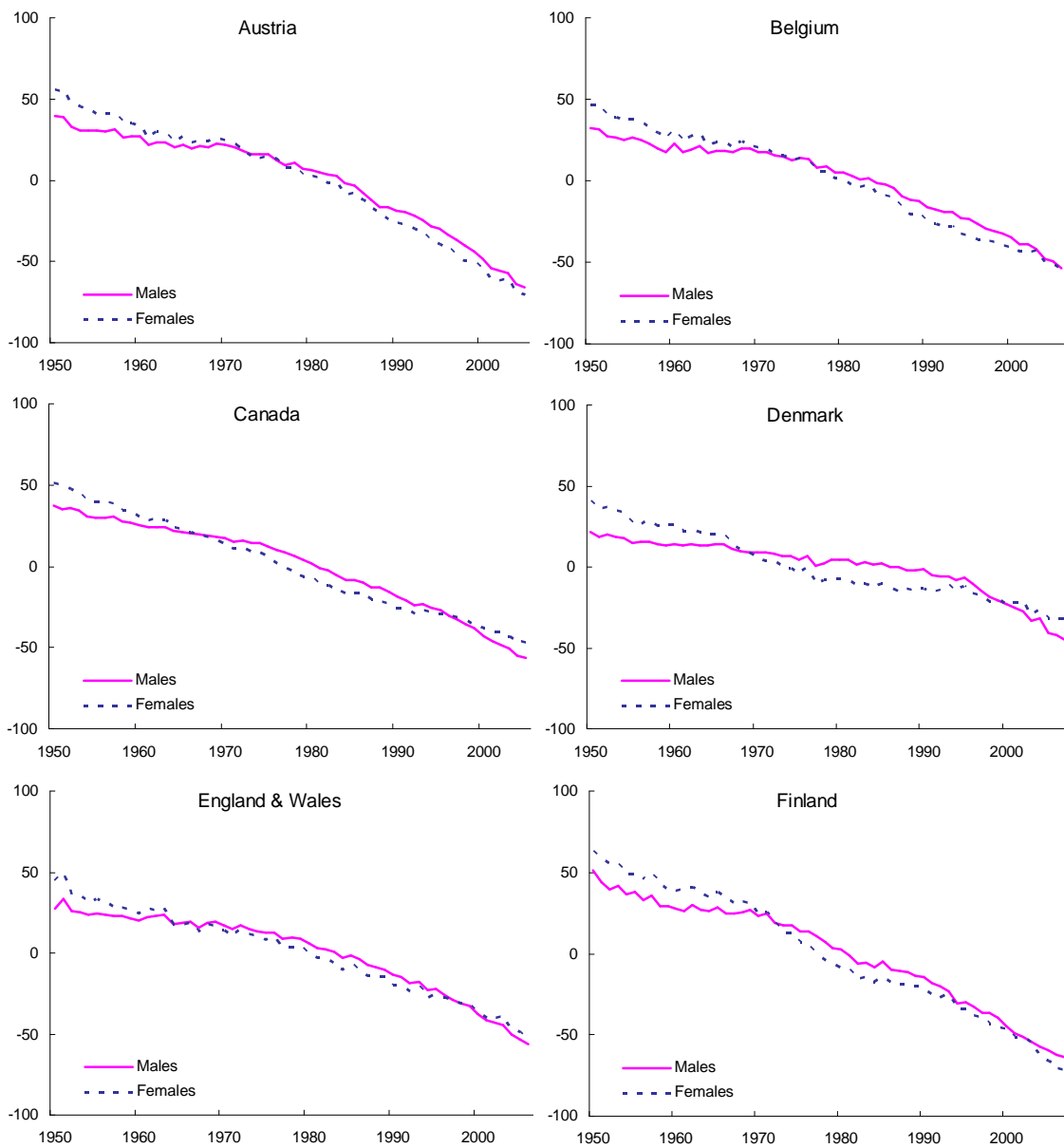
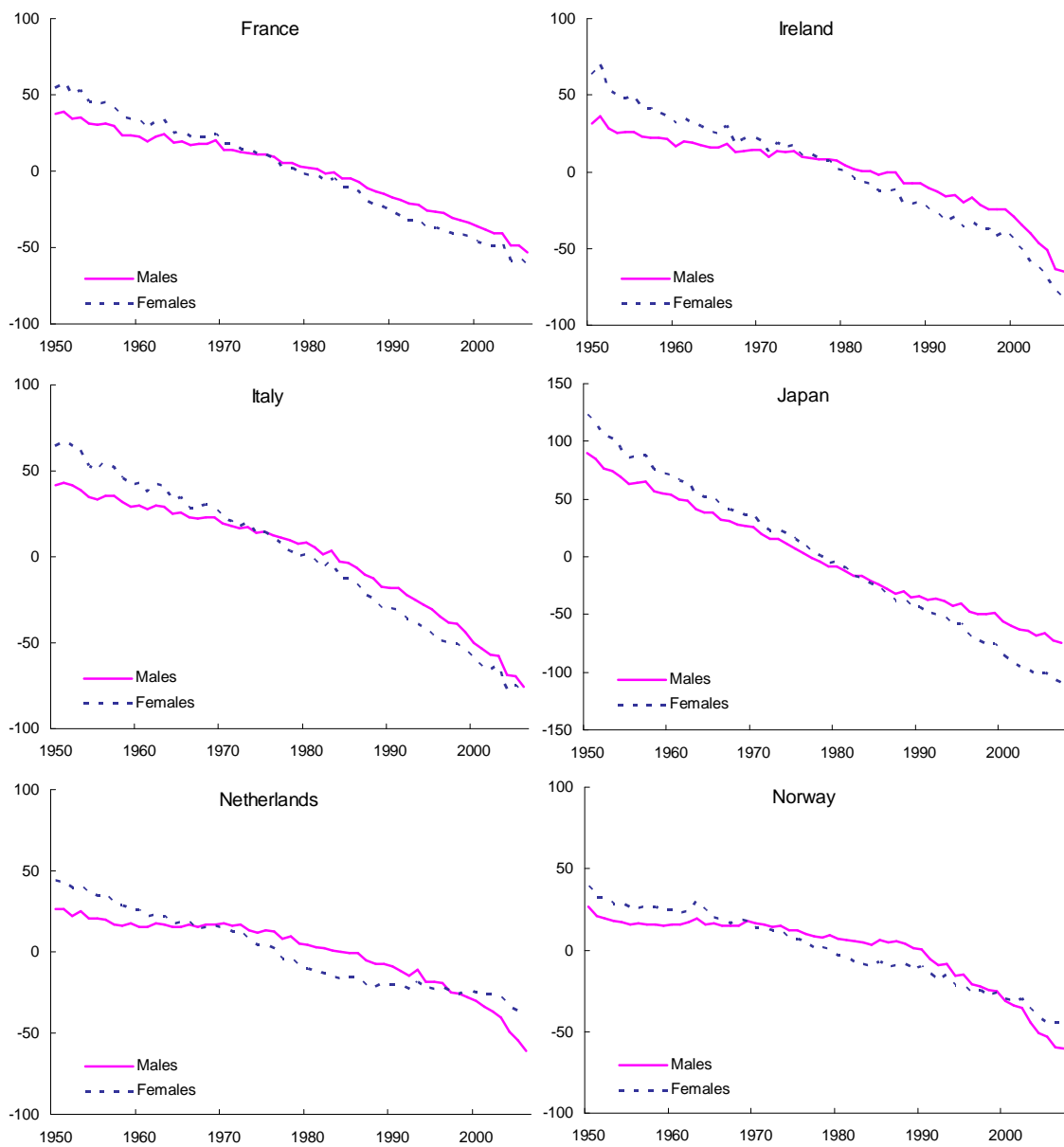*Figure 1.a.* Estimated mortality indices for selected countries.

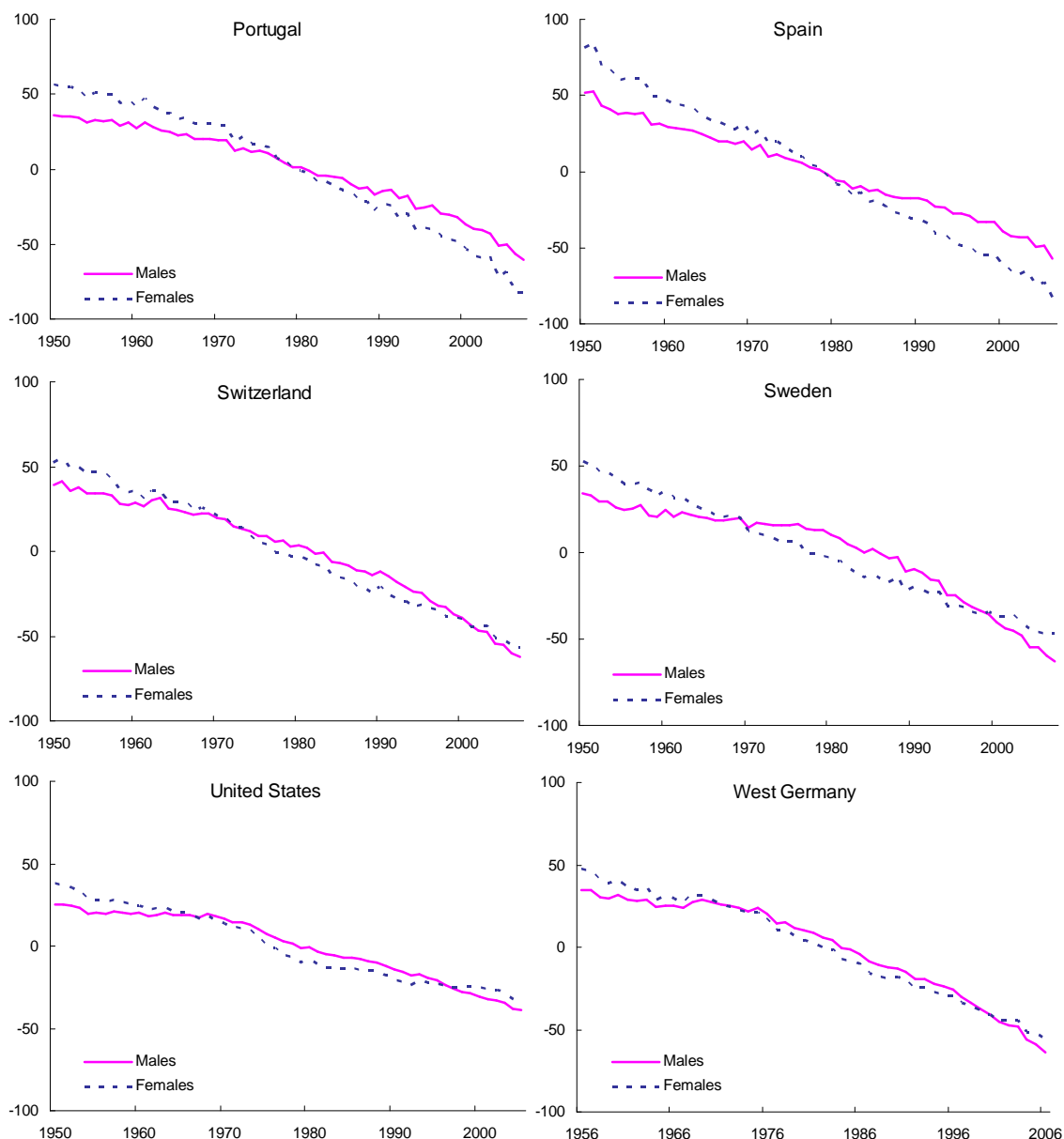*Figure 1.b.* Estimated mortality indices for selected countries.

Figure 1.c. Estimated mortality indices for selected countries.

However, as explained in the previous section, if the time-series process has a unit-root, there is a risk that such apparent structural changes may be nothing more than illusions in the data. The problem of spurious breaks was first raised, from a graphical perspective, by Hendry and Neale (1991). A statistical explanation of the phenomena was given by Nunes et al. (1995, 1996). Next, we present formal and valid tests for the genuine presence of structural changes in the mortality indices.
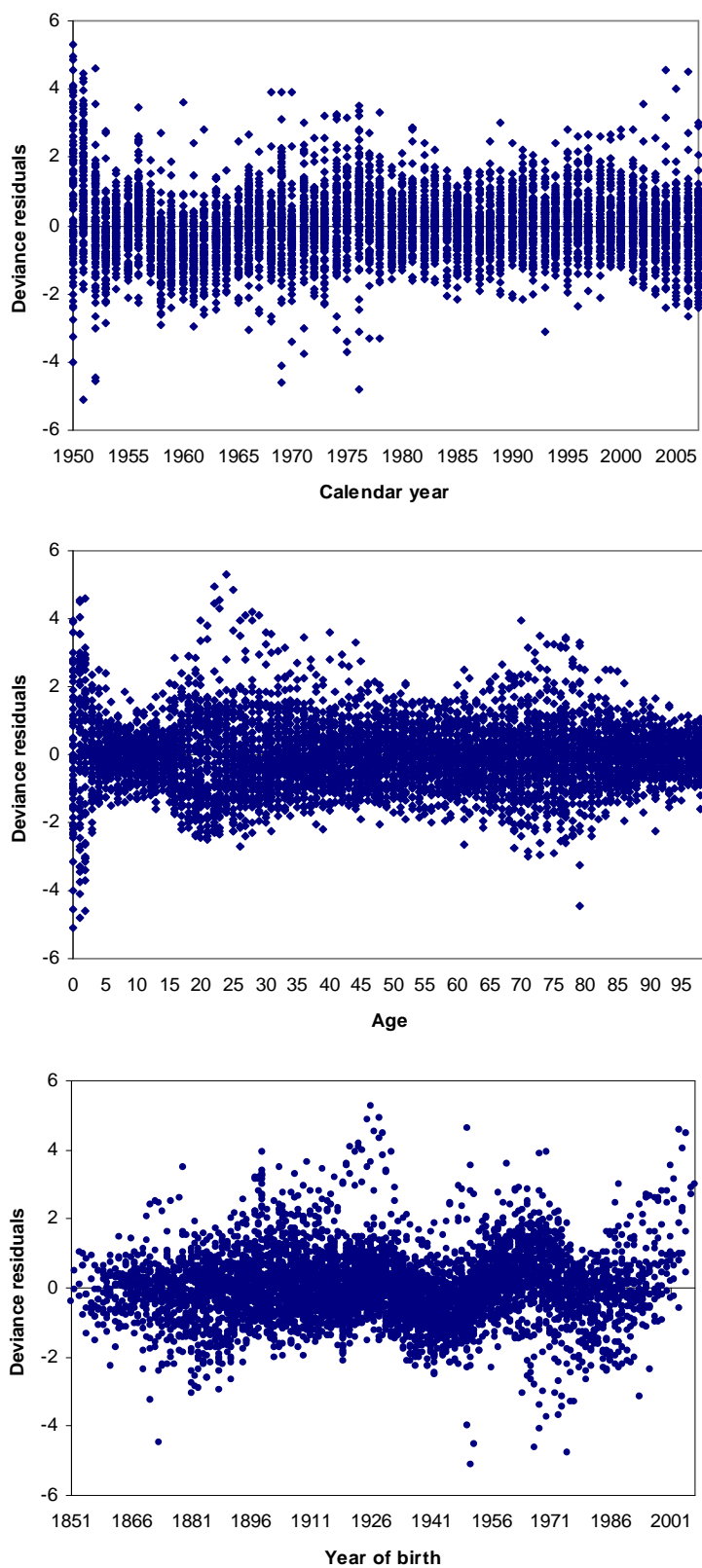
Figure 2. Plots of Deviance Residuals of the Poisson Lee-Carter model for Portuguese Males.

2.4.2. Structural change and unit-root tests

Before conducting a time series analysis of the mortality indices, we start by looking at potential outliers that may be present in the data. We follow the method proposed by Chen and Liu (1993) modified by Goméz and Maravall (1997) which is implemented in the program TRAMO, and included in the DEMETRA software distributed by Eurostat. A similar approach is also used in Li and Chan (2007), who propose an outlier-adjusted model for k_t for Canada and the United States, and Li and Chan (2005) for the UK and the Scandinavian countries. This method is able to detect four types of outliers: additive outlier (AO) which affects only the level of a single observation, innovational outlier (IO) affecting all observations beyond some date through the dynamics of the process, level shift (LS) affecting the level of all observations after a given time, and temporary change (TC) affecting all observations after a given point but with an exponentially decaying effect. It is important to note that this outlier analysis is able to detect changes in the level of the series but not in the trend slope. In TRAMO, the critical value used to detect the outliers is determined by the number of observations in the series based on simulation experiments. We also exclude dates within five observations of the start or end of the sample.

Very few outliers are actually found. For West Germany females a LS in 1968; for Sweden males an IO in 1989; for France males a TC in 1958 and AOs in 1961 and 1969; and for Denmark males a TC in 1977. Therefore, for only 4 out of the 36 series analyzed, were outliers found. This suggests that in the second half of the 20th century and beginning of the 21st, there is little evidence of major disrupting events that would require some outlier correction. If our analysis had started in the beginning of the 20th century, the outlier analysis would have captured significant events, such as the two World Wars or the Spanish flu. Lin and Chan (2005, 2007) find several outliers throughout the 20th century. However, if the authors had used the value recommended in TRAMO according to the number of observations in the series, fewer outliers would have been found. For instance, in the case of England and Wales, all outliers would occur in the 1st half of the 20th century. Therefore, in what follows, we opted for using the original series without using any outlier correction.

The traditional Box-Jenkins approach to forecasting would begin by analysing the stationarity of each series. For all the estimated k_t, the corresponding empirical autocorrelation functions approach zero very slowly while, on the contrary, the empirical partial autocorrelation functions cut off abruptly at lag one. The same behaviour is obtained after removing a linear trend from the k_t series. These are typical behaviours of non-stationary unit-root series and, according to the usual Box-Jenkins approach, a clear indication that the k_t series should be first-differenced, leading therefore to ARIMA(p,1,q) forecasting models. However, as explained above, this analysis is not valid in the presence of structural changes.

We now follow the sequential approach described in Section 3 to identify the correct forecasting model in the potential presence of a structural change. The results of the different structural change and unit-root tests for males and females are summarized in Table 1.

We begin by looking at the results of the t_0^* and t_1^* structural change tests. Since there is no reason or evidence pointing to simultaneous changes in the level and slope of the downward trend in mortality, the structural change tests are performed using model A in HLT, which considers only a change in the trend slope. For males, according to both tests, there is evidence of a structural change in the trend slope in all countries, except Finland and Spain. In these two countries, evidence is favourable to a structural change only in the case of the t_0^* test. The results of the robust t_λ^* test confirm the presence of a structural change in all countries except Finland and Spain.

For females, the situation is quite different. Evidence is in general favourable to a structural change in the trend slope when using the t_0^* test, except for Finland, Spain, and Switzerland. However, for the t_1^* test, which is valid only if a unit-root is present, evidence of a structural change is favourable only for Austria, Ireland, and Japan. The results of the t_λ^* test confirm the presence of a structural change only in Austria, Ireland, Italy, Japan, and Sweden.

Table 1. *Structural change and unit-root tests*

Males	Structural Change Tests			Unit-Root Tests		Conclusion	
Country	t_0^*	t_1^*	t_λ^*	ADF-GLS	ADF-GLS-break	I(0)/I(1)	Break Date
Austria	17.21**	4.85**	4.49**	-0.93	-1.83	I(1)	1983
Belgium	15.91**	3.70**	3.16**	-0.64	-1.72	I(1)	1976
Canada	14.97**	5.59**	4.77**	-0.26	-2.01	I(1)	1975
Denmark	26.10**	7.40**	10.83**	-0.46	-3.24***	I(0)	1994
England &	17.40**	5.01**	4.27**	-0.64	-1.90	I(1)	1979
Finland	8.98**	2.23	1.90	-1.41	-1.70	I(1)	-
France	12.41**	3.13**	3.93**	-1.15	-2.39	I(1)	1985
Ireland	10.04**	8.02**	6.81**	-1.02	-1.40	I(1)	1999
Italy	18.10**	5.80**	4.95**	-0.95	-2.51	I(1)	1983
Japan	8.92**	3.44**	2.94**	-0.80	-1.41	I(1)	1955
Netherlands	9.78**	5.37**	4.58**	-0.68	-1.18	I(1)	2000
Norway	19.54**	5.64**	4.81**	-1.27	-2.00	I(1)	1988
Portugal	8.32**	3.77**	3.18**	-0.10	-1.83	I(1)	1996
Spain	3.49**	2.77	2.32	-2.24	-3.20***	I(1)	-
Sweden	22.86**	5.35**	4.83**	-1.21	-1.66	I(1)	1988
Switzerland	10.03**	4.00**	3.41**	-0.92	-1.79	I(1)	1990
United States	14.58**	4.55**	4.04**	-0.78	-2.79	I(1)	1968
West Germany	13.72**	4.88**	4.16**	-0.83	-1.73	I(1)	1975

Females	Structural Change Tests			Unit-Root Tests		Conclusion	
Country	t_0^*	t_1^*	t_λ^*	ADF-GLS	ADF-GLS-break	I(0)/I(1)	Break Date
Austria	12.82**	3.10**	2.66**	-0.94	-1.42	I(1)	1983
Belgium	7.69**	2.00	1.71	-1.43	-2.00	I(1)	-
Canada	5.23**	1.97	1.68	-1.77	-2.80	I(1)	-
Denmark	6.46**	2.55	2.17	-1.25	-2.31	I(1)	-
England & Wales	5.90**	2.32	2.19	-1.41	-2.05	I(1)	-
Finland	2.44	1.91	1.63	-1.53	-1.83	I(1)	-
France	3.97**	1.40	1.42	-1.78	-2.61	I(1)	-
Ireland	7.12**	5.22**	4.46**	-1.74	-1.38	I(1)	1999
Italy	12.05**	2.61	5.41**	-0.78	-1.77	I(1)	1983
Japan	5.75**	4.26**	4.19**	-1.46	-4.32***	I(0)	1955
Netherlands	6.52**	1.81	1.54	-1.16	-1.80	I(1)	-
Norway	4.29**	1.57	1.34	-2.24	-2.57	I(1)	-
Portugal	9.86**	2.88	2.46	-0.43	-2.20	I(1)	-
Spain	1.59	1.45	1.24	-3.02	-2.96	I(1)	-
Sweden	5.19**	1.44	3.35**	-1.00	-3.51***	I(0)	1984
Switzerland	1.94	1.60	1.36	-1.92	-2.28	I(1)	-
United States	3.76**	1.96	1.67	-1.50	-2.05	I(1)	-
West Germany	7.26**	2.34	1.99	-1.56	-2.05	I(1)	-

Notes: ** denotes rejection of the null hypothesis of no structural change at the 5% level. *** denotes rejection of the null hypothesis of a unit-root at the 5% level. I(0) denotes trend-stationarity and I(1) denotes a unit-root process. The following critical values were used: 2.56 for t_0^* and t_λ^*, and 3.00 for t_1^* (Table 1 in HLT); -3.19 for ADF-GLS (Table I in Elliot et al., 1996); and a set of critical values that depend on the break fraction (between -3.45 and -3.09) for ADF-GLS-break (Table 1 in Carrion-i-Silvestre et al., 2008).trend-stationarity and I(1) denotes a unit-root process. The following critical values were used: 2.56 for t_0^* and t_λ^*, and 3.00 for t_1^* (Table 1 in HLT); -3.19 for ADF-GLS (Table I in Elliot et al., 1996); and a set of critical values that depend on the break fraction (between -3.45 and -3.09) for ADF-GLS-break (Table 1 in Carrion-i-Silvestre et al., 2008).

The next step consists of testing for a unit-root. In every case that the result of the t_λ^* test points to a structural change, the appropriate unit-root test to use is the ADF-GLS test allowing for a change and denoted as ADF-GLS-break in Table 1. For males, except for Finland and Spain, this is the appropriate unit-root test to use, while for females it should be used only for Austria, Ireland, Italy, Japan, and Sweden. The ADF-GLS test not allowing for a break should be used in all the other cases. In almost all cases, the evidence supports the unit-root hypothesis. The few exceptions are the cases of males in Denmark and females in Japan and Sweden.

Finally, for the cases where in the first step a structural change was detected, the date when that change occurred is estimated in accordance with the conclusion of the appropriate unit-root test. Wherever a unit-root is present, the selected estimator of the structural change date corresponds to the date where the t_1^* statistic in (11) attains its maximum. In the few cases where a unit-root is rejected, the selected estimator of the structural change date corresponds to the date where the t_0^* statistic in (10) attains its maximum. These dates are presented in the last column of Table 1. Although not presented, for almost every case where a structural change is identified, the corresponding estimated magnitude of the change in the trend slope, given by γ in (8) or (9), is negative. The only exceptions are the cases of males and females in Japan and females in Sweden.

In summary, there is significant evidence supporting the presence of structural changes in the evolution of male mortality associated with a more accentuated decline in the overall mortality rate in recent years for almost every country considered. In contrast, evidence of structural changes in female mortality is found only for a few countries. It is interesting to note that in the case of the European countries these structural changes have all taken place in the second half of the sample, that is, after the mid-1970s.

The evidence of a unit-root found in almost every case considered implies that the uncertainty regarding the future evolution of mortality will increase with the forecast horizon. If instead the alternative of stationarity was found, confidence bands would not grow with the forecast horizon. We illustrate this in the next section.
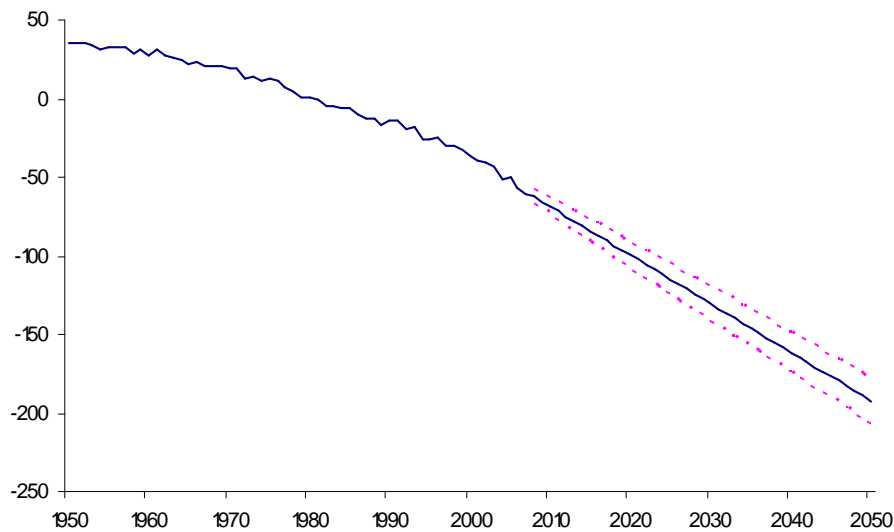
2.4.3. Forecasting Portuguese male mortality

To illustrate the impact of the results of the unit-root and structural change tests obtained in the previous sub-section in terms of mortality and life expectancy forecasts, we consider as an example the case of the Portuguese male population. We start by analysing the fit of the Poisson regression model by plotting the standardised deviance residuals against calendar year, age and year of birth (Figure 2). The overall pattern of well balanced positive and negative residuals and the absence of trends indicate a good adjustment of the model in terms of capturing period and age effects. There is only a slight ripple effect in the plot against year of birth, indicating eventual cohort effects that are not contemplated by this model.

According to the results of the tests presented in Table 1, the adequate forecasting model to consider is an ARIMA(p,1,q) allowing for a change in the drift in 1996, that is, an ARMA(p,q) model for the first-differences of k_t including as deterministic regressors a constant term and a step dummy variable as in equation (9). The orders p and q of this model are identified by analysing the residual autocorrelation functions, and using the Ljung-Box Q-tests and the Akaike and Schwarz information criteria. The final estimated model is the following ARIMA(0,1,1) model allowing for the structural change in 1996:

$$k_t - k_{t-1} = -1.37 - 1.72 DU_t^{1996} + u_t, \text{ with } u_t = \varepsilon_t - 0.53\varepsilon_{t-1}, \qquad (17)$$

where $DU_t^{1996} = 1$ if $t > 1996$ and $DU_t^{1996} = 0$ if $t \leq 1996$, and ε_t is a white noise. More detailed estimation results for this model appear in the first column of Table 2. Using this model we obtain forecasts of the mortality index k_t for the period 2008-2050. These are presented in Figure 3 together with the corresponding 95% confidence bands. The computed confidence intervals only account for the variability coming from the fitted ARIMA model: parameter estimation errors and future shocks. Other sources of variability could be allowed for by using bootstrap methods as in Brouhns, Denuit, and Van Keilegom (2005), Koissi et al. (2006), and Renshaw and Haberman (2008). As expected, the model predicts a decline in the mortality index at a pace that is consistent with the more pronounced decline in mortality in the final years of the 20th century and

Figure 3. Portuguese males mortality index (1950-2007) and forecasts from an
ARIMA(0,1,1) with a break in drift in 1996 (2008-2050)

beginning of the 21st century. The fact that the confidence bands widen is a direct consequence of the unit-root non-stationarity of this model.

To better understand the impact of allowing for a structural change in the mortality index by following the approach proposed in this paper, we also considered models estimated with different specifications of the structural change and the unit-root. Estimation results for these models also appear in Table 2. In Figure 6 we plot the resulting point forecasts for the period 2008-2050 for some of these models. The optimal model forecasts the greater decline in the mortality index.

We begin by analysing the random walk with drift model, which was found to be the optimal ARIMA model for k_t in the original work of Lee and Carter for the US population and in many other applications of their method. This model appears in the second column of Table 2. However, according to the residual autocorrelation functions and the Ljung-Box Q-statistics, this model does not fit well the Portuguese mortality index. The unit-root model without allowing for a structural change that best fits the mortality index series is an ARIMA(1,1,0) model, which appears in the third column of Table 2. An ARIMA(0,1,1) model was also reasonable but did not fit the model as well. The point forecasts of these models were very similar but with the simple random walk with drift model generating wider confidence bands.

Table 2. *Estimated forecasting models for the Portuguese male mortality index*

	(1) ARIMA(0,1,1) break in 1996	(2) ARIMA(0,1,0) no break	(3) ARIMA(1,1,0) no break	(4) ARMA(1,0) break in
Constant	-1.37** (0.16)	-1.70** (0.35)	-1.69** (0.23)	38.57** (2.90)
DU_t^{1996}	-1.72** (0.38)			
t				-0.92** (0.16)
DT_t^{1973}				-1.12** (0.21)
$AR(1)$			-0.40** (0.13)	0.59** (0.13)
$MA(1)$	-0.53** (0.12)			
Log-Likelihood	-121.65	-132.65	-127.86	-125.74
AIC	4.45	4.77	4.64	4.63
SBC	4.56	4.81	4.71	4.78
Q(4) p-value	0.509	0.011	0.840	0.113
No. Obs.	56	56	56	56
Forecasts of life expectancy at birth				
2030	80.9 [80.2 , 81.7]	78.7 [76.8 , 80.5]	78.7 [77.4 , 79.9]	78.8 [78.2 , 79.2]
2050	84.8 [84.0 , 85.6]	81.1 [78.7 , 83.3]	81.1 [79.5 , 82.6]	81.6 [81.1 , 82.1]
Forecasts of life expectancy at age 65				
2030	20.0 [19.5 , 20.5]	18.5 [17.3 , 19.7]	18.5 [17.7 , 19.3]	18.5 [18.2 , 18.9]
2050	22.7 [22.1 , 23.3]	20.1 [18.5 , 21.6]	20.1 [19.1 , 21.1]	20.5 [20.1 , 20.8]

The corresponding forecasts from this model appear in Figure 4. The first thing to note is that, in this model, the forecasted rate of decline in mortality is, by construction, basically given by the average rate of decline during the whole estimation period. Therefore, the more accentuated decline in mortality since the end of the 20th

century until the end of the estimation sample is not translated into the future projections. Consequently, the projected decline in mortality is much less when compared with the optimal model, which allowed for a structural change in 1996. Another important difference is that the confidence bands for the projected mortality index are wider for this model. This is a consequence of the poorer fit of this model relative to the optimal structural change model.
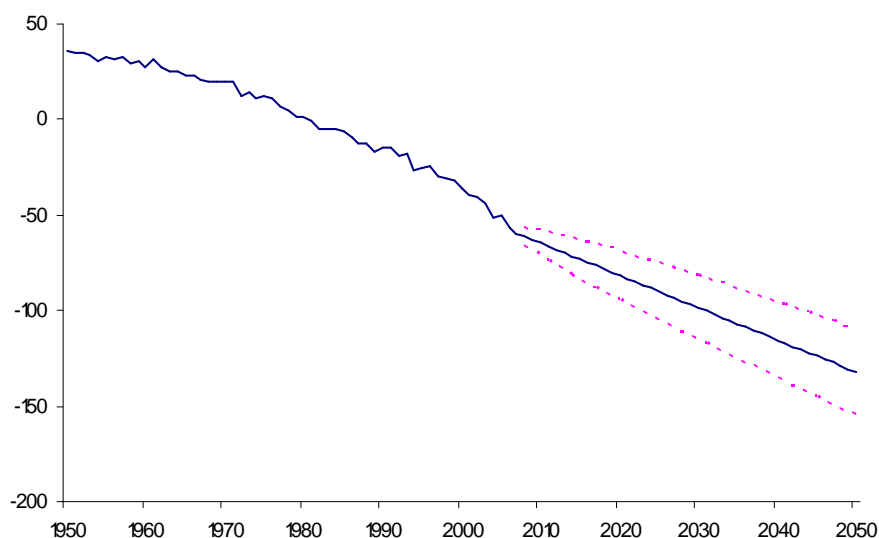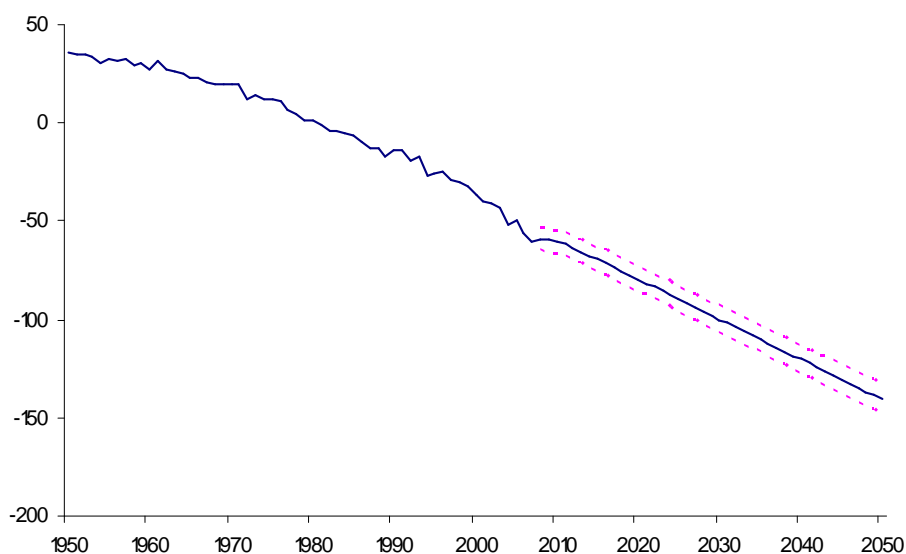

Figure 4. Portuguese males mortality index (1950-2007) and forecasts from an ARIMA(1,1,0) with no break (2008-2050)

We also considered trend-stationary models. When no break in trend was allowed for, the analysis of the residuals suggested an ARMA(1,1) model. However, this model resulted in an autoregressive root very close to 1, which basically confirms the previously obtained unit-root tests results. Because the model also includes a linear trend, the estimated near unit-root would, in fact, lead to a quadratic trend. We conclude that this model is not able to adequately describe the mortality index.

Finally, we have considered a trend-stationary model allowing for a change in the trend slope as in equation (8). According to the t_0^* test statistic, the estimated break date is in 1973. The best model in this case is an ARMA(1,0) including as deterministic regressors a constant term, a linear trend, and a change dummy defined as $DT_t^{1973} = t - 1973$ if $t > 1973$ and $DT_t^{1973} = 0$ if $t \leq 1973$. Estimation results are presented in the

fourth column of Table 2. The corresponding mortality index forecasts are graphed in Figure 5. The forecasted decline in mortality is also not as great as in the optimal forecasting model. Regarding the confidence bands, as expected, these do not grow as the forecasting horizon grows, since this model assumes stationarity around a broken trend. Thus, for longer horizons, this model will produce forecast intervals with the least amplitude, but obviously underestimating future uncertainty, given the results of the sequential testing procedure that led to the optimal forecasting model.

Figure 5. Portuguese males mortality index (1950-2007) and forecasts from an ARMA(1,0) with a break in trend slope in 1973 (2008-2050)

The mortality index forecasts are also used to produce forecasts of the age-specific forces of mortality $\mu_x(t)$ using the previously estimated α_x and β_x in equation (3). From these, we obtain forecasts for life expectancy at birth and at age 65 for each model (see Brouhns et al., 2005). These appear in Table 2 (see also Figure 7). In accordance with the forecasts obtained for the mortality index, the optimal model with a unit-root and a change in 1996 gives the highest projected gains in life expectancy during the next four decades. Life expectancy at birth is expected to increase by almost 9 years and life expectancy at age 65 by 6 years.
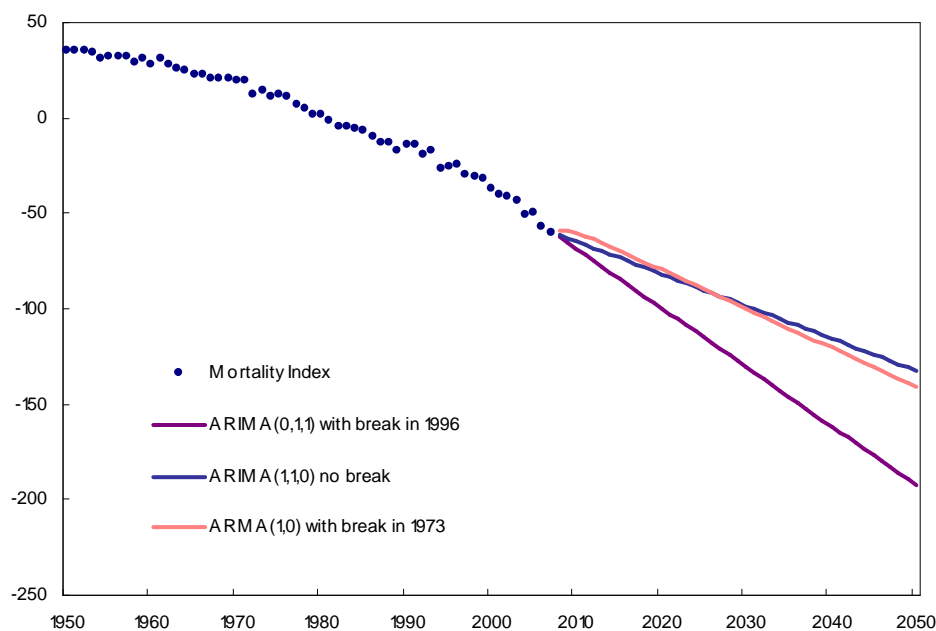
Figure 6. Portuguese males mortality index (1940-2007) and forecasts for several models (2008-2050)

The predicted value for life expectancy at birth in 2050 is 84.8 years with a 95% confidence interval of 84.0 to 85.6 years. Life expectancy at age 65 is predicted to grow from 16.7 years in 2007 to 22.7 years in 2050 with a ± 0.6 years 95% confidence interval. Also, in agreement with the results obtained regarding the confidence bands for the mortality forecasts, the amplitude of the confidence bands for the 2050 life expectancy forecasts is substantially smaller for the optimal model than for the unit-root models without allowing for a structural change, but larger when compared with the trend-stationary model with a break.
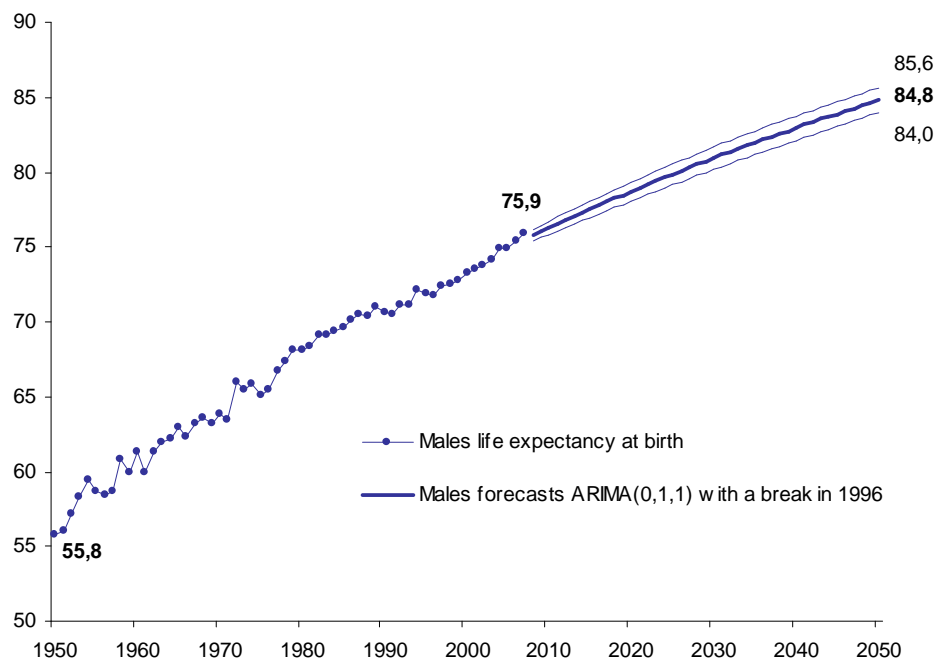
Figure 7. Portuguese males life expectancy at birth (1950-2007) and forecasts from an ARIMA(0,1,1) with a break in drift in 1996 (2008-2050).

2.5.   Discussion of Results and Conclusions

In the previous sections, we have considered the problem of forecasting future mortality and life expectancy in the possible presence of a structural change in the context of the Lee-Carter model. The proposed approach avoids the problems of other unit-root tests and is valid even in the presence of a structural change in the trend slope of the mortality index. Specifically, the results of tests for a change in the trend and for the presence of a unit-root are used to identify the most efficient ARIMA model for the mortality index that should be used to carry out the projections.

We apply the proposed procedure to post-1950s time-series mortality data for 18 developed countries and find significant evidence of structural changes in the rate of decline in the overall male mortality rate for almost every country considered. In contrast, female mortality decline has remained stable for the majority of the countries. These findings are consistent with results appearing in some recent studies.

Booth et al. (2006) compare the original Lee-Carter method with the variant proposed by Booth et al. (2002) to determine the optimal fitting period, and with the Lee and Miller (2001) approach which sets the starting period in 1950. Using data by sex for 10 developed countries (Australia, Canada, Denmark, England and Wales, Finland, France, Italy, Norway, Sweden, and Switzerland), the authors find that these variants are more accurate and that the Booth et al. (2002) variant had the best performance for 15 of the 20 populations, with optimal fitting periods starting in the late sixties or during the seventies. Since the fitting period in the study ends in 1985, the authors could not identify changes in the linear pattern after this year. The authors also find that the average overall absolute forecast error in the period from 1986 till the end of the century for the female populations is smallest for the fitting period starting in 1950, while for males is smallest for shorter fitting periods. Shang et al. (2010) extend the results of Booth et al. (2006) and include more recent data till 2004 for 14 countries, identifying consistent linear mortality declines for most countries starting in the late 1960s, during the 1970s and 1980s.

Carter and Prskawetz (2001) also found deviations from linearity in the data for Austria from 1947 to 1999. They find a substantial difference in life expectancy forecasts, for both males and females, when using the 1976-1999 sub-sample, which generates higher life expectancies when compared with the forecast based on the full sample. However, the authors refrain from identifying a date for the structural change. Renshaw and Haberman (2003a) while applying the Lee-Carter model to England and Wales for 1950-1998, found that the estimated k_t profile appears linear for females but seems to depart from linearity for males. For those, the pattern in the estimated k_t essentially comprises two linear segments, hinged in the mid-1970s. For females, however, the pattern in the estimated k_t is essentially linear throughout the time span concerned. These findings were consistent with a preceding exploratory analysis of trends in crude death rates developed by the authors, which showed that for males there was a pronounced increase in the rate of improvement in mortality, stemming from the 1970s, in most age bands. This is a feature that was also noted by Wilmoth (2000) and Valin and Meslé (2004) for Western countries. In a study of Belgian male mortality from 1948 to 2001, Denuit and Goderniaux (2005) also consider the possibility of a

change in the trend of mortality decline. Using an ad-hoc rule based on the R^2 statistic, they actually find that the optimal starting point to consider is the year of 1970.

The fact that a structural change may have occurred earlier in the case of the US and only later for most European countries is also brought up in a study of long-term mortality forecasters by Waldron (2005). The author remarks that a period of rapid mortality improvement started in the US in 1968, but only in 1982 for most of Europe. Wilmoth (2000) highlights the rapid decline in cardiovascular diseases (CVD) in the US since 1968 explained by a number of factors: a decrease in adult smoking, reduced blood pressure levels, more control of hypertension, a decrease in the consumption of saturated fats and cholesterol, medical advances, and improved medical care, among others. The author also points the early decline in cancer mortality that occurred in Japan since the 1960s. Cutler, Deaton and Lleras-Muney (2006) also refer that the decline in CVD in the US in 1960s was also accompanied by an important reduction in infant mortality due to improved neonatal medical care for low weight newborns. It is also pointed out that US females started smoking later than men, and were slower to quit. In some European countries, female smoking was still rising.

We also show the importance of considering structural changes when making predictions. By studying the case of Portuguese male mortality, we illustrate that accounting for such structural changes in a forecasting model in fact leads to a major impact in mortality and life expectancy forecasts over the next decades. Of course, one thing that our approach is not able to do, as in the case of similar extrapolative methods, is to predict the occurrence of new structural changes in the future. That would require extending the demographic model by including explanatory variables capable of predicting major changes that might affect future mortality, such as advances in medicine, occurrence of new diseases, changes in lifestyles, or the improvement or degradation in socio-economic conditions. However, the use of such information also raises a number of difficult problems. An example of including the effects of economic conditions on mortality is considered by Hanewald (2009).

Our proposed sequential testing approach to detect unit-roots and structural changes was applied in the context of the typical Lee-Carter model. However, it should be possible to also apply it to other variants of this model that allow, for instance, more than one latent mortality index in order to capture dissimilar evolutions of mortality at

different ages as in Renshaw and Haberman (2003b). Another extension of the proposed approach would be to consider the possibility of more than one structural change occurring during the analyzed period. These possible extensions are left for future research.

ACKNOWLEDGMENTS

REFERENCES

Alho, J. (2000). Discussion. *North American Actuarial Journal*, 4(1), 91-93.

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

Booth, H, Hyndman, R.J., Tickle, L., and Jong, P. (2006). Lee-Carter Mortality Forecasting: a Multi-country Comparison of Variants and Extensions. *Demographic Research*, 15, 289 – 310.

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under Conditions of Variable Mortality Decline. *Population Studies*, 56, 325– 336.

Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. ADSRI Working Paper No.3.

Brouhns, N., Denuit, M. and Van Keilegom, I. (2005). Bootstrapping the Poisson Log-bilinear Model for Mortality Forecasting. *Scandinavian Actuarial Journal*, 3, 212 -224.

Brouhns, N., Denuit, M. and Vermunt, J.K. (2002). A Poisson Log-linear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics*, 31, 373-393.

Bongaarts, J. (2006), How long will we live?. *Population and Development Review*, 32(4), 605-628.

Carrion-i-Silvestre, J. L., Kim, D. and Perron, P. (2008). GLS-based Unit Root Tests with Multiple Structural Breaks both Under the Null and the Alternative Hypotheses. *Econometric Theory,* 25, 1754-1792.

Carter, L.R. and Prskawetz, A. (2001). Examining Structural Shifts in Mortality Using the Lee-Carter Method. Max Planck Institute for Demographic Research, WP 2001-007.

Chan, W.-S., S.-H. Li, and S.-H Cheung (2008). Testing Deterministic Versus Stochastic Trends in the Lee-Carter Mortality Indexes and Its Implications for Projecting Mortality Improvements at Advanced Ages. Paper presented at the Living to 100 and Beyond Symposium, Orlando, FL, January 7–9, 2008.

Chen, C. and Liu, L.-M. (1993). Forecasting Time Series with Outliers. *Journal of Forecasting*, 12, 13-35.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006), The Determinants of Mortality. *Journal of Economic Perspectives*, 20(3), 97–120.

Debón, A., Montes, F. and Puig, F. (2008). Modelling and Forecasting Mortality in Spain. *European Journal of Operational Research* 189, 624-637.

Denuit, M., and Goderniaux, A.-C. (2005). Closing and projecting lifetables using log-linear models. *Bulletin of the Swiss Association of Actuaries*, 1, 29-49.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time-series with a Unit-root. *Journal of the American Statistical Association*, 74, 427-431.

Elliot, G., Rothenberg, T.J. and Stock, J.H. (1996). Efficient Tests for an Autoregressive Unit-root. *Econometrica*, 64, 813-836.

Fígoli, M.G.B. (1998). Modelando e Projectando a Mortalidade no Brasil. *Revista Brasileira de Estudos de População*, 15(1), 95-113.

Gómez, V., and Maravall, A., (1997). Programs Tramo and Seats, Instructions for the User (Beta Version: November 1997).

Guillen, M. and Vidiella-i-Anguerra, A. (2005). Forecasting Spanish Natural Life Expectancy. *Risk Analysis*, 25(5), 1161-1170.

Hanewald, K. (2009). Mortality modeling: Lee-Carter and the macroeconomy, SFB 649 Discussion Papers, SFB649DP2009-008, Humboldt University, Berlin, Germany.

Harris, D., Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009). Testing for a Unit-root in the Presence of a Possible Break in Trend. *Econometric Theory*, 25, 1545-1588.

Harvey, D.I, Leybourne, S.J. and Taylor, A.M.R. (2009). Simple, Robust and Powerful Tests of the Changing Trend Hypothesis. *Econometric Theory*, 25, 995-1029.

Hendry, D.F., and Neale, A.J. (1991). A Monte Carlo Study of the Effects of Structural Changes on Tests for Unit-roots, in "Economic Structural Change, Analysis, and Forecasting", P. Hackl and A. H. Westlund (eds.), Springer-Verlag, 95-119.

Horiuchi, S. and Wilmoth, J. R. (1998). Deceleration in the Age Pattern of Mortality at Older Ages. *Demography*, 35(4), 391-412.

Human Mortality Database, HMD. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), Available at

www.mortality.org or www.humanmortality.de (data downloaded on 17 Oct. 2008).

Human Mortality Database, HMD (2007). Methods Protocol for the Human Mortality Database. J.R. Wilmoth, K. Andreev, D. Jdanov, and D.A. Glei with the assistance of C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, P. Vachon. Available at: www.mortality.org/Public/Docs/MethodsProtocol.pdf.

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

Kim, H., Fay, M., Feuer, E. and Midthune, D. (2000). Permutation tests for jointpoint regression with applications to cancer rates. *Statistics in Medicine*, 19, 335-351.

Koissi, M., Shapiro, A., Högnäs, G. (2006). Evaluating and extending the Lee–Carter model for mortality forecasting confidence interval. *Insurance: Mathematics & Economics*, 38(1), 1-20.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P.J. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3), 159–178.

Lee, R. (2000). The Lee-Carter Method for Forecasting Mortality, with various Extensions and Applications. *North American Actuarial Journal*, 4(1), 80-93.

Lee, R. and Carter, L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87(419), 659-675.

Lee, R. and Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38(4), 537-549.

Li, S.H. and Chan, W.S. (2005). Outlier Analysis and Mortality Forecasting: the United Kingdom and Scandinavian Countries. *Scandinavian Actuarial Journal*, 3, 187-211.

Li, S.H. and Chan, W.S., (2007). The Lee-Carter Model for Forecasting Mortality, Revisited. *North American Actuarial Journal*, 11, 68-89.

Lundstrom, H. and Qvist, J. (2004). Mortality Forecasting and Trend Shifts: an Application of the Lee-Carter Model to Swedish Mortality Data. *International Statistical Review*, 72(1), 37-50.

Monnier, A. (2004). L'Union européenne à l'heure de l'élargissement. *Population*, 59(2), 361-384.

Nunes, L.C., Kuan, C.-M. and Newbold, P. (1995). Spurious Change. *Econometric Theory*. 11(4), 736-749.

Nunes, L.C., Newbold, P. and Kuan, C.-M. (1996). Spurious Number of Changes. *Economics Letters*. 50(2), 175-178.

Perron, P. (1989). The Great Crash, The Oil Price Shock, and the Unit-root Hypothesis. *Econometrica*, 57(6), 1361-1401.

Perron, P. (1997). Further Evidence on Changing Trend Functions in Macroeconomic Variables. *Journal of Econometrics*, 80, 355-385.

Perron, P. (2006). Dealing with Structural Breaks. In "Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory", K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.

Perron, P. (2008). Econometrics of Structural Change. In "The New Palgrave Dictionary of Economics", 2nd ed, S. Durlauf and L. Blume (eds.), Palgrave Macmillan.

Perron, P., and Rodríguez, G. (2003). GLS Detrending, Efficient Unit-root Tests and Structural Change. *Journal of Econometrics*, 115, 1-27.

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

Renshaw, A.E. and Haberman, S (2003b) .Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics Economics*, 33(2), 255-272.

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

Shang, H. L., R. J. Hyndman, and H. Booth (2010). A comparison of ten principal component methods for forecasting mortality rates. Working paper 08/10. Department of Econometrics & Business Statistics, Monash University.

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

Stoto, M.A. (1983). The Accuracy of Population Projections. *Journal of the American Statistical Association*, 78(381), 13-20.

Tuljapurkar, S., N. Li and C. Boe (2000). A Universal Pattern of Mortality change in the G7 Countries. *Nature*, 405, 789-792.

Vallin, J. and Meslé, F. (2004). Convergences and divergences in mortality. A new approach to health transition. *Demographic Research*, Special Collection 2, article 2. Downloadable at www.demographic-research.org.

Waldron, H. (2005), Literature Review of Long-Term Mortality Projections. *Social Security Bulletin*, 66(1), 16-30.

White, M. W. (2002). Longevity Advances in High-Income Countries, 1955-96. *Population and Development Review*, 28(1), 59-76.

Wilmoth, J.R. (1993). Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality change. Technical Report, University of California, Berkeley. U.S.A.

Wilmoth, J.R. (1996). Mortality Projections for Japan: A comparison of four methods. In "Health and mortality among elderly populations", G. Caselli and A. Lopez (eds.), Oxford University Press, New York, 266-287.

Wilmoth, J. R. (1998). Is the pace of Japanese mortality decline converging toward international trends? *Population and Development Review*, 24(3), 593-600.

Wilmoth, J. R. (2000). Demography of longevity: past, present and future trends. *Experimental Gerontology*, 35, 1111-1129.

Yu, B., Huang, L., Tiwari, R.C., Feuer, E.J., and Johnson, K.A. (2009). Modelling population-based cancer survival trends by using join point models for grouped survival data. *Journal of the Royal Statistical Society Series A,* 172(2), 405-425.

Zivot, E. and Andrews, D.W.K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, 10(3), 251-270.

3. PATTERNS OF MORTALITY DECLINE IN PORTUGAL SINCE 1950

Abstract

In this paper, we show how a detailed descriptive analysis of the patterns of mortality decline can be used in the process of formulating and validating parsimonious mortality models. In particular, we study the characteristics of mortality behaviour in Portugal from 1950 to 2007 by exploring different descriptive analysis and visualizations of mortality data, and use the Lee-Carter model, as well as some of its extensions, in order to identify the major patterns in the evolution of male and female mortality over time for different ages. In our analysis of Portuguese data, we show that an extension of the mortality model adding a second order age-period term is able to provide a better fit to the data by capturing important characteristics of the data, in particular the non-linear decline in log-death rates for some ages over the period considered.

3.1.   Introduction

The pattern of mortality behaviour in Portugal, as in other developed countries, has changed dramatically throughout the twentieth century to this day. During this period, with the exception of the influenza epidemic crisis in 1918, one saw a general decline in the level of mortality, a dramatic reduction in infant mortality, the increase of survival at more advanced ages, and extraordinary gains in the life expectancy of the population. The ongoing mortality decline and increasing longevity has huge consequences for the individual as well as for society as a whole in questions regarding public sector pension provisions, saving for old age, health care provision, and changing family structures, among many others.

Understanding mortality dynamics over the age and time dimensions is crucial for any demographic analysis. Towards this end, it is important to look at the historical time series data from different perspectives and to use appropriate mortality models. Booth (2006) and Booth and Tickle (2008) provide excellent overviews of the recent development and state of the art in stochastic mortality modelling and forecasting methods. The most prominent method was initially proposed by Lee and Carter (1992) using U.S. mortality data. This model has seen several extensions (see, e.g., Lee and Miller, 2001; Brouhns, Denuit and Vermunt, 2002; or Renshaw and Haberman, 2006), and has been applied to model death rates in many countries, including the G7 countries (Tuljapurkar, Li and Boe, 2000), Brazil (Fígoli, 1998), Japan (Wilmoth, 1998), Austria (Carter and Prskawetz, 2001), Australia (Booth, Maindonald and Smith, 2002), and Spain (Debón, Montes and Puig, 2008). Moreover, variations of the Lee-Carter model have been employed to model and forecast other demographic variables such as fertility rates or migration flows (Girosi and King, 2008; Hardle and Mysickova, 2009).

The main objective of this paper is to better understand the characteristics of mortality behaviour in Portugal from 1950 to 2007. We explore different descriptive analysis and visualizations of mortality data and use the Lee-Carter model, as well as some of its extensions, in order to identify the major patterns in the evolution of male and female mortality over time for different ages.

We show how a detailed descriptive analysis of the patterns of mortality decline can be used in the process of formulating and validating parsimonious mortality models.

In our analysis of Portuguese data, a model with two age-period interaction terms is able to provide a good fit to the data and to improve upon the simple Lee-Carter model by capturing the non-linear decline in log-death rates for some ages over the period considered.

The paper is set out as follows. In the following section, we present the relevant data and conduct a descriptive analysis of the main trends in Portuguese mortality. Section 3 summarizes the Lee-Carter model and several of its extensions. The results of the application of the Poisson Lee-Carter model are presented and discussed in section 4. Section 5 analyzes the differences in estimation results obtained for several sub-periods. Estimation results for the extension of the model allowing for two factors are considered in section 6. A summary of the results obtained and concluding remarks are provided in the last section of the paper.

3.2.   Mortality in Portugal from 1950 to 2007

The demographic data used to study the dynamics of mortality decline in the Portuguese population are the size of the population under analysis and the number of deaths that occur in the population at different ages and points in time. These data allow calculating other indicators such as death rates and life expectancies which we analyse in this section. We use data from 1950 to 2007 obtained from the Human Mortality Database (HMD). This database contains information from official sources in different countries and provides them for scientific analyzes in a uniform format. Data contained in the HMD are known for their reliability and accuracy and offer a valuable resource for mortality research. In what concerns the quality of the data on the numbers of deaths, after 1960 these are reasonably accurate. Prior to this date, they exhibit patterns of age heaping at ages ending in zero and should be used with more caution (see Canudas-Romo, Coelho and Pina, 2008). Data on population has some more quality issues, which are evident prior to the 1981 Census. Population count data are actually unknown. Population is estimated by the Instituto Nacional de Estatística (INE) in-between census periods (which occur every ten years), taking into account recorded births and deaths, and estimated net migration. These data are available, from INE, by

single year of age up to an open interval 85 and older. The survivor ratio methods are used by the HMD to estimate the population aged 85 and older (for details see Wilmoth, Jdanov and Glei, 2007, and Canudas-Romo et al., 2008).
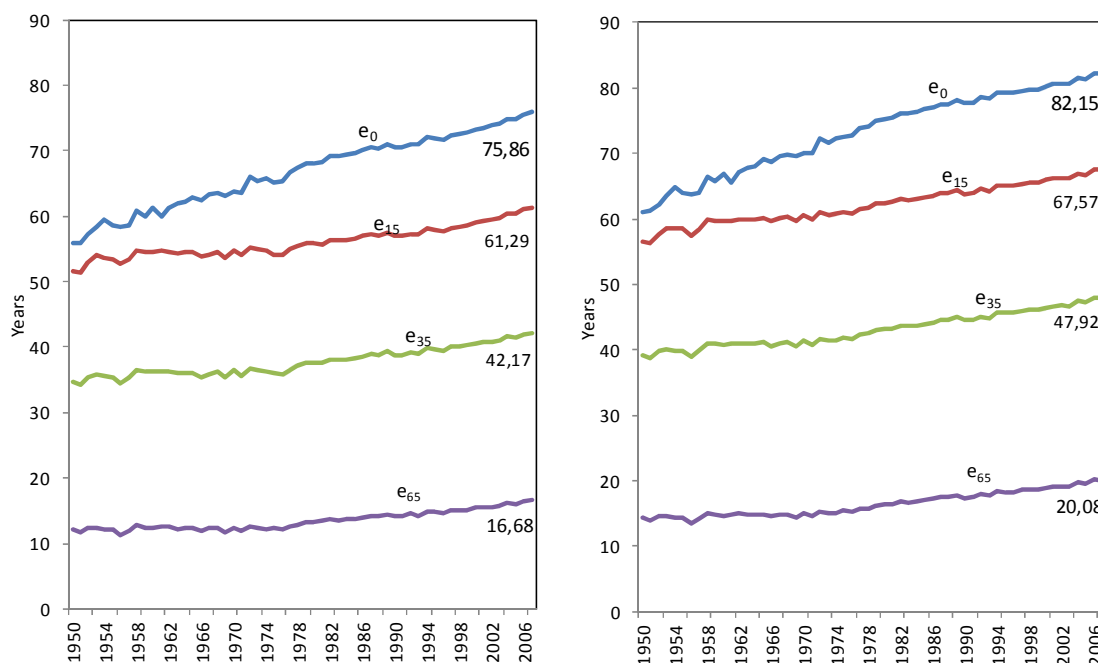

Figure 1. Life expectancy at birth, 15, 35 and 65 years old for male (left panel) and female (right panel) population, 1950 – 2007, Portugal.

Source: Human Mortality Database.

Between 1950 and 2007, life expectancy at birth in Portugal increased 20.1 years for men and 21.2 years for women, attaining, respectively, 75.9 and 82.2 years, which represents remarkable gains in terms of expected longevity for both sexes (see also Figure 1). Within the European Union (EU15), Portugal is the country whose life span has extended the most in the last six decades, although it was also the country with one of the lowest life expectancies at birth throughout the period (see Figure 2 and Table 1).
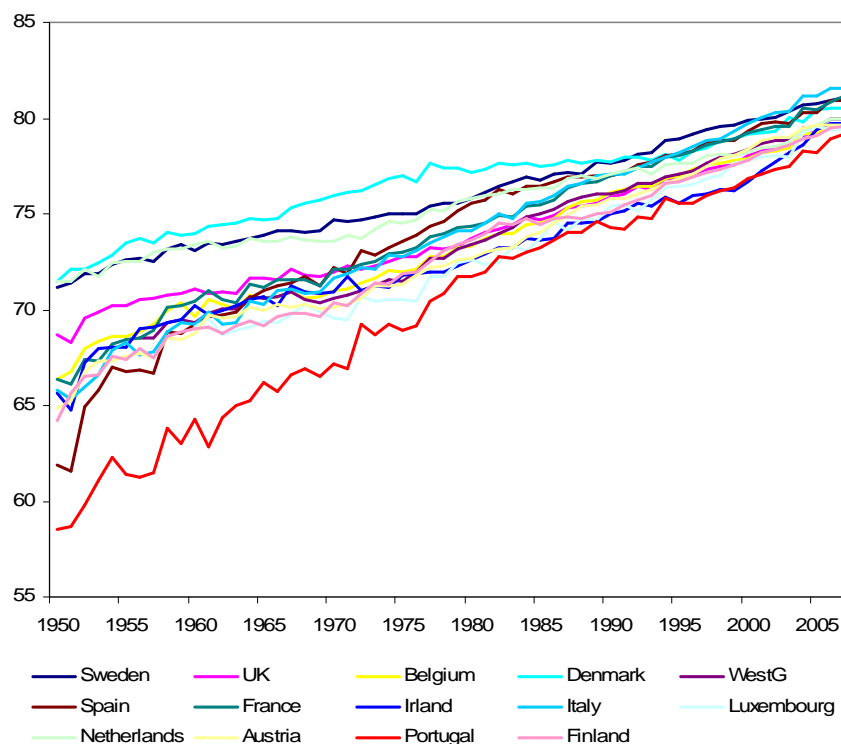
Figure 2. Life expectancy at birth for EU15 countries (except Greece), total population, 1950 - 2007.

Source: Human Mortality Database.

Table 1: *Life Expectancy at Birth for EU15 Countries (except Greece) for Male and Female Populations*

Country	Reference year	Life Expectancy at Birth	
		Males	**Females**
Austria	2005	76,7	82,2
Belgium	2006	76,5	82,2
Denmark	2007	76,1	78,4
Finland	2007	75,9	82,9
France	2007	77,4	84,4
Irland	2006	77,3	81,9
Italy	2006	78,6	84,1
Luxembourg	2006	76,7	81,8
Netherlands	2006	77,6	81,9
Portugal	**2007**	**75,9**	**82,2**
Spain	2006	77,6	84,1
Sweden	2007	78,9	83,0
UK	2006	77,2	81,5
WestG	2006	77,2	82,3

Source: Human Mortality Database.

These advances in life expectancy are perceptible at all ages. Female life expectancy has increased 11.1 years at age 15, 8.8 at 35 years old, 5.6 at 65, and 2.2 at 80. Male life expectancy increased 9.8, 7.5, 4.5 and 1.9 years at the same ages. The longevity gains over the whole period have been more favourable to women and the difference between life expectancy of men and women is higher in 2007 than the one in 1950. The annual evolution of this difference, however, varies with age. Until 1995, the gap in life expectancy at birth, 15 and 35 years old, steadily increased. From 1996 onwards this trend changed course due to the acceleration of male longevity improvements while women showed slower gains in life expectancy. At 65 years, the difference between life expectancy of men and women has continued to grow until the end of the 20th century showing an apparent tendency to stabilize in more recent years.
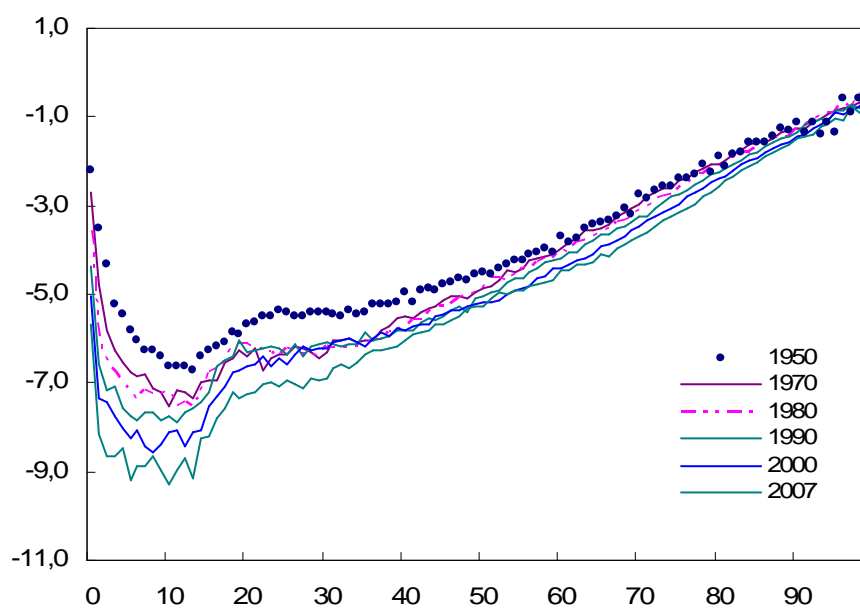
Figure 3. Male log death rates by age for selected years, Portugal.

The overall mortality age profiles presented in Figures 3 and 4 are consistent with the typical pattern found in most other developed countries, starting with a high infant mortality, followed by a rapid decline throughout the early childhood years. Mortality then increases to a local maximum between ages 15 and 30, especially among men. This feature is designated as the "accident hump" and reflects accident mortality for males and accident plus maternal mortality for the female population (see Heligman

and Pollard, 1980). The rates then steadily increase across older ages. Over time, the mortality age profile tends to move down reflecting the overall improvements in all age-specific deaths rates.
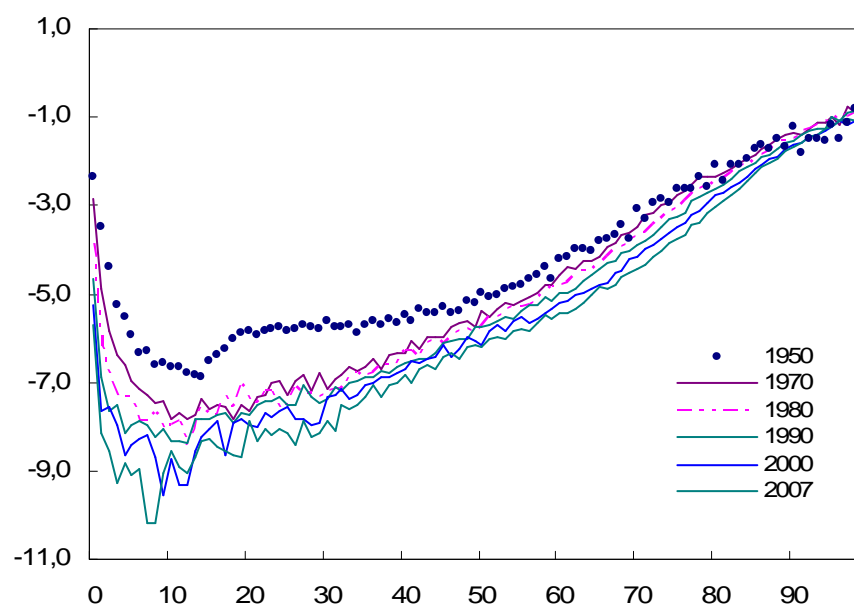

Figure 4. Female log death rates by age for selected years, Portugal.

Although age-specific death rates, over the period 1950 to 2007, declined at all ages for both men and women, the pace of the decline differs between ages, sexes and over periods of time as can be seen from Figure 5.[1] The most significant gains in mortality occurred at younger ages, i.e., in children under one year of age and children between 1 and 4 years of age, mainly as a result of reducing the number of deaths due to infectious diseases (see Rodrigues, Moreira and Fernandes, 2004). In 1950, 97 per 1000 female live births and 112 per 1000 male live births did not survive their first year of life. In 2007 these figures were reduced to about 3 deaths per 1000 live births. As can be seen from Figure 6, where the average annual rate of mortality decline over the whole period is plotted for different age groups, important improvements in mortality were also obtained by individuals at their teens and twenties. The improvements for the

---

[1] One should note that death rates for some age groups, particularly older ages, show somewhat erratic fluctuations and high variability due to small exposures to risk and inaccuracies in the register of ages in the official statistics.

elderly were more modest, but still significant, explaining the increase in life expectancy at age 65 in Figure 1.
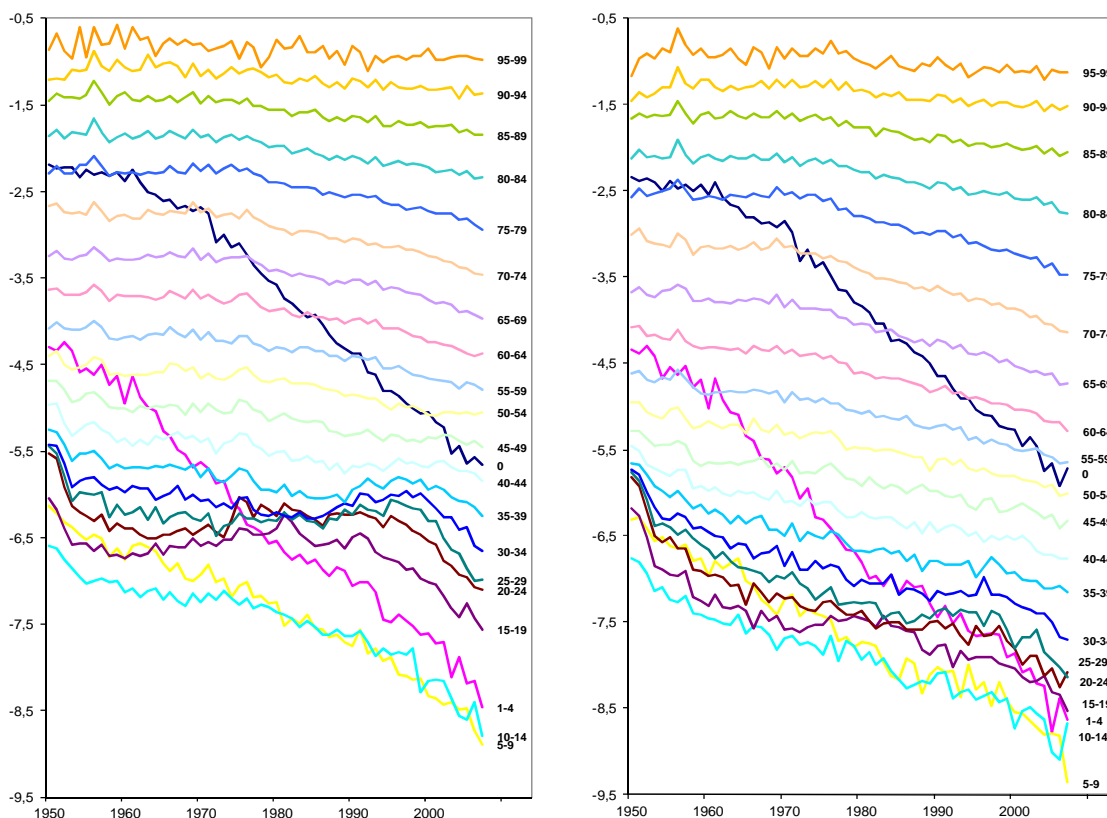
Figure 5. Age-groups log death rates for males (left panel) and females (right panel) over time, 1950 – 2007, Portugal.

The analysis of the average annual rates of mortality decline by age-groups for different sub-periods as depicted in Figures 7 and 8, or over time for different age groups (using nine-year rolling averages) as presented in Figures 9 and 10, shows some fairly different time paths. The 1950s were characterised by a general decline in mortality at all ages below 60 years old, with the greatest improvements for people then aged 15 to 29, followed by children aged 1 to 14 years old. The improvements in death rates from 1960 till the end of the century were relatively more modest at young adult ages, particularly for men that at some ages even experienced increases in mortality in some decades. During the 1970s, the mortality of children from 1 to 4 years old has declined at the fastest rates, closely followed by infant mortality. From 1980 to 2000,

infant mortality decreases at the fastest rate and gains in male mortality in the late teens and initial twenties begin to become evident. In the first years of the 21st century, it is noteworthy the rise in the rates of decline of mortality at almost all ages. This is particularly evident for men up to 40 years old. In this last period, the pace of mortality decline at older ages, i.e., over 65 years, also grew.
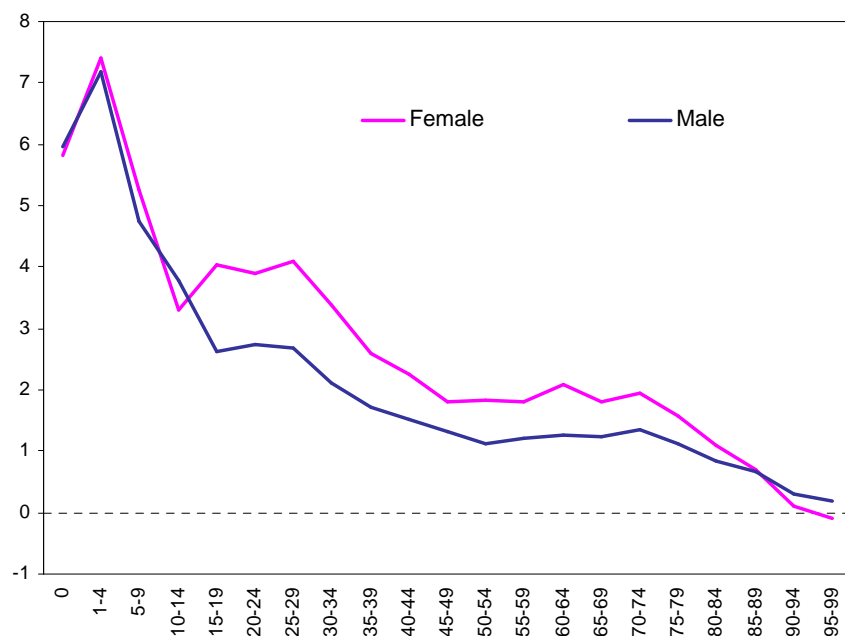
Figure 6. Age-groups average annual rate of mortality decline (in %), 1950 – 2007, Portugal.

The analysis of the graphical pattern of mortality improvement rates in Figure 7 further suggests that over time a specific group of male individuals might have experienced a non-declining or even increased mortality. In particular, from 1960 to 1980 this happened to young male adults between 15 to 29 years. In the 1980s, increases in death rates were experienced by males from 20 to 39 years. In the next decade, males aged between 30 and 49 years didn't benefit from improvements in mortality or even experienced a deterioration of death rates. The age groups showing lower improvements in death rates are apparently moving upwards along time.
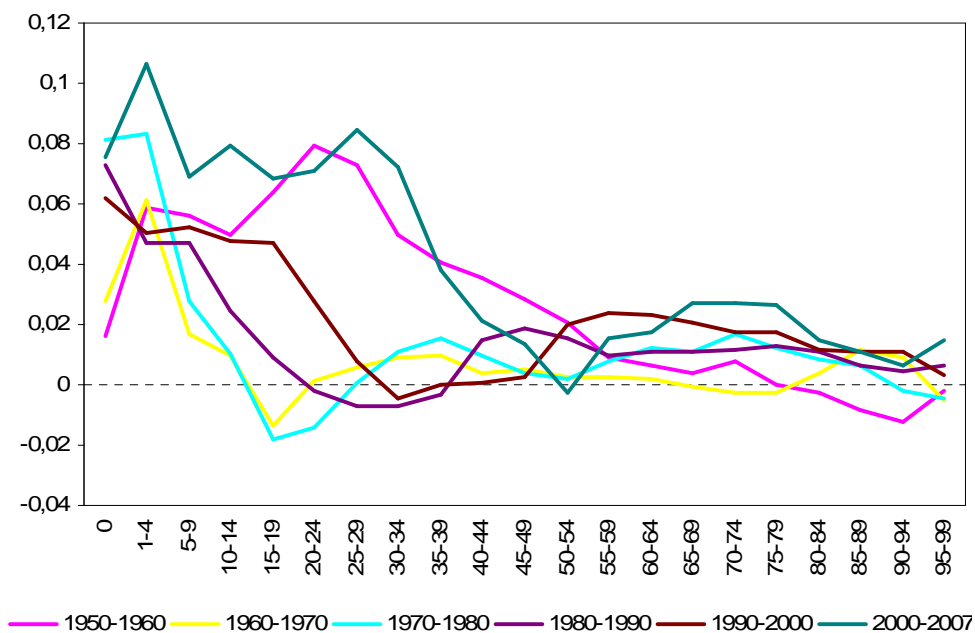
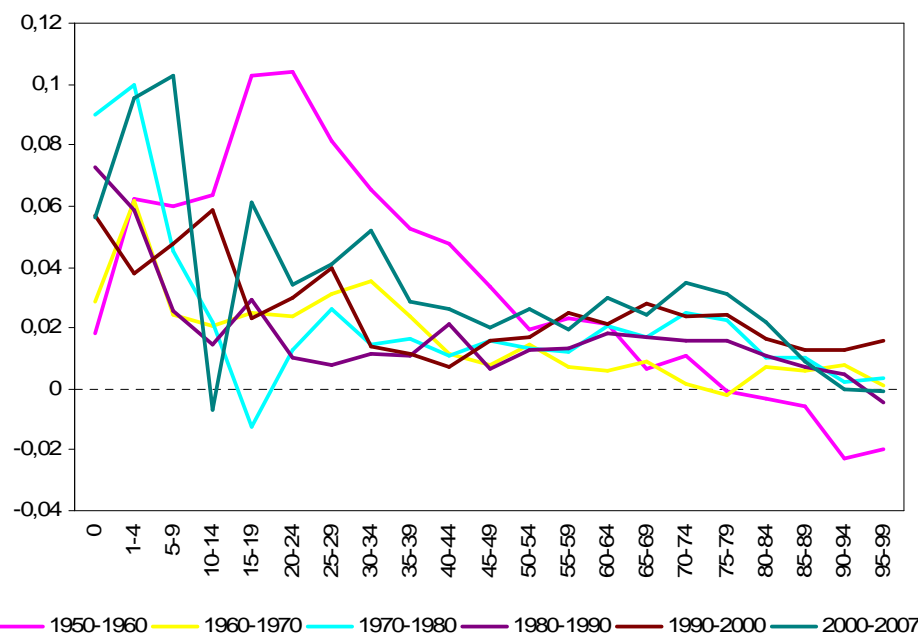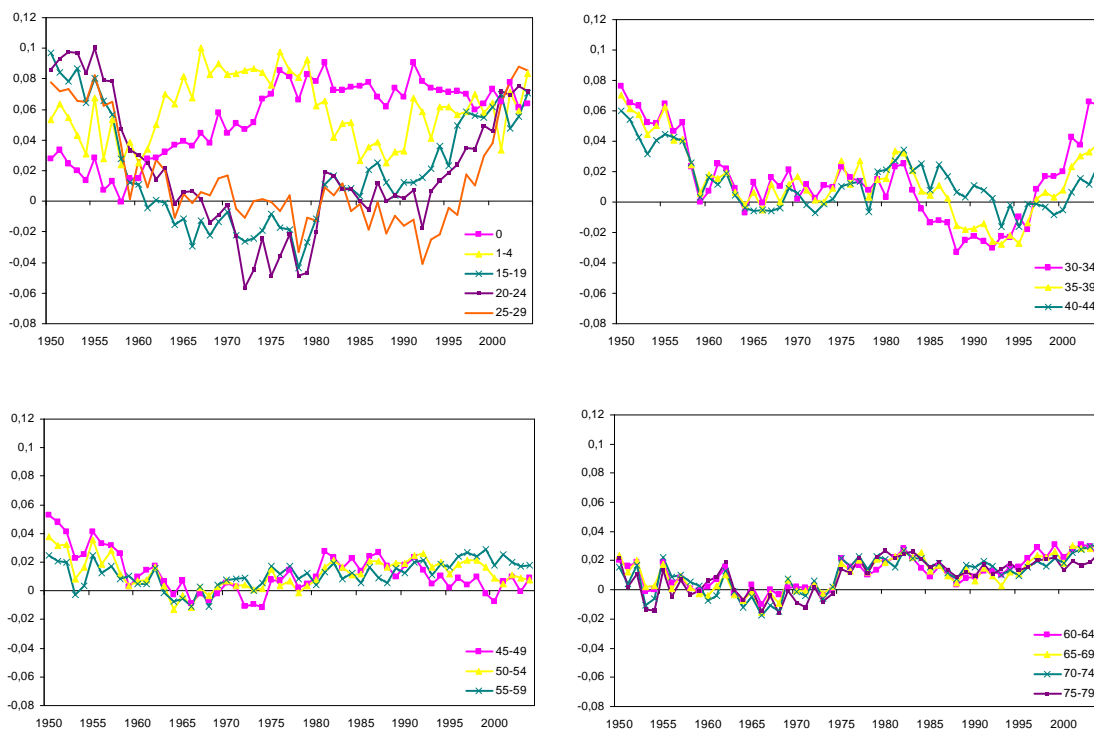*Figure 7.* Average annual rate of male mortality decline by age-groups for selected decades, Portugal.

Figure 8. Average annual rate of female mortality decline by age-groups for selected decades, Portugal.

However, this pattern becomes less evident with increasing ages. This particular behaviour suggests that specific generations of males may have been experiencing mortality conditions particularly difficult compared to others during some time. To better understand this, Figure 11 graphs the average annual improvement rates for the period 1950 to 2007 calculated by year of birth using data for ages 0 to 99. The graph suggests that people born around the 1950s, especially men, have experienced particular poor improvements in mortality when compared with adjacent generations. As discussed by Willets (1999, 2004), this particular behaviour may be due either to "period-specific" or "cohort-specific" influences.

Figure 9. Average annual rate of male mortality decline, nine-year rolling averages, for selected age groups, Portugal.

Figure 10. Average annual rate of female mortality decline, nine-year rolling averages, for selected age groups, Portugal

Table 2 shows the contributions of different age groups to the increase in life expectancy at birth. These were calculated using the Arriaga's (1984) method. The age group from 1 to 19 years old was the largest contributor to the increase in life expectancy at birth during the 1950s. After 1960 and till the end of the 1980s, the major contributions to the increase in life expectancy come from improvements in mortality at age 0. In the 1960s, the age group from 1 to 19 years was the second largest contributor to the increase in life expectancy, but in the following two decades this place is occupied by those aged 60 to 79 years. This older age group becomes the one with the larger contribution to the increase in life expectancy at birth in subsequent years.
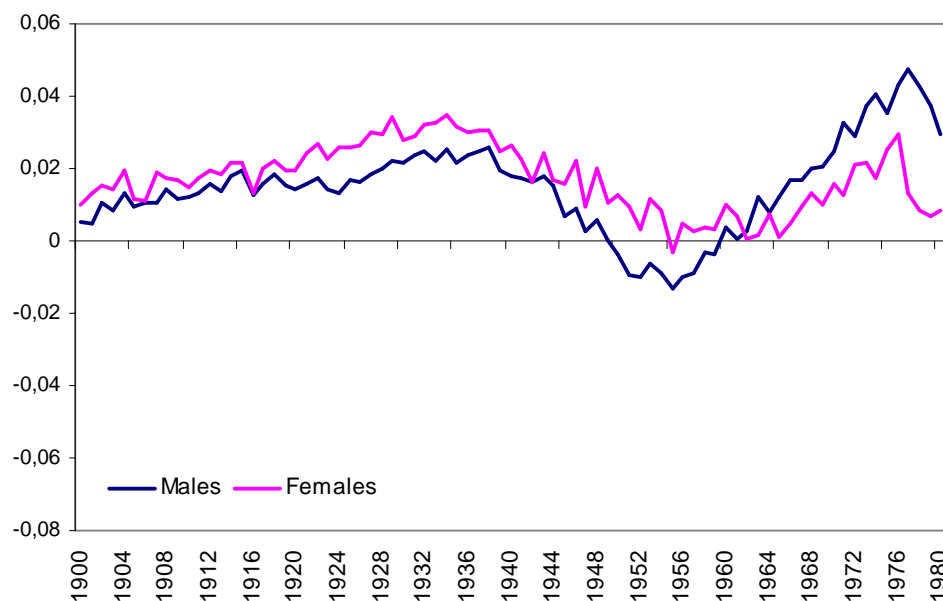
Figure 11. Average annual rate of mortality decline by year of birth, Portugal population data from 1950 to 2007, ages 0 to 99

Table 2. *Age-contribution to Change in Life Expectancy Over Time in Portugal, by Sex and Period (in Years)*

Age-groups	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000	2000-2007	Total Change 1950-2007
Females							
0	1,16	1,44	2,61	0,91	0,36	0,15	6,63
1-19	2,36	0,97	0,74	0,27	0,19	0,14	4,67
20-39	1,33	0,31	0,15	0,09	0,16	0,16	2,20
40-59	0,70	0,22	0,28	0,25	0,33	0,27	2,05
60-79	0,41	0,15	1,09	0,81	1,10	0,88	4,44
80+	-0,08	0,12	0,17	0,21	0,42	0,29	1,14
Total	**5,88**	**3,21**	**5,05**	**2,53**	**2,56**	**1,89**	**21,12**
e0 at start of period	61,03	66,91	70,12	75,17	77,7	80,26	
Males							
0	1,09	1,50	2,64	1,06	0,45	0,22	6,97
1-19	2,11	0,83	0,59	0,34	0,40	0,27	4,53
20-39	1,39	0,12	0,06	-0,09	0,15	0,72	2,34
40-59	0,74	0,12	0,19	0,48	0,43	0,25	2,22
60-79	0,21	-0,04	0,67	0,62	1,03	0,93	3,42
80+	-0,03	0,06	0,07	0,12	0,19	0,18	0,59
Total	**5,49**	**2,6**	**4,23**	**2,53**	**2,65**	**2,57**	**20,07**
e0 at start of period	55,79	61,28	63,88	68,11	70,64	73,29	

Analysis of the distribution of deaths by age, plotted in Figures 12 and 13, show the increasing concentration of deaths around the mode of the distribution, with this mode moving toward older ages through time. These two aspects are reflected,
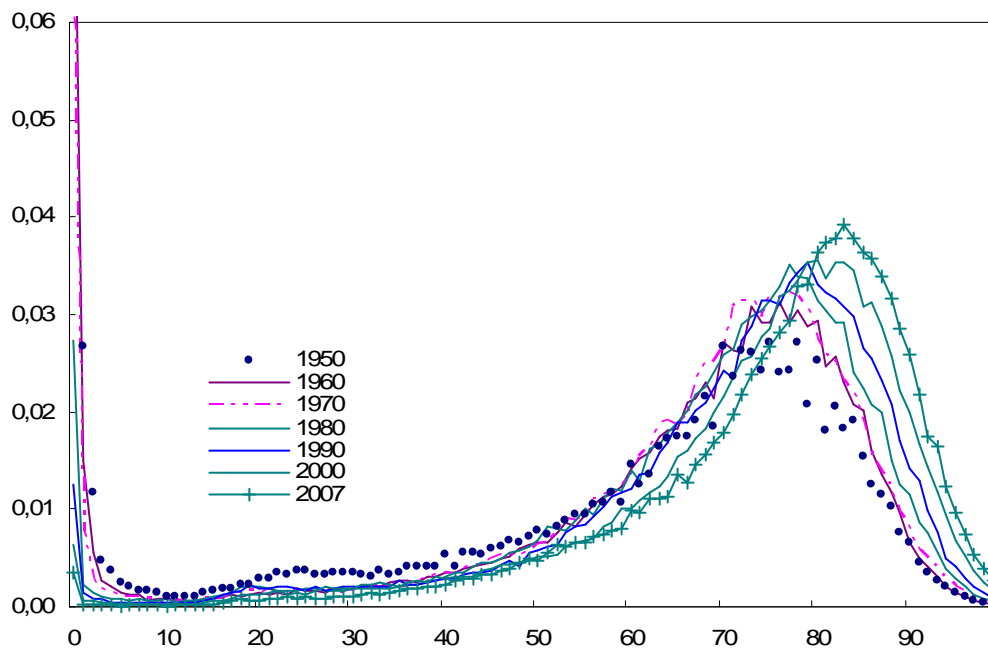

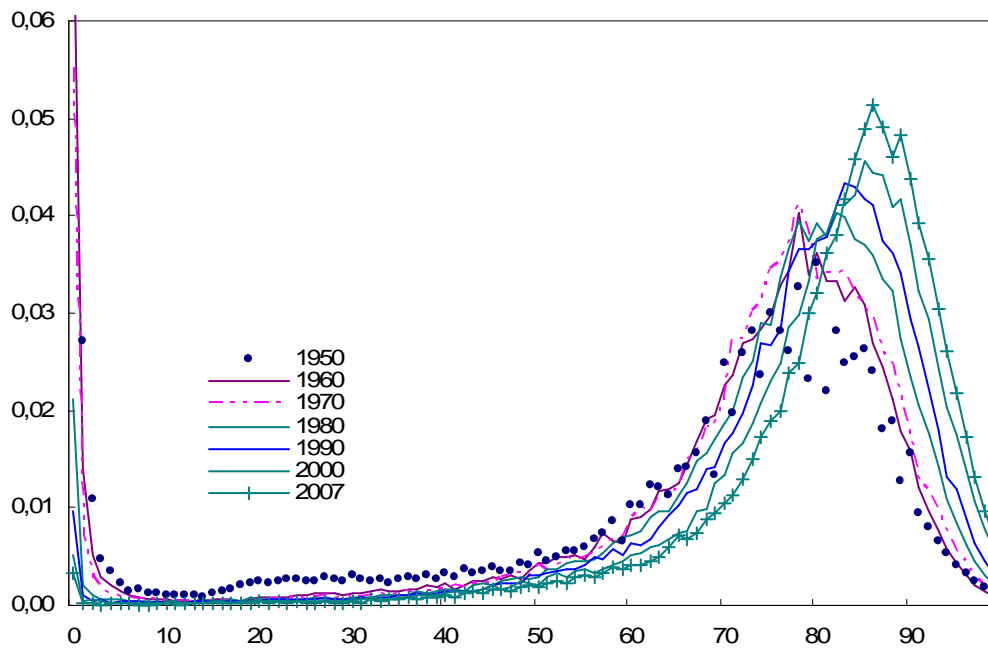*Figure 12.* Distribution of the male deaths over ages for selected years, Portugal

Figure 13. Distribution of the female deaths over ages for selected years, Portugal.

respectively, in the phenomena of "rectangularization" and "expansion" of the survival curve, plotted in Figures 14 and 15, which represents the probability of surviving from birth to various ages. In other words, individuals are surviving in greater proportions to older ages and at these ages, both men and women, can expect to live on average longer.
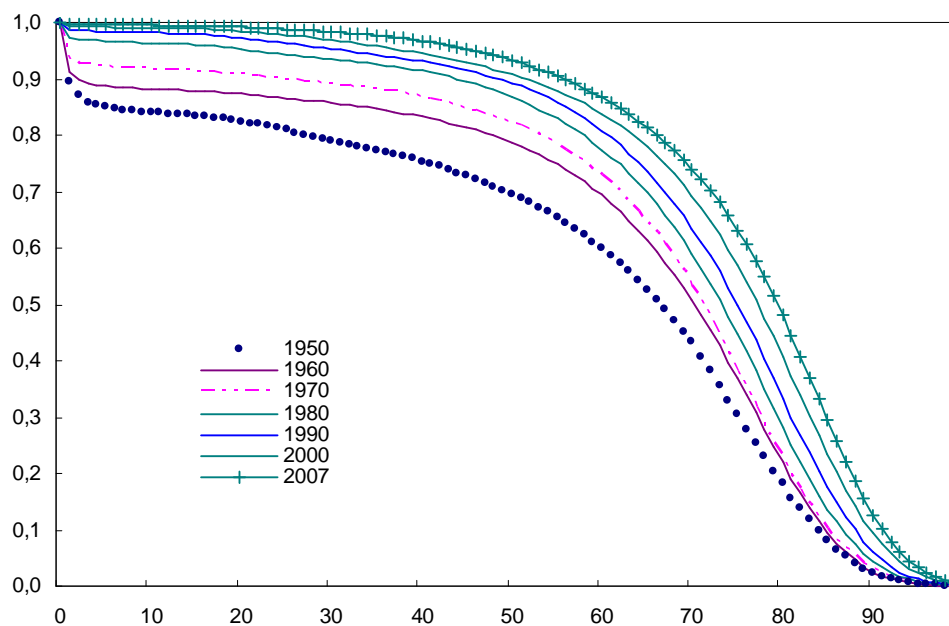

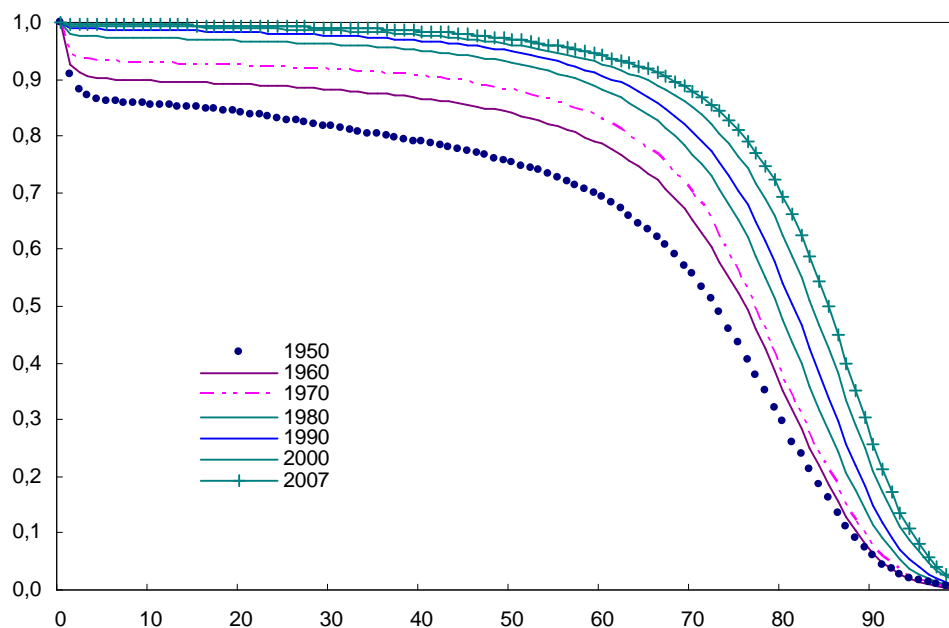*Figure 14.* Male survival curve for selected years, Portugal

Figure 15. Female survival curve for selected years, Portugal

3.3.   The Lee-Carter Method and Extensions

The most popular stochastic mortality method is the one proposed by Lee and Carter (1992). It combines a demographic model of age-specific death rates with statistical time-series forecasting methods. Let $\mu_x(t)$ denote the mortality force at age x during calendar year t and $D_{x,t}$ denote the number of deaths recorded at age x during year t from an exposure-to-risk $E_{x,t}$.[2] We assume that age-specific mortality rates are constant within bands of age and time, but allowed to vary from one band to the next. Specifically, given any integer age x and a calendar year t, it is assumed that

$$\mu_{x+y}(t+\tau) = \mu_x(t) \ \text{ for } \ 0 \le y, \tau < 1. \tag{1}$$

Under (1), estimates of $\mu_x(t)$, denoted as $\hat{\mu}_x(t)$, are given by the following ratio:

$$\hat{\mu}_x(t) = \frac{D_{x,t}}{E_{x,t}}. \tag{2}$$

The Lee-Carter (1992) model assumes that the observed force of mortality for any age x is driven by a common time-varying component, denoted by k_t, which is also referred to as the overall mortality index. More precisely, the following relation is assumed:

$$\ln \hat{\mu}_x(t) = \alpha_x + \beta_x k_t + \varepsilon_x(t) \tag{3}$$

where $\hat{\mu}_x(t)$ is given by (2), α_x are the age-specific parameters that affect the overall level of $\ln \hat{\mu}_x(t)$ over time, β_x are the age-specific parameters that characterize the sensitivity of $\ln \hat{\mu}_x(t)$ to changes in the mortality index k_t, and $\varepsilon_x(t)$ represent error terms capturing particular age-specific historical influences not explained by the model. The errors are assumed to have mean 0 and constant variance.

Lee and Carter (1992) propose a two-stage procedure to estimate the model given by equation (3). First, a least-squares solution to estimate α_x, β_x and k_t is found. Since the model is underdetermined, some normalization constraints are imposed to obtain a unique solution: the sum of the β_x over all ages equals one and the sum of the

---

[2] Estimates of the population exposed to the risk of death are based on population estimates, with a correction reflecting the timing of deaths (see Human Mortality Database).

k_t over time equals zero. As a consequence, the α_x will be equal to the average values of $\ln \hat{\mu}_x(t)$ over time. The solution can be found using the singular value decomposition. Since the model is written in terms of log mortality, the observed total number of deaths in each year will not equal the sum of the fitted deaths by age. To ensure this equality, the k_t are estimated a second time, taking the α_x and β_x estimates from the first step, such that for each year t, given the actual age distribution for the population, the implied number of deaths equals the observed number of deaths.

The homoskedasticity assumption on the error term $\varepsilon_x(t)$, implied by the singular value decomposition estimation of the parameters α_x, β_x and k_t, has been considered fairly unrealistic since the logarithm of the observed force of mortality is much more volatile at old ages, where the number of deaths and of those exposed to risk are relatively small (see Wilmoth, 1993, and Alho, 2000). To circumvent this problem, Brouhns *et al.* (2002) propose an alternative model that keeps the Lee-Carter log-bilinear form for the force of mortality but replaces the least-squares approach by a Poisson regression for the number of deaths. Specifically, they consider the following model describing the distribution of $D_{x,t}$:

$$D_{x,t} \sim Poisson\big(E_{x,t}\mu_x(t)\big) \tag{4}$$

with

$$\mu_x(t) = \exp(\alpha_x + \beta_x k_t), \tag{5}$$

where the parameters α_x, β_x and k_t keep the meaning originally attributed by the Lee-Carter model. These parameters can be estimated by maximum likelihood still subject to the model identification constraints that the sum of the β_x over all ages equals unity and the sum of the k_t over time equals zero. The Poisson likelihood is optimised adapting the iterative fitting method due to Goodman (1979), as described in Brouhns et al. (2002). Lee and Carter (1992) also show that, given the estimated parameters, the Box-Jenkins methodology can be used to forecast the mortality index k_t and, consequently, life expectancies.

Booth et al. (2002) and Renshaw and Haberman (2003) argue that the interaction between age and time in the demographic model can be captured better by adding a second log-bilinear term to equation (5), which would become

$$\mu_x(t) = \exp\left(\alpha_x + \beta_{x,1}k_{t,1} + \beta_{x,2}k_{t,2}\right). \tag{6}$$

The identification conditions of the model are now described by:

$$\sum_t k_{t,1} = \sum_t k_{t,2} = 0, \tag{7}$$

$$\sum_x \beta_{x,1} = \sum_x \beta_{x,2} = 1. \tag{8}$$

To estimate the parameters, the iterative procedure developed by Goodman (1979) is reformulated in order to include the additional parameters (for details see Renshaw and Haberman, 2003).

In the process of statistical modelling, residuals provide information regarding assumptions about error terms and the appropriateness of the model. They measure the departure of fitted values from actual values of the dependent variable, and can be used to detect model misspecification, outliers or observations with poor fit. In particular, visual inspection of the residuals can potentially indicate the nature of misspecifications and ways that it may be corrected, as well as provide a feel for the magnitude of the effect of the misspecification.

For linear models, a residual is easily defined as the difference between the actual and the fitted value. For the classical linear regression model, with normally distributed homoskedastic error, in large samples the residual has the desirable properties of being symmetrically distributed around zero with constant variance. For nonlinear models the very definition of a residual is not unique and several residuals have been proposed (see Cameron and Trivedi, 1998). For count data there is no one residual that has zero mean, constant variance and symmetric distribution. With a Poisson random component, deviance residuals have been considered the most appropriate to monitor the quality of the fit (Brounhs, Denuit and Keilegom, 2005). These are defined as:

$$r_{x,t}^D = sign\left(D_{x,t} - \hat{D}_{x,t}\right)\sqrt{2\left\{D_{x,t}\ln\left(\frac{D_{x,t}}{\hat{D}_{x,t}}\right) - \left(D_{x,t} - \hat{D}_{x,t}\right)\right\}} \tag{9}$$

where

$$\hat{D}_{x,t} = E_{x,t}\exp\left(\hat{\alpha}_x + \hat{\beta}_x\hat{k}_t\right) \tag{10}$$

Renshaw and Haberman (2006) also propose using the standardised deviance residuals:

$$sign\left(D_{x,t} - \hat{D}_{x,t}\right)\sqrt{\frac{2\left\{D_{x,t}\ln\left(\frac{D_{x,t}}{\hat{D}_{x,t}}\right) - \left(D_{x,t} - \hat{D}_{x,t}\right)\right\}}{Dev\left(D_{x,t},\hat{D}_{xt}\right)\Big/v}} \qquad (11)$$

where $Dev\left(D_{x,t},\hat{D}_{xt}\right)$ is the deviance statistic, a measure of goodness-of-fit of the model (see definition below), and v = (number of data cells – number of independent parameters) corresponds to the degrees-of-freedom.

A graphical analysis can be used to check that the residuals are independent and identically distributed (Brounhs et al., 2005; Koissi, Shapiro and Högnäs, 2006). A shaded contour plot of the residuals over ages and years can also be used to locate where the model cannot properly capture mortality trends and may reveal peculiar period shocks, cohort effects or specific age-time interaction effects not captured by the model (Booth et al., 2002; Koissi et al., 2006).

The deviance statistic $Dev\left(D_{x,t},\hat{D}_{xt}\right)$, as already mentioned, is a common measure of goodness-of-fit of the model and may be used as a measure of the overall performance of the model:

$$Dev\left(D_{x,t},\hat{D}_{x,t}\right) = \sum_{t}\sum_{x}\omega_{x,t}2\left\{D_{x,t}\ln\left(\frac{D_{x,t}}{\hat{D}_{x,t}}\right) - \left(D_{x,t} - \hat{D}_{x,t}\right)\right\}, \qquad (12)$$

with

$$\omega_{x,t} = \begin{cases} 1, & E_{x,t} > 0, \\ 0, & E_{x,t} = 0. \end{cases} \qquad (13)$$

The model performance may also be accessed by the percentages of total deviance explained by the model. Cameron and Windmeijer (1997) for Poisson regression models proposed a pseudo R^2 based on the decomposition of the deviance, denoted by R^2_{Dev},

$$R^2_{Dev} = \frac{\sum_t \sum_x \left\{ D_{x,t} \ln\left(\frac{\hat{D}_{x,t}}{\bar{D}} \right) - \left(D_{x,t} - \hat{D}_{x,t} \right) \right\}}{\sum_t \sum_x \left\{ D_{x,t} \ln\left(\frac{D_{x,t}}{\bar{D}} \right) \right\}} . \tag{14}$$

R^2_{Dev} can be interpreted as the relative reduction in deviance due to covariates in the model and shares the characteristics of the R^2 for the common linear regression. Proportions of the variance accounted for by the model (ratio of the variance of differences between the actual and fitted number of deaths to the variance of the actual ones) over the years at different ages are also computed by Brouhns et al. (2002) to evaluate model performance.

3.4. Results of the Poisson Lee-Carter Model applied to 1950 - 2007

In this section we present and discuss the estimation results of the Poisson extension of the Lee-Carter model, hereafter designated as PLC model, for the Portuguese data, separately for men and women, from 1950 to 2007. Figures 16, 17 and 18 plot the estimated $\hat{\alpha}_x$, $\hat{\beta}_x$, and $\hat{k}_t$, respectively. The $\hat{\alpha}_x$'s describe the general shape of the age-specific death rates, and correspond to the time-average of the logarithms of deaths rates that were presented in Figures 3 and 4. As expected, these are relatively high at infant and childhood ages, declining very rapidly with age to reach its minimum at around ages 10 to 12. The values of $\hat{\alpha}_x$ then increase to a local maximum around the
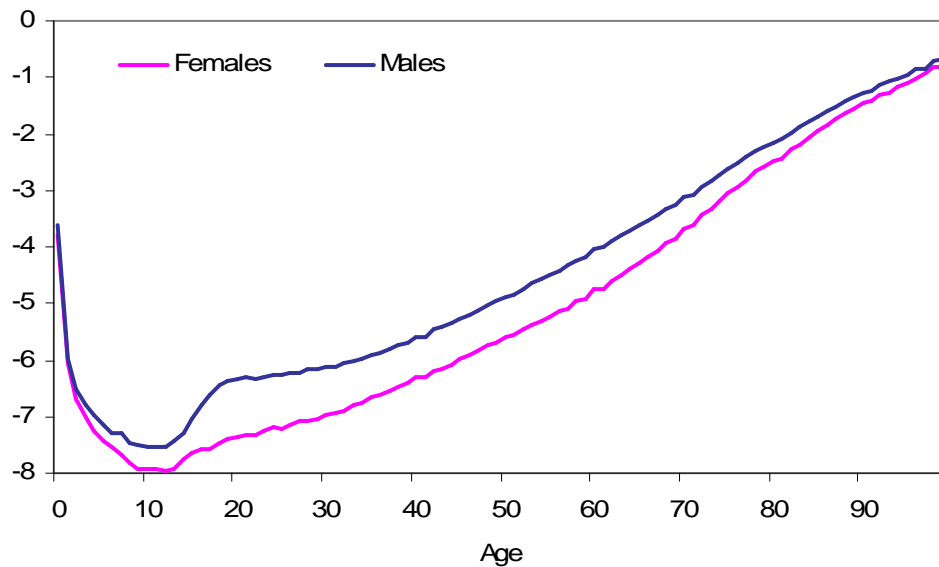
Figure 16. PLC model estimates: $\hat{\alpha}_x$ for ages from 0 to 99, 1950 – 2007, Portugal

late teens and early twenties, predominantly among men, corresponding to the "accident hump". Then, the $\hat{\alpha}_x$ values follow an unavoidable nearly linear increase into the middle and older ages, reflecting the higher mortality at more advanced ages.
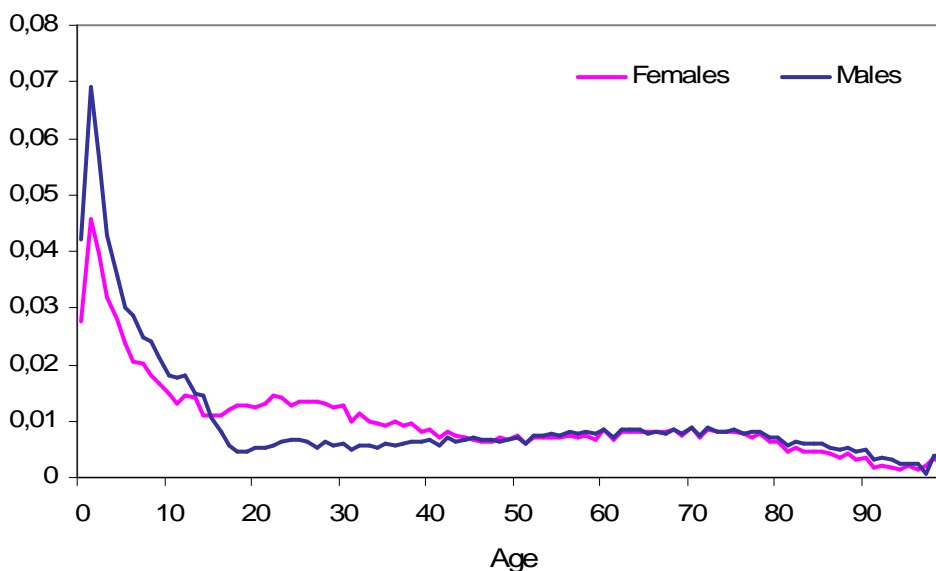

Figure 17. PLC model estimates: $\hat{\beta}_x$ for ages from 0 to 99, 1950 – 2007, Portugal

Figure 18 plots estimates of $\hat{k}_t$, for females and males. As shown, $\hat{k}_t$ declines roughly linearly from 1950 to 2007, reflecting the overall improvements in mortality over time. The downward trend in $\hat{k}_t$ is in general more pronounced for women.
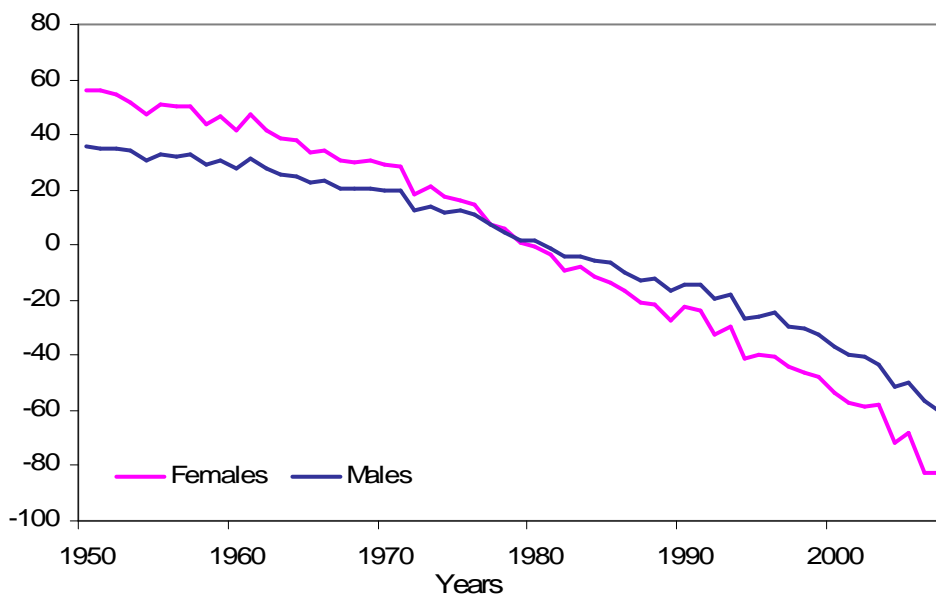

*Figure 18.* PLC model estimates: $\hat{k}_t$, 1950-2007, Portugal

Since all age-specific death rates are driven by this overall time index, they move up and down together, although not necessarily by the same amount. The estimated parameters $\hat{\beta}_x$ appearing in Figure 17 tell us at which ages do death rates decline more rapidly and which decline more slowly over time in response to the overall decline in $\hat{k}_t$. The estimates $\hat{\beta}_x$ and $\hat{k}_t$ are thus complementary, i.e., it is the value of the product of both that results in temporal variations in age-specific death rates. The larger values of $\hat{\beta}_x$ are attained at infant and younger ages, confirming the pattern of mortality decline by age-groups observed in Figure 6. The values of $\hat{\beta}_x$ decrease with age but remain positive. Significant declines in mortality have occurred also among females in their twenties, while males in late teens and twenties have had weak improvements in mortality.

The actual and estimated log-death rates for a selected set of ages are presented in Figure 19. In general, the PLC model captures the overall trends in mortality, with the estimated age-specific death rates showing a decreasing behaviour over time consistent with the observed age-specific death rates behaviour. The exceptions are those ages with a less regular behaviour where some difficulties of adjustment are evident (ages 20 and 30 for instance).

The standardized deviance residuals over ages and years are plotted in a heat-map presented in Figure 20. Negative residuals occur where the model overestimates death rates and positive residuals where the model underestimates the death rates. Systematic patterns are visible, in particular for men, showing a structure hardly consistent with the hypothesis of independence of the residuals. Among other things, there is an apparent cohort effect reflected in the salient diagonal with similar values. The deviance residuals for women are, in absolute value, lower (as can be confirmed by the different scales used in Figure 20) and although there is some clustering, there is no clear evidence of a cohort effect.

The percentages of total deviance explained by the model are high. The values of R^2_{Dev} are 95.5% for males and 95.6% for females, clearly conveying that the model provides a good overall fit to the data. However, the percentages of the total deviance of the number of deaths explained by the model at different ages shown in Figure 31, evidence the difficulties of the model for late teens and young adult ages, particularly among males.
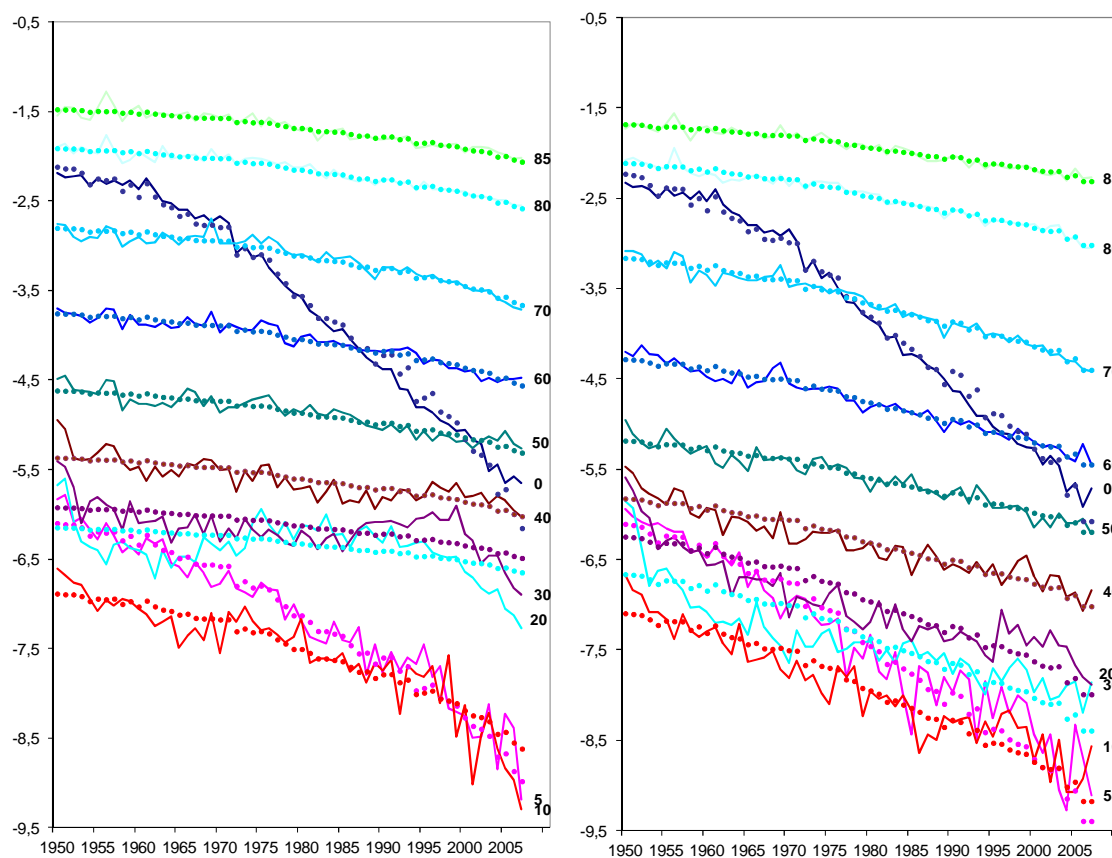
Figure 19. Actual (solid line) and fitted (dotted line) log death rates from PLC model for males (left panel) and females (right panel) for selected ages, Portugal

Summarizing the results, we find that the Poisson Lee-Carter model in general fits well the data and is able to capture the major patterns of mortality decline. However, there are some difficulties in fitting at some particular ages. Also, the presence of some clustering in the residuals may suggest the presence of some age-time interactions or a cohort effect not accounted for by this model.
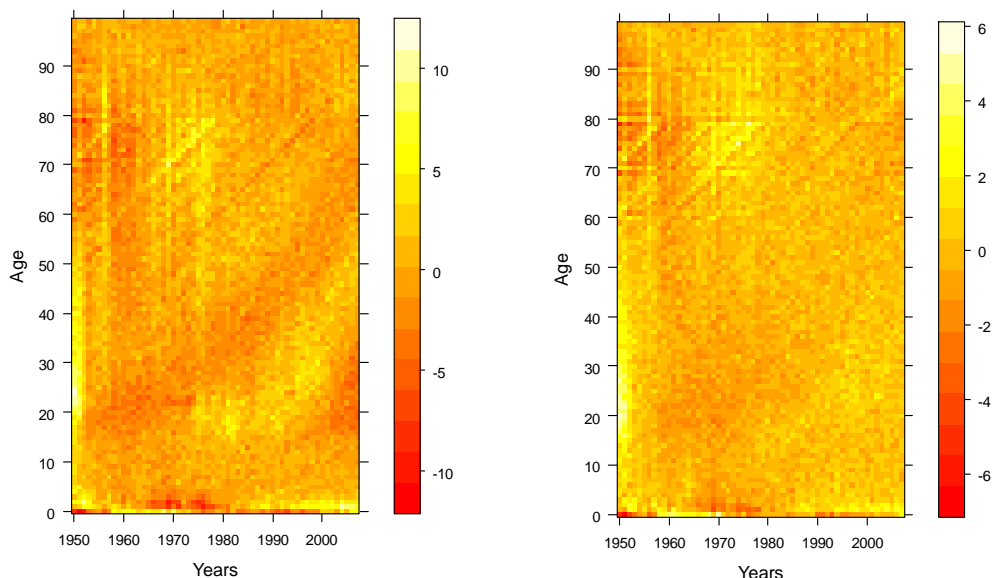
Figure 20. Standardized deviance residuals from PLC model for males (left panel) and females (right panel), 1950-2007, Portugal

3.5.   Results of the Poisson Lee-Carter Model applied to Sub-periods

As discussed in Section 2, an analysis of the age-specific paths of mortality decline over the period 1950 to 2007 shows rather different improvement patterns in different sub-periods. This cannot be captured by the PLC model estimated in the previous section which assumes that the age-specific patterns of mortality decline are all similar, as they are driven by the common overall mortality index k_t and the vector of parameters β_x is invariant over time. In fact several empirical studies find that the β_x parameters are not necessarily fixed in time, which may cause a lack of fit of the model for longer periods. An example is Lundstrom and Qvist (2004) who examine changing trends in the Swedish mortality decline during the twentieth century. Booth at al. 2002 partially addresses this problem when choosing the fitting period to be used for forecasting purposes. Also Lee and Miller (2001), Carter and Prskawetz (2001) and Koissi and Shapiro (2009) propose alternative solutions to accommodate the age-specific variation in the Lee-Carter model.

To address this problem, we apply the PLC model to four sub-periods: 1950-1964, 1965-1979, 1980-1994, and 1995-2007, and estimate the parameters $\hat{\alpha}_x$, $\hat{\beta}_x$ and $\hat{k}_t$ in each case. Results are presented in Figures 21-25.
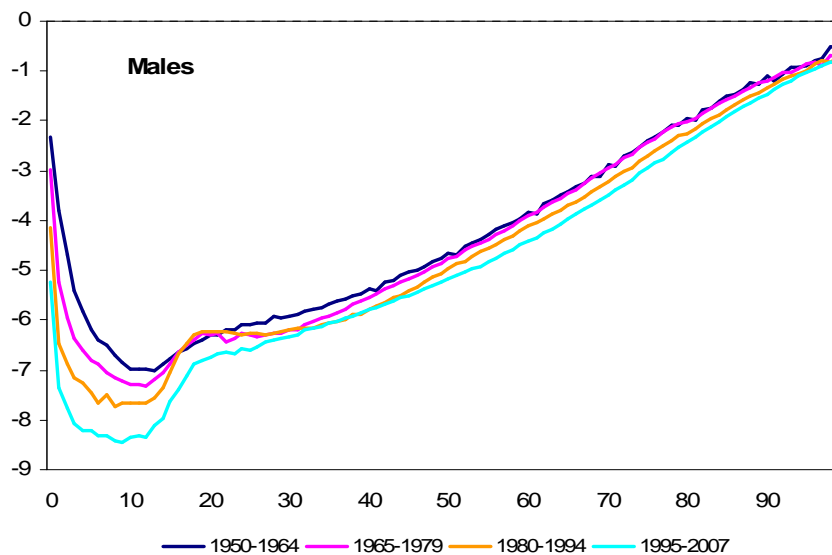

Figure 21. PLC model estimates for males: $\hat{\alpha}_x$, for each of the four sub-samples considered, Portugal.
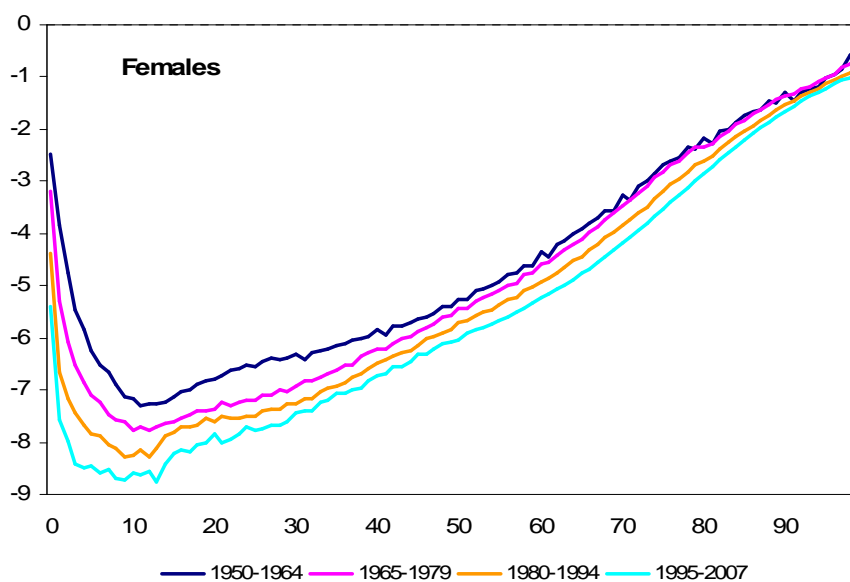

Figure 22. PLC model estimates for females: $\hat{\alpha}_x$, for each of the four sub-samples considered, Portugal.
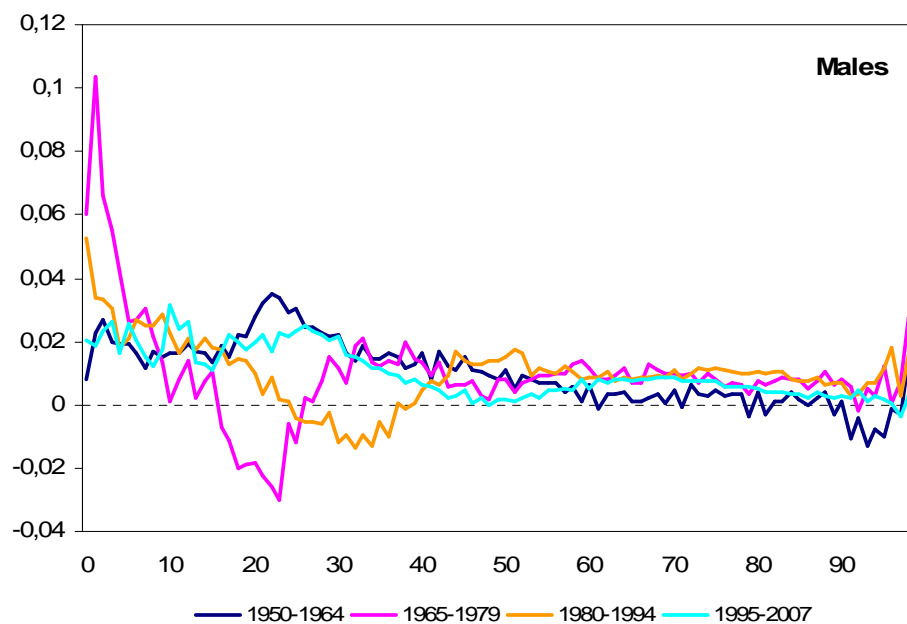
Figure 23. PLC model estimates for males: $\hat{\beta}_x$, for each of the four sub-samples
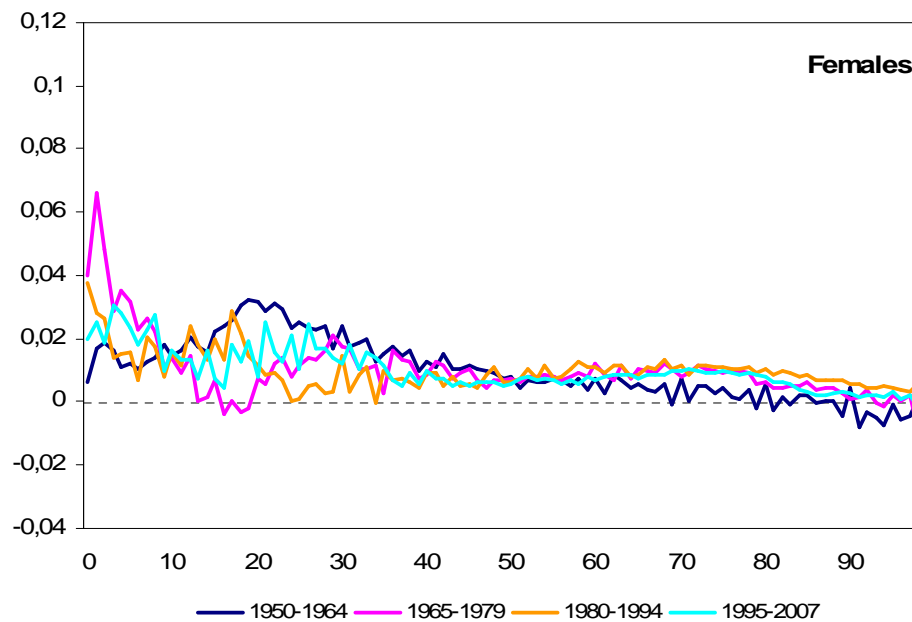
considered, Portugal

Figure 24. PLC model estimates for females: $\hat{\beta}_x$, for each of the four sub-samples

considered, Portugal

The behaviour of the estimated overall age profiles, $\hat{\alpha}_x$, presented in Figures 21 and 22, shifting down over time, is consistent with the pattern already observed in Figures 3 and 4, reflecting the decline in death rates at all ages with the exception of the increase in male mortality between 15 and 25 years old in the period 1980 -1994.

The trajectories of the $\hat{k}_t$, presented in Figure 25 for all sub-periods, show a well-defined decreasing trend. However, the rhythm of decline between 1965 and 1994 is relatively more moderate, especially for men. In the latest period, there is evidence of a more pronounced decline in mortality trends. In a study of the mortality trends in developed countries, Coelho and Nunes (2011) also find evidence of a structural change in the overall rate of mortality decline for Portugal in this period.
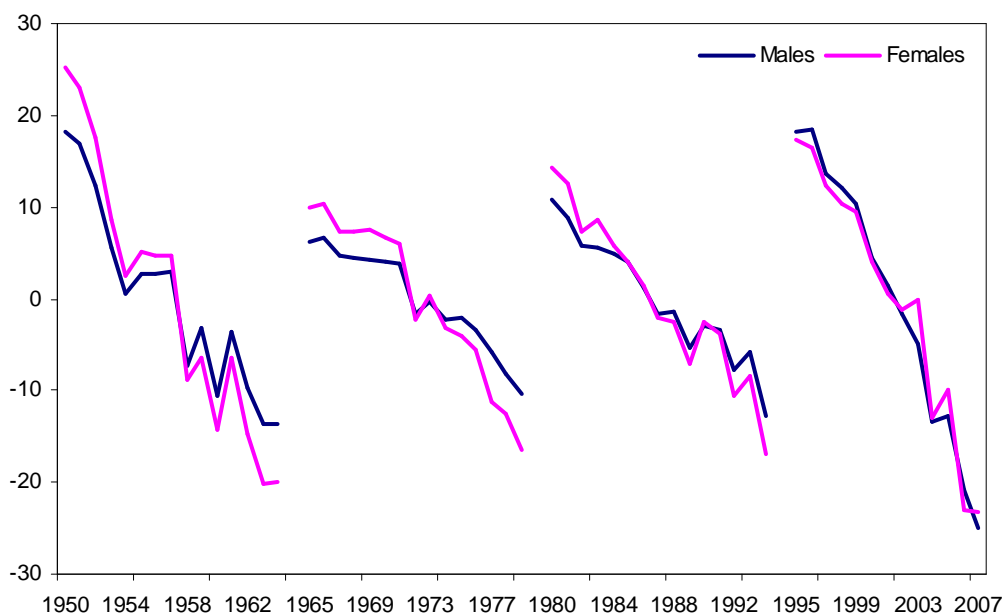
Figure 25. PLC model estimates for males and females: $\hat{k}_t$, for each of the four sub-samples considered, Portugal

Figures 23 and 24 depict the age-specific parameters $\hat{\beta}_x$, for men and women, estimated for each sub-time period. The rate of improvement in mortality at each age changes substantially over time, confirming the descriptive analysis in Section 2. In the period 1950 to 1964, the death rates of young adults declined more rapidly than at other ages. From 1965 to 1979, males aged 15 to 25 experience strong deteriorations in

mortality. The relative deterioration in male mortality shifts into older ages (ages 25 to 40) in the period 1980-1994. In this period, the greatest relative mortality declines occur at younger ages as well as at older adult ages. The period 1995 – 2007 is marked by relative improvements in mortality most evident in the younger ages that extend up to about 40 years old.

3.6.   Results of the Double Bilinear Poisson Lee-Carter Model

The patterns observed in the residual analysis of the PLC model, especially for men, suggest possible age-time interactions not fully captured by the model, similar to the pattern found by Booth et al. (2002) for Australia. Moreover, the results of applying the PLC model to different sub-periods confirm the findings of the descriptive analysis in Section 2 that the rate of mortality decline over time did not follow the same pattern for all ages. This suggests applying the extension of the PLC model which adds a second bilinear term to the demographic model, henceforth designated as PLC2, in order to better capture the time trend behaviour in the age-specific death rates. Estimates of the parameters $\hat{\beta}_x^1$, $\hat{\beta}_x^2$, $\hat{k}_t^1$ and $\hat{k}_t^2$ for this extended model are plotted in Figures 26 to 29. We omit the estimated $\hat{\alpha}_x$ since they were almost identical to the ones obtained with the PLC model considered in Section 4.

The estimated $\hat{k}_t^1$ show an almost linear downward trend for both men and women. The values of $\hat{\beta}_x^1$ are positive for all ages (except for ages above 90), which associated with the decrease in $\hat{k}_t^1$, translate into a roughly exponential decrease of death rates. This first term is thus very similar to the $\hat{k}_t$ obtained in the PLC model. The rate of decline captured in $\hat{k}_t^1$ is attenuated or accentuated by the effect of $\hat{k}_t^2$. This second term starts with a decreasing trend in the 1950s. In the 1960s, it is more or less constant for women, after which it shows a growing trend. In the case of men, $\hat{k}_t^2$ presents an increasing trend after 1960, stabilizing near the end of the twentieth century. The values of $\hat{\beta}_x^2$ are larger for young adult ages, both men and women, which means

that the impact of the introduction of this second term is more focused at these particular ages.
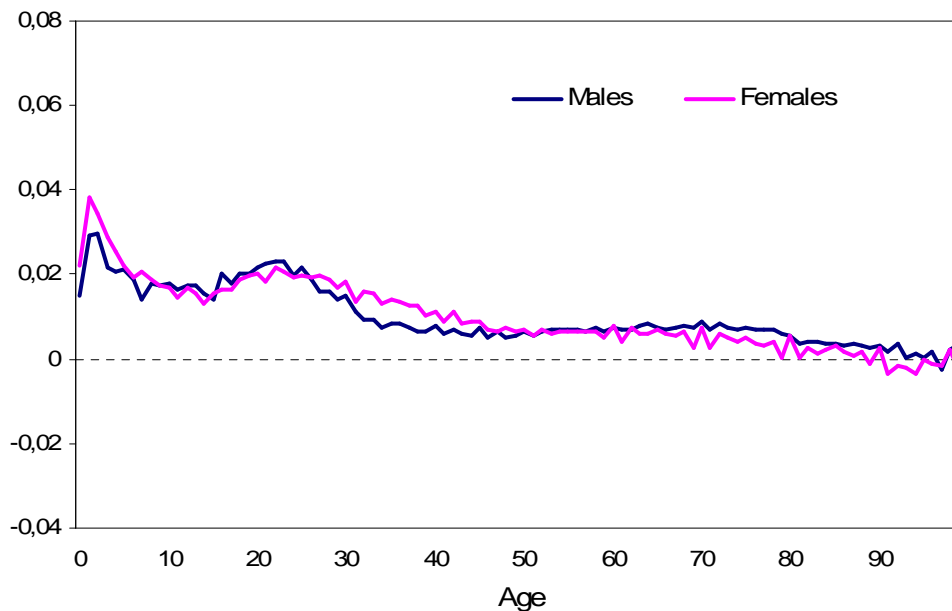

Figure 26. PLC2 model estimates: $\hat{\beta}_x^1$ for ages from 0 to 99, 1950 – 2007, Portugal.
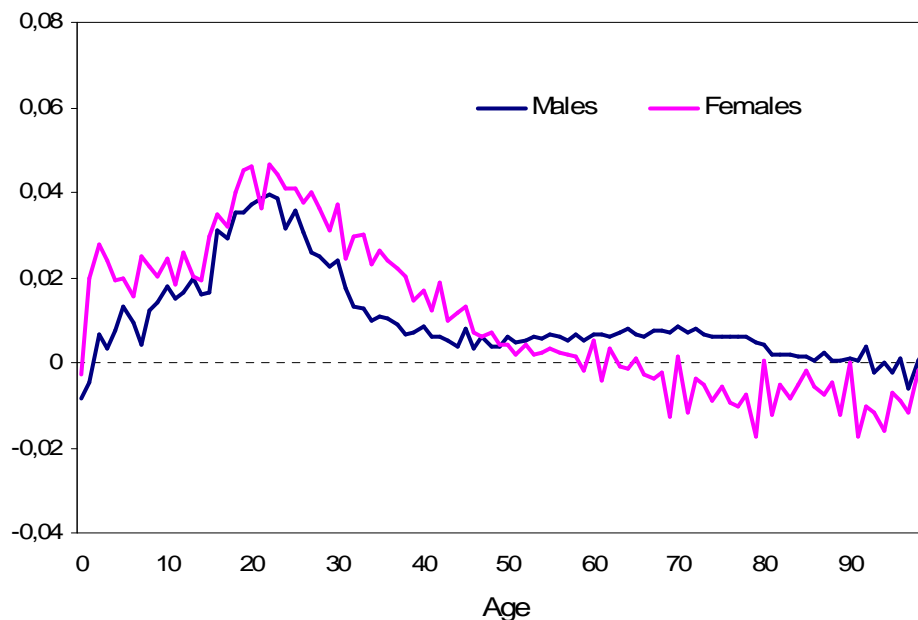

Figure 27. PLC2 model estimates: $\hat{\beta}_x^2$ for ages from 0 to 99, 1950 – 2007, Portugal.

For men, the growing trend of $\hat{k}_t^2$ from 1960 till the end of the century associated with positive values of $\hat{\beta}_x^2$, specially for young adult ages, results in the attenuation of the effect of the first factor, i.e. reducing the rate of decline of the death rates in that period for those ages. This also implies that the rhythm of decline of male mortality is accentuated at the end of the twentieth century and beginning of the twenty first century.
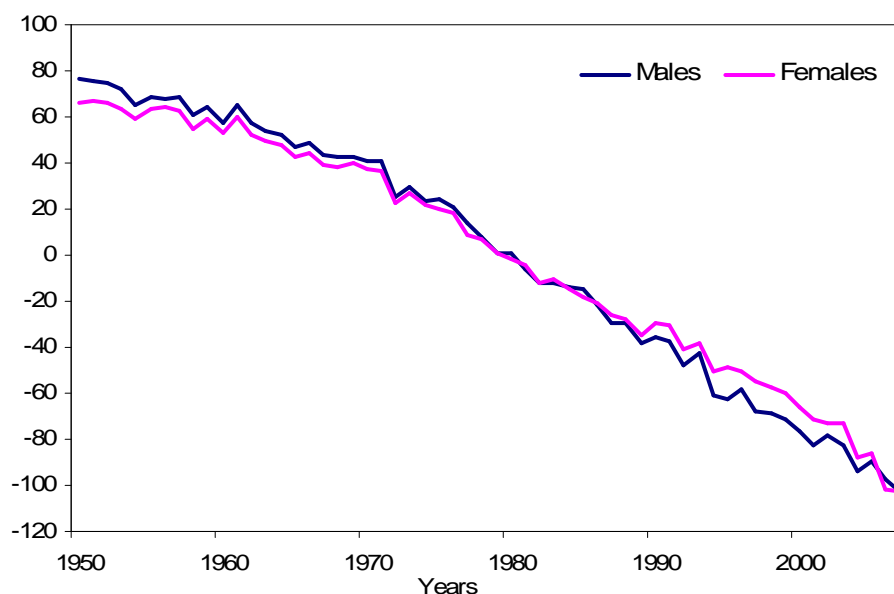
Figure 28. PLC2 model estimates: $\hat{k}_t^1$, 1950-2007, Portugal.

The impact of the introduction of a second term for women is less significant since $\hat{k}_t^2$ varies less over time. The major effect is to reduce the gains determined by the first factor between ages 15 to 30 years since 1970. It should be noted that in the case of women, $\hat{\beta}_x^2$ takes negative values for ages above 60 years which attenuates the rate of decline of death rates for these ages in the first years of the period in analysis while accentuating the decline afterwards.
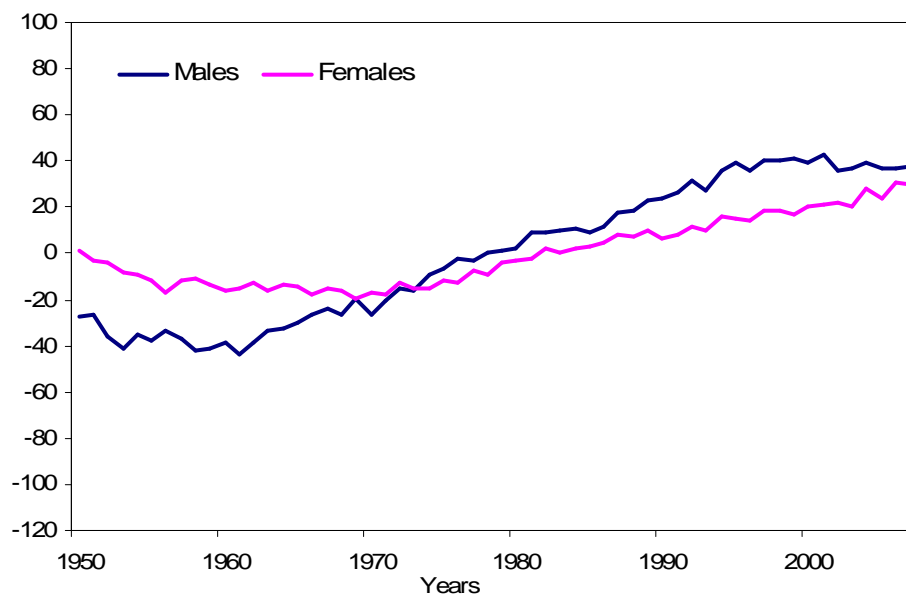
Figure 29. PLC2 model estimates: $\hat{k}_t^2$, 1950-2007, Portugal.

Overall, the inclusion of the second term improves slightly the fit of the model. The percentages of total deviance explained by the model with two factors are higher for both men and women, the values of R^2_{Dev} are 96,8% for males and 97,6% for females. As can be seen from Figure 31, the fit is improved especially at young adult ages. Although systematic patterns in the residuals are still found for men and women, as can be seen in Figure 30, they have a relatively smaller dimension when compared with the PLC model considered in Section 4, in particular for men.

We conclude that the second term enables the model to capture the non-linearities in the decline of log-death rates for particular ages, in particular young adults. There is also some evidence of some possible structural change at the end of the twentieth century with a more accentuated decline in male mortality thereafter.

Figure 30. Standardized deviance residuals from PLC2 model for males (left panel) and females (right panel), 1950-2007, Portugal

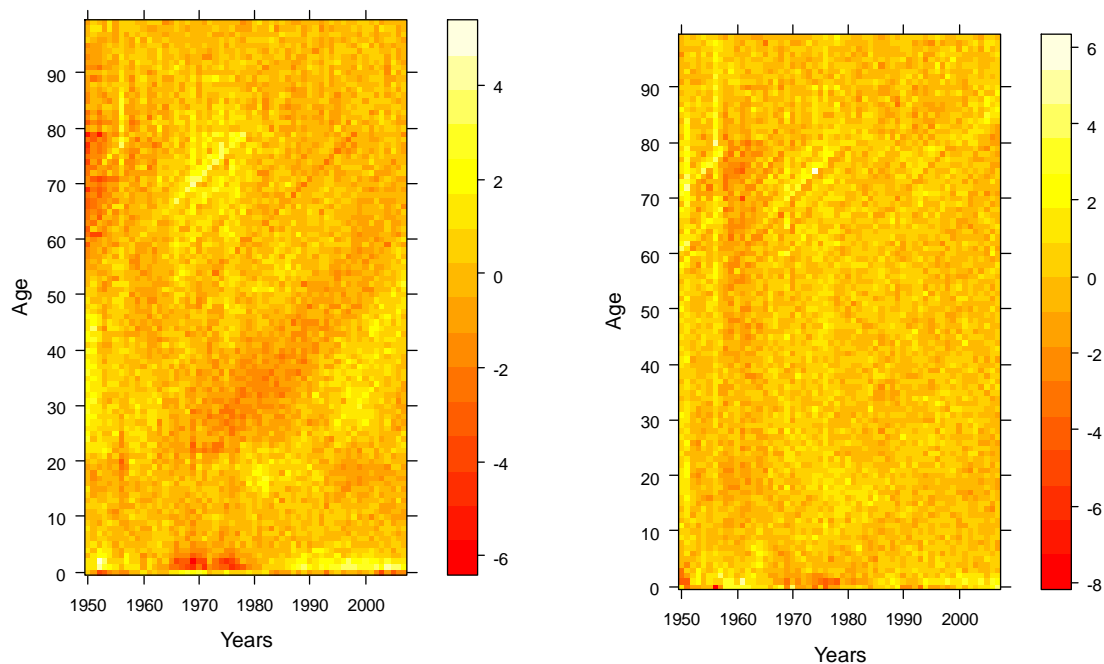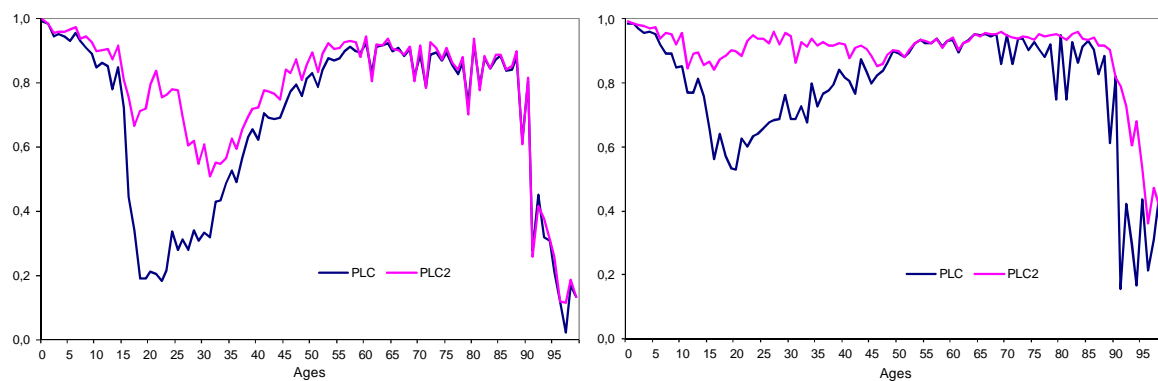*Figure 31.* Proportions of total explained deviance of the number of deaths by age from PLC and PLC2 models for males (left panel) and females (right panel), 1950-2007, Portugal.

3.7. Conclusion

In this paper, we analyse the pattern of mortality decline in Portugal since 1950. The general shape of the age specific death rates is consistent with the typical pattern found in other countries, and has moved down over time reflecting the overall decreases in mortality at all ages. These improvements have been greater at lower ages, especially infant mortality, than at higher ages. However the rate of improvement in death rates showed considerable variability over the period considered, particularly in the case of teens and young adults. In more recent years, increases in life expectancy at birth are originating more and more from improvements in mortality at older ages. Also, the beginning of the 21st century denotes an apparently different pattern in mortality decline, particularly evident for men up to 40 years old. Evidence of some apparent cohort effect for people born between 1950 and 1960, have been noticed, especially for men, that seem to have experienced particular poor improvements in mortality when compared with adjacent cohorts.

We estimate several extensions of the Lee-Carter mortality model in order to capture the main characteristics of the decline in mortality. The results obtained from using the Poisson Lee-Carter model with one interaction term show an almost linear declining trend of mortality over time. The model also captures the greater decline in death rates at younger ages in response to changes in this general declining time path. Although the model fits reasonably well the data in general terms, for some specific ages the fit is not so good, as is the case of men between 20 and 35 years. From an analysis of the pattern of the residuals, an eventual 1950-1960 cohort effect is also not appropriately captured by the model. The Lee-Carter model was also estimated for different sub-periods. The results confirm that the patterns of mortality improvements at each age have changed over time. Finally, we considered the model incorporating an extra age-time interaction component. The inclusion of this second term is able to improve the fit of the model by capturing the non-linearity in the decline of log-death rates at some ages.

We conclude that a careful descriptive analysis of the patterns of mortality decline is essential to help formulate and validate parsimonious mortality models. In our

analysis of Portuguese data, we discussed how an extension of the mortality model was able to provide a better fit to the data by capturing important characteristics of the data.

However, there were a number of issues that we left for subsequent studies. A better understanding of the major patterns of mortality decline and the choice of more appropriate mortality models is important not only for historical purposes but also for forecasting purposes. Alternative formulations of the stochastic model will have consequences in terms of predictions of future mortality and life expectancy. An evaluation of the different models in terms of forecasting validity and performance is therefore an important issue to investigate. We also leave for future research some questions that were raised during our analysis, namely studying the eventual presence of a cohort effect.

REFERENCES

Alho, J. (2000). Discussion. *North American Actuarial Journal*, 4(1), 91-93.

Arriaga, E. (1984). Measuring and Explaining the Change in Life Expectancies. Demography, 21(1), pp. 83-96.

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

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under Conditions of Variable Mortality Decline. *Population Studies*, 56, 325– 336.

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

Booth, H., and Tickle, L. (2008), Mortality modelling and forecasting: A review of methods, *ADSRI Working Paper nº. 3.* Available at: htt://adsri.anu.edu.au/.

Brouhns, N., Denuit, M., and Vermunt, J.K. (2002). A Poisson Log-linear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics*, 31, 373-393.

Brouhns, N., Denuit, M., and Van Keilegom, I. (2005). Bootstrapping the Poisson Log-bilinear Model for Mortality Forecasting. *Scandinavian Actuarial Journal*, 3, 212 -224.

Cameron, A. C., and Trivedi, P. K. (1998). Regression Analysis of Count Data. Cambridge University Press.

Cameron, A. C., and Windmeijer, F. A. G. (1997). An R-squared Measure of Goodness of Fit for Some Common Nonlinear Regression Models. *Journal of Econometrics*, 77, 329 -342.

Canudas-Romo, V., Coelho, E., and Pina, C. (2008). About Mortality Data for Portugal. Available at www.mortality.org.

Carter, L.R., and Prskawetz, A. (2001). Examining Structural Shifts in Mortality Using the Lee-Carter Method. Max Planck Institute for Demographic Research, WP 2001-007.

Coelho, E. and Nunes, L. C. (2011). Forecasting Mortality in the Event of a Structural Change. *Journal of the Royal Statistical Society A* 174, Part 3, pp 713-736.

Debón, A., Montes, F., and F. Puig. (2008). Modelling and Forecasting Mortality in Spain. *European Journal of Operational Research* 189: 624–637.

Fígoli, M.G.B. (1998). Modelando e Projectando a Mortalidade no Brasil. *Revista Brasileira de Estudos de População*, 15(1), 95-113.

Girosi, F., and King, G. (2008). *Demographic Forecasting*, Princeton University Press, Princeton.

Goodman, L.A. (1979). Simple Models for the Analysis of Association in Cross-Classification Having Ordered Categories. *Journal of American Statistical Association* 74, 537-552.

Hardle, W., and Mysickova, A. (2009). Stochastic Population Forecast for Germany and its Consequence for the German Pension System. SFB 649 Discussion Papers SFB649DP2009-009, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

Heligman, L., and Pollard, J. H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries* 107: 49-80.

Human Mortality Database, HMD. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), Available at www.mortality.org or www.humanmortality.de (data downloaded on [17/10/2008]).

Koissi, M., Shapiro, A., and Högnäs, G. (2006). Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38, 1-20.

Koissi, M. and Shapiro, A. (2009). The Lee-Carter Model under the Condition of Variable Age-specific Parameters. Available at: http://www.soa.org/ library/proceedings/arch/2009/arch-2009-iss1-koissi- shapiro.pdf.

Lee, R. and Carter, L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87(419), 659-675.

Lee, R. and Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38(4), 537-549.

Lundstrom, H. and Qvist, J. (2004). Mortality Forecasting and Trend Shifts: an Application of the Lee-Carter Model to Swedish Mortality Data. *International Statistical Review*, 72(1), 37-50.

Renshaw, A., and Haberman, S., (2003). Lee–Carter mortality forecasting with age specific enhancement. *Insurance: Mathematics and Economics* 33 (2), 255–272.

Renshaw, A. E., and Haberman, S., (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 58: 556–570.

Rodrigues, T., Moreira, M. J. G., and Fernandes, A. A. (2004). Social Changes and Better Health Conditions of the Portuguese Population 1974 – 2000. *Hygiea Internationalis*, volume 4, nº1.

Tuljapurkar, S., Li, N., and Boe, C. (2000). A Universal Pattern of Mortality Decline in the G7 Countries. *Nature*, 405, 789-792.

Wilmoth, J.R. (1993). Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality change. Technical Report, University of California, Berkeley. U.S.A.

Wilmoth, J.R. (1998). Is the Pace of Japanese Mortality Decline Converging Toward International Trends?. Population and Development Review, Vol. 24 n.º3 (sep.1998), pag. 593 – 600.

Wilmoth, J.R., Andreev, K., Jdanov, D., and Glei D.A. (2007) Methods Protocol for the Human Mortality Database. Human Mortality Database. Available at: http://www.mortality.org/Public/Docs/MethodsProtocol.pdf.

Willets, R.C. (1999). Mortality in the next millennium. Paper presented to the Staple Inn Actuarial Society.

Willets, R.C. (2004). The cohort effect: insights and explanations. British Actuarial Journal, 10, 833-877.

4. COHORT EFFECTS IN MORTALITY MODELLING

Abstract

In Portugal, as in other developed countries, substantial mortality gains have occurred at all ages in the last decades. However, a previous exploratory study of the patterns of mortality decline in Portugal over the period 1950 – 2007 (Coelho and Nunes, 2010) shows that the rhythm at which mortality has been improving differs between ages and over periods of time. In particular, an apparently odd pattern of mortality improvement was detected in the male population, suggesting that over time a specific group of male individuals might have experienced a non-declining or even increasing mortality.

In this paper we use age-period-cohort (APC) models to investigate the feasibility of a cohort effect in the Portuguese male mortality. Besides the classical additive APC model, the possibility of an interaction of period and cohort with age is considered. The Bayesian Information Criterion (BIC) is used to select a parsimonious model that adequately fits the data. The APC with age-period and age-cohort interactions is found to be the model that best captures male mortality patterns over the period 1950 – 2007. A simulation study is conducted to show the consequences of estimating an incorrect model. The results also confirm that BIC can be used to select the correct model.

Key words: Age-Period-Cohort model, Bayesian Information Criterion, cohort effects, period effects.

4.1. Introduction

In Portugal, as in other developed countries, substantial mortality gains have occurred at all ages in the last decades. However, a previous exploratory study of the patterns of mortality decline in Portugal over the period 1950 – 2007 (Coelho and Nunes, 2010) shows that the rhythm at which mortality has been improving differs between ages and over periods of time. In particular, an apparently odd pattern of mortality improvement was detected in the male population, suggesting that over time a specific group of male individuals might have experienced a non-declining or even increasing mortality[3]. Willets (1999, 2004) found a similar behavior for England and Wales, attributing it to "cohort-specific" influences.

The residual analysis of the Poisson Lee-Carter model (Lee and Carter, 1992; Brouhns, Denuit and Vermunt, 2002) applied to Portuguese male mortality data, by Coelho and Nunes (2011), also suggests that particular age-time influences are not fully captured by the model. These results lead us to consider the possibility of a cohort effect not yet identified in the temporal pattern of male mortality in Portugal.

Besides England and Wales, cohort effects are also found in Spain (Cleries, Martínez, Valls, Pareja, Esteban, Gispert, Moreno, Ribes, and Borràs, 2009), Denmark, and in the male populations of Austria, France, Italy, Japan, Netherlands and Switzerland (Andreev and Vaupel, 2005).

In this paper we use age-period-cohort (APC) models to investigate the feasibility of a cohort effect in the Portuguese male mortality. To our knowledge, all studies of the mortality evolution in Portugal have focused on the impacts of age and period, and relatively little is known about possible cohort effects. The objective of the APC analysis is to disentangle the effects of age, period and cohort on some outcome, in our case, mortality. The separation of age, period and cohort patterns in mortality decline can contribute in important ways to an understanding of the recent mortality decline as well as for forecasting future trends. In the prospect of a relevant cohort effect, the possibility of an interaction with age, that is cohort-specific age effects, is also considered. This is done by applying the extension of the Lee-Carter model

---

[3] The average annual mortality improvement rates for Portuguese males plotted by year of birth suggest that people born between 1950 and 1960, especially men, have experienced particular poor improvements in mortality when compared to adjacent generations.

including a cohort effect associated with an age interaction term as proposed by Renshaw and Haberman (2006).

The Bayes Information Criterion (BIC) is used to select the most parsimonious model that best captures mortality patterns (Cairns, Blake, Dowd, Coughlan, Epstein, Ong and Balevich, 2009). As such, we conclude for cohort specific influences in the data if the selected model includes cohort effects. A simulation study allows us to confirm if BIC is indeed selecting the correct model and to explore what are consequences of estimating an incorrect model.

In the first section we present the theoretical framework of the classical APC model, the well-known problem of model identification, classical estimation methods and some recent developments. Then the introduction of interactions in the model is discussed and the general model of Renshaw and Haberman (2006) is presented. Using data on the annual number of deaths and population size by age over the period 1950 – 2007 empirical results are presented. Finally, the results of the simulation study are shown, followed by final considerations.

4.2. Age-Period-Cohort Analysis

4.2.1. Introduction

The discussion of the importance of cohort influences on mortality risks dates back to the 1920's (Hobcraft, Menken and Preston, 1982; Richards, 2008; Murphy, 2010). According to Collins (1982) and Hobcraft et al. (1982), it was Derrick, in 1927, who first argued that age-specific death rates plotted by year of birth provided a more consistent basis for projecting mortality patterns rather than the more familiar year of death. While analyzing age-specific mortality rates for England and Wales from 1841 to 1925, Derrick found that logarithms of mortality rates plotted against year of birth were strikingly parallel and concluded that almost all the temporal change in mortality is due to the birth cohort influence (Hobcraft et al., 1982). In 1934, Kermack, McKendrick and McKinlay reached similar conclusions for the regularity of the cohort effect that changes the relative risk of mortality by the same proportion at most ages (Murphy,

2010; Hobcraft et al., 1982). The term "cohort" was used and defined for the first time in 1939 in a posthumous publication on tuberculosis by Frost (Case, 1956). Using age-specific death rates from tuberculosis in Massachusetts from 1880 to 1930, Frost showed graphically that age-specific rates seem to be more regular for cohorts than for periods and demonstrated that age patterns of mortality from tuberculosis look quite different for cohorts and for periods (Hobcraft et al., 1982).

The graphical display of the data can reveal probable relationships. However a formal approach to modeling data, such as regression analysis is thought to be helpful. According to Hobcraft et al. (1982), the first proper identified APC model was specified by Greenberg, Wright, and Sheps (1950), where the age effects were parameterized through a beta distribution. While the statement of the APC model is straightforward, the estimation of the model poses a statistical problem, known as the identification problem, due to the linear relationship between age, period and cohort. Two decades later, K. Mason, W. Mason, Winsborough and Poole (1973) address the identification problem in the APC model and, using a multiple classification framework, propose a constrained-based approach, founded on a priori and theoretical knowledge, to estimate the age, period and cohort model parameters. Following the constrained-based approach proposed by Mason et al. (1973), several other approaches have been proposed in order to solve the identification problem of the APC model. Among them are the analyses ignoring sequentially one of the three dimensions and the search for an estimable function of the parameters (Kupper and Janis, 1980; Rodgers, 1982; Holford, 1983; Kupper, Janis, Karmous and Greenberg, 1985; Holford, 1991; Clayton and Schifflers, 1987; Fu, 2000; Yang, Fu, and Land, 2004; Kuang, Nilsen and Nilsen, 2008a).

Holford (1983) focused on finding estimable functions of the parameters which are invariant as to the particular constraint applied, and in particular proposed a partition of age, period and cohort effects into linear and curvature components, the latter components being uniquely determined, and noting that the second differences are a well known measure of curvature. Clayton and Schifflers (1987) also focused their attention on what functions of the death rates could be meaningfully estimated, in particular on the ratios of relative risks which are identifiable. In the logarithmic scale these ratios correspond to second order differences. Zero values for the second order differences indicate that the log-risk versus calendar time curve is locally a straight line,

while positive or negative values indicate convex or concave relationships, respectively. The second differences have the important property of representing characteristics specifically attributable to age, period or cohort without any arbitrary repartition of the linear components.

Although literature on age-period-cohort models is rich with attempts to solve the identifiability problem, Clayton and Schifflers (1987) and Carstensen (2007) agree that such attempts are useless, since is not possible to surmount the problem of having two variables as well as their sum in the same linear model. According to them it is not possible to uniquely estimate the absolute effects of age, period and cohort, unless the researcher is willing to make some assumptions on the relative importance of the effects. Carstensen (2007) proposed what he called as a sensible parameterization to overcome the identifiability problem of the APC model. A sensible parameterization depends on the analysis of the subject matter, it is meaningful, understandable and recognizable, practically estimable, and allows reconstruction of the fitted rates from the estimated parameters.

Carstensen (2007) also proposed an extension of the APC model considering explicitly data tabulated by Lexis triangles, and parametric smooth functions, such as splines, natural splines or fractional polynomials, to model age, period and cohort effects.

Using the second order differences of the three effects as measures of curvature, Kuang et al. (2008b) propose an alternative approach to the identification problem in the APC model. Their approach consists in transforming the design matrix so that curvatures (the second differences) of age, period and cohort are estimated directly.

4.2.2. The Classical Additive Age-Period-Cohort Model

The purpose of the APC model is to estimate the respective contributions of the effects of age (at the event), period (year of occurrence) and cohort (year of birth) on the evolution of the variable of interest.

Age is considered the most important source of variation in vital rates (Hobcraft et al., 1982; Clayton and Schifflers, 1987). It represents the mortality risks related to

biological and physiological factors, which are associated to the aging process, and which affect equally all individuals of the same age, regardless of their cohort and observation period.

Period effects represent contemporaneous conditions that change the level of mortality risk at all ages. Period effects may be associated, for example, to advances in medical knowledge, the development of new diagnostic methods, the access to health facilities, improved sanitary conditions, improvement in life conditions; environmental conditions that are susceptible to affect all cohorts at all ages.

Cohort effects represent the influence of past conditions on current mortality risks, which differentially affects age groups. The cohort could be a generation of individuals, a group of individuals born during the same year (birth cohort), or a group of people that lived through important comparable historical events or structural environment changes in the same period (Wilmoth, 2001, Willekens and Scherbov, 1991). The birth cohort effect is better thought of as a generational effect, because it is not limited to exposures that might occur at the time of birth (Holford, 2006). The cohort represents the events that affect in a similar way all individuals of the same cohort regardless of the age of the individuals and observation period.

4.2.2.1. Model formulation

Let D_{ij} denote the number of deaths recorded at age i and calendar year j, and E_{ij} the number of persons at risk of death at age i and calendar year j, where $i \in [i_1, i_I]$ and $j \in [j_1, j_T]$. Consider a rectangular data array, (D_{ij}, E_{ij}), comprising the observed number of deaths D_{ij} with matching exposure to risk of death E_{ij}. Age-specific death rates are calculated by the ratio of the number of deaths and exposure-to-risk estimates in matched intervals of age and time (see Human Mortality Database (2007)), $m_{ij} = \dfrac{D_{ij}}{E_{ij}}$.

Cohort or year of birth z is defined by $z = j - i \in [j_1 - i_I, j_T - i_1]$.

Let the dependent variable y_{ij} denote some monotonic transformation (e.g., linear, logarithm, or logit) of the observed death rates, m_{ij}, at age i and year j, which is modeled as a linear additive function of age, period, and cohort effects. The APC model may be written as:

$$y_{ij} = \delta + \alpha_i + \beta_j + \gamma_z + \varepsilon_{ij} \qquad (1)$$

where the parameter δ is an overall constant that establishes a general level to y_{ij}; α_i, β_j, and γ_z denote the effects for age group i, period j, and cohort z, respectively; and ε_{ij} is the random error component with $E(\varepsilon_{ij}) = 0$.

The random distribution of ε_{ij} is tied to the assumptions on the stochastic nature of y_{ij}. Most age-period-cohort models have been approached as a class of generalized linear models (GLM) (Robertson and Boyle, 1998a; Robertson, Gandini and Boyle, 1999). The Poisson distribution is generally a valid distribution to model counts of demographic events, such as deaths (Brillinger, 1986), and log-linear Poisson models are widely used in demography and epidemiology.

4.2.2.2. The identification problem

The major problem with APC models is the exact linear relationship between age, period and cohort, $z = j - i$, which can be viewed as a special case of collinear regressors. Since there are infinitely many linear transformations leading all to the same estimated occurrence (incidence or mortality) rate, it is impossible to estimate a unique set of parameters.

The identification problem, in that arbitrary constants and an arbitrary linear trend can be added to the elements of the three components parameters without changing the predicted values of y_{ij}, can be illustrated using arguments of group theory as in Kuang et al. (2008a) and in Carstensen (2007). The predicted values of y_{ij} given equation (1) are a function of the parameters δ, α_i, β_j and γ_z, with $z = j - i$, which is invariant to the group transformations g defined by

$$
g : \begin{pmatrix} \alpha_i \\ \beta_j \\ \gamma_z \\ \delta \end{pmatrix} \mapsto \begin{Bmatrix} \alpha_i + a + \lambda \cdot i \\ \beta_j + b - \lambda \cdot j \\ \gamma_z + c + \lambda \cdot (j - i) \\ \delta - a - b - c \end{Bmatrix} \qquad (2)
$$

where a, b, c and λ are arbitrary real numbers. These means that if the estimated parameters $\hat{\delta}$, $\hat{\alpha}_i$, $\hat{\beta}_j$ and $\hat{\gamma}_z$ are adjusted in such a way that:

$$
\alpha_i^* = \hat{\alpha}_i + a + \lambda \cdot i
$$

$$
\beta_j^* = \hat{\beta}_j + b - \lambda \cdot j
$$

$$
\gamma_z^* = \hat{\gamma}_z + c + \lambda \cdot z
$$

$$
\delta^* = \hat{\delta} - a - b - c
$$

then the predicted values of y_{ij} using model (1) are the same, for the two sets of parameters ($\hat{\delta}$, $\hat{\alpha}_i$, $\hat{\beta}_j$, $\hat{\gamma}_z$) and (δ^*, α_i^*, β_j^*, γ_z^*), whichever the choice of a, b, c and λ, since $z = j - i$:

$$
\begin{aligned}
y_{ij}^* &= \delta^* + \alpha_i^* + \beta_j^* + \gamma_z^* \\
&= \hat{\delta} - a - b - c + \hat{\alpha}_i + a + \lambda \cdot i + \hat{\beta}_j + b - \lambda \cdot j + \hat{\gamma}_z + c + \lambda \cdot z \\
&= \hat{\delta} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_z + (-a - b - c + a + b + c) + \lambda(i - j + z) \\
&= \hat{\delta} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_z + \lambda((i - j) + (j - i)) \\
&= \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_z = \hat{y}_{ij}.
\end{aligned}
$$

Since the different sets of parameters age, period and cohort lead to identical estimated values for the dependent variable, there is no way of selecting between the alternative solutions based on the goodness of fit quality.

A regression model that includes complete sets of dummy variables for each age, period and cohort factors is not estimable since the columns of the regressors matrix (matrix X in OLS) will be linear functions of one another and consequently the matrix $X'X$ is not invertible. Usually in regression models that assume unique effects for each category within each dimension and each dimension is represented exhaustively by its categories, a location or normalization constraint, such as taking one level of age, period

and cohort as a reference by setting one of each three scales parameters to zero or sum-to-zero constraints to center the parameters so that they sum to zero, are sufficient to identify the model. Such constrains, however, are insufficient to break the linear dependency between age, period, and cohort in the age-period-cohort model.

Arbitrary linear trends may be added to any of the factors with compensating trends in the other two (Carstensen, 2007). Unless one has prior reasons for fixing overall trends in any of the effects, only departures from linearity can be detected (Holford, 1983; Clayton and Schifflers, 1987).

4.2.2.3.   The estimation approach of Kuang, Nilsen and Nilsen

Acknowledging the identification problem of the APC model, several estimation strategies have been proposed. Following the suggestion of Holford (1983) and Clayton and Schifflers (1987), Kuang, et al. (2008a) propose an alternative approach to the identification problem in the APC model using second order differences, which can be uniquely estimated. Their proposal consists in transforming the design matrix so that curvatures (the second differences) of age, period and cohort are estimated directly.

Kuang et al. (2008a) consider the identification problem and propose a maximal invariant parameter based on the second differences and three initial points designated a canonical parameterization. Considering the APC model for the logarithm of mortality rates, $y_{ij} = \log m_{ij}$:

$$y_{ij} = \delta + \alpha_i + \beta_j + \gamma_z + \varepsilon_{ij} \tag{3}$$

with age $i = i_1,\ldots, i_I$, year $j = j_1,\ldots,j_T$, and cohort $z=j-i= j_1- i_I,\ldots, j_T - i_1$, Kuang et al. (2008a) show that it can be written in terms of the second differences and three initial points:

$$y_{ij} = y_{i_1,j_1} + (i-1)\left(y_{i_1+1,j_1} - y_{i_1,j_1}\right) + (j-1)\left(y_{i_1,j_1+1} - y_{i_1,j_1}\right) + a_{ij} \tag{4}$$

with

$$a_{ij} = \sum_{s=i_1+2}^{i}(i-s+1)\Delta^2\alpha_s + \sum_{s=j_1+2}^{j}(j-s+1)\Delta^2\beta_s + \sum_{s=j_1-i_I+2}^{j-i}(j-i-s+1)\Delta^2\gamma_s$$

where Δ^2 denotes the second difference operator. In (4), y_{i_1,j_1} determines the level, $\left(y_{i_1+1,j_1} - y_{i_1,j_1}\right)$ and $\left(y_{i_1,j_1+1} - y_{i_1,j_1}\right)$ determine linear effects, and the double differenced parameters determines the non-linear effects. Second differences have the advantage of being interpretable. A nil value indicate the absence of change in the local trend, a positive value or a negative value correspond respectively to an acceleration or deceleration of the local trend.

The canonical parameter proposed is

$$\xi = \left(y_{i_1,j_1}, y_{i_1+1,j_1}, y_{i_1,j_1+1}, \Delta^2\alpha_{i_1+2}, \ldots, \Delta^2\alpha_{i_I}, \Delta^2\beta_{j_1+2}, \ldots, \Delta^2\beta_{j_T}, \Delta^2\gamma_{j_1-i_I+2}, \ldots, \Delta^2\gamma_{j_T-i_1}\right) \quad (5)$$

Kuang et al. (2008a) demonstrate that ξ gives a unique parameterization of y_{ij} [4]. For estimation, a design matrix for the canonical parameter ξ can be deducted from (4).

The uniqueness of ξ implies that the design matrix has full column rank. So the model may be estimated by least squares regression or generalized least squares regression (linear or Poisson regression).

The original parameters α_i, β_j, and γ_{j-i} can be constructed from ξ, imposing some identification restrictions on the parameters. To do so the researcher should look at the estimated parameters in conjunction with information external to the data. For instance, one can set the constant, the first age, cohort and period parameters equal to zero and an additional equality constraint on the first two calendar year parameters in order to get estimates for the parameters levels, $\delta = \alpha_1 = \gamma_1 = \beta_1 = \beta_2 = 0$.

Carstensen (2007), Kuang et al. (2008a) remark the importance of the graphical representation of the three components estimated effects, and the fact that the first differences as well as the absolute levels can be moved around between the three components.

---

[4]Kuang et al. (2008a) show that the three corner points ($y_{i_1,j_1}, y_{i_1+1,j_1}, y_{i_1,j_1+1}$) can be replaced by three other points $\left(y_{i',j'}, y_{i'',j''}, y_{i''',j'''}\right)$ and the parameterization of y_{ij} by ξ remains unique (exactly identified) if the matrix B, based on the new set of three points, and defined as $B = \left(b_{i'j'} \quad b_{i''j''} \quad b_{i'''j'''}\right)'$ where $b_{ij} = \left(1, i-1, j-1\right)$, is invertible.

4.2.3. Age-by-Period and Age-by-Cohort Interactions

The classical APC model is a linear, additive, fixed effects model. It assumes that age effects are the same for each period and cohort, period effects are the same for each age and cohort, and cohort effects are the same for each age and period. The empirical and theoretical adequacy of the classical APC model that allows no interaction terms have been questioned (Glenn, 1976; Rodgers, 1982; Wilmoth, 1990). Glenn (1976) argues that there are often substantive reasons for assuming the existence of such interactions. Following the criticism of Glenn (1976), and taking into consideration the discussion on the conceptual nature of a cohort in Ryder's (1965), in that he explains that the cohort can change over time, not only due to composition changes, but because different cohorts live through different social events, or the same events at different ages, and these events may change the cohort nature in several lifelong ways, Fienberg and Mason (1982) describe several ways to extend APC models to include interactions.

The inclusion of interaction terms was also proposed as a solution to the identifiability problem of the classical additive APC model. Moolgavkar, Stevens and Lee (1979) proposed, and James and Segal (1982) extensively analyzed, a non-linear APC model with an age interaction term associated to the period effects. The non-linear model for the time effect proposed by Moolgavkar et al. (1979) is just one example of interaction terms in the APC model.

According to Holford (1991, 2005) nonlinear models with age-period and/or age-cohort interactions are especially interesting, because in most situations they will yield a unique set of parameters, without the use of arbitrary constraints. However, the authors remark that the additive APC model is a special case of this form, and special caution is necessary. Also they note that intrinsically nonlinear models often have complex regions where the different parameters give nearly identical fitted values to the data, resulting in unstable parameter estimates.

The well known Lee-Carter model (Lee and Carter, 1992) includes a bilinear term in the age-period model, but no cohort effect. Renshaw and Haberman (2006), based on the work of James and Segal (1982), generalized the Lee-Carter model to incorporate cohort effects with an interaction term with age. The APC model is

generalized to allow age profiles to change over time, period effects to have different influences on people of different ages, and cohorts to change from one period to the next.

4.2.3.1.   The Lee-Carter model

Lee and Carter (1992) propose the following model for the logarithm of death rates, $y_{ij} = \log m_{ij}$:

$$y_{ij} = \alpha_i + \beta_i k_j + \varepsilon_{ij} \tag{6}$$

where α_i represents the general shape of the age-specific mortality profile (main age effects), k_j represent the main period effects (underlying time trend), β_i are age-specific parameters that measure the sensitivity of $\log m_{ij}$ to changes in the temporal index of mortality k_j, and ε_{ij} represent the error terms which are assumed to have zero mean and constant variance.

The structure is invariant to the group transformations g defined by:

$$g : \begin{pmatrix} \alpha_i \\ \beta_i \\ k_j \end{pmatrix} \mapsto \left\{ \begin{array}{l} \alpha_i - c\beta_i \\ \dfrac{\beta_i}{d} \\ d(k_j + c) \end{array} \right\} \tag{7}$$

where c and d are arbitary constants with $d \neq 0$. So k_j is determined up to a linear transformation, β_i only up to a multiplicative constant and α_i only up to a linear transformation.

In order to ensure the identifiability of the model two constraints must be imposed. Lee and Carter (1982) propose using the following identifying constraints $\sum_{i=i_1}^{i_I} \beta_i = 1$ and $\sum_{j=j_1}^{j_T} k_j = 0$, which imply that α_i is an average of $\log m_{ij}$ taken over time.

The homoscedasticity assumption on the error term ε_{ij} have been considered fairly unrealistic (Wilmoth, 1993; Alho, 2000) and Brouhns et al. (2002) proposed an

extension to the Lee-Carter model that keeps the log-bilinear functional form, but replaces the least squares approach with a Poisson regression for the number of deaths.

4.2.3.2.   The Renshaw-Haberman model

Renshaw and Haberman (2006) incorporated cohort effects into the model, adding an additional pair of bilinear terms $\beta_i^{(2)}\gamma_{j-i}$, where γ_{j-i} represent cohort effects and $\beta_i^{(2)}$ measure the corresponding interaction with age. The model proposed by Renshaw and Haberman, henceforth designated RH model, for the logarithm of death rates, $y_{ij} = \log m_{ij}$, can be written as follows:

$$y_{ij} = \alpha_i + \beta_i^{(1)}k_j + \beta_i^{(2)}\gamma_{j-i} + \varepsilon_{ij} \tag{8}$$

The structure is invariant to the group transformations g defined by:

$$g : \begin{pmatrix} \alpha_i \\ \beta_i^{(1)} \\ k_j \\ \beta_i^{(2)} \\ \gamma_{j-i} \end{pmatrix} \mapsto \begin{cases} \alpha_i - d\beta_i^{(1)} - f\beta_i^{(2)} \\ \dfrac{\beta_i^{(1)}}{c} \\ c\left(k_j + d\right) \\ \dfrac{\beta_i^{(2)}}{e} \\ e\left(\gamma_{j-i} + f\right) \end{cases} \tag{9}$$

where c, d, e and f are arbitrary constants with $c \neq 0$ and $e \neq 0$.

To deal with the identifiability problem of the model, due to the relationship between the three time components and to ensure a unique set of parameters, Renshaw and Haberman (2006) adopted a two-stage fitting strategy in which α_i is estimated first, according to the original Lee-Carter model (the average over time of the logarithm of death rates), followed by an iterative fitting algorithm to estimate the remaining parameters subject to the parameter constraints:

$$\sum_{i=i_1}^{i_I} \beta_i^{(1)} = 1, \quad \sum_{i=i_1}^{i_I} \beta_i^{(2)} = 1 \text{ and } k_{j_1} = 0.$$

The last constraint is used when setting the starting values and is subsequently maintained in the iterative fitting algorithm, while the first two constraints are imposed once the algorithm has converged.

According to Haberman and Renshaw (2009), estimating α_i first by conditioning on a predetermined static life table allows the preservation of a realistic representation of the main age effects which match the typical shape of a static life table on the log scale.

Renshaw and Haberman (2006) had already noted that convergence of the iterative fitting procedure was slow when fitting age-period-cohort with interactions. Cairns et al. (2009) also found that parameter values in the iterative scheme converge very slowly to their maximum likelihood estimates, which according to them suggests that some sort of identifiability problem remains. Arguing that the model proposed by Renshaw and Haberman (2006) lacked robustness in the parameter patterns, Cairns et al. (2009), proposed a different algorithm to fitting the Renshaw-Haberman model and impose the following constraints to ensure identifiability:

$$\sum_{j=j_1}^{j_T} k_j = 0, \quad \sum_{i=i_1}^{i_I} \beta_i^{(1)} = 1, \quad \sum_{i=i_1}^{i_I} \sum_{j=j_1}^{j_T} \gamma_{j-i} = 0 \ , \text{ and } \quad \sum_{i=i_1}^{i_I} \beta_i^{(2)} = 1 \ .$$

Imposing the sum of period and cohort parameters equals one, implies that the estimate for α_i will be (at least approximately) equal to the mean over time of the logarithm of the age-specific death rates, $\log m_{ij}$ (Cairns et al., 2009).

4.2.3.3.   The Lee-Carter model with additive cohort effects

The Lee-Carter model may be extended to include an additive main cohort effect, or alternatively we can set the age-cohort interaction parameters, in the RH model $\beta_i^{(2)}$, equal to a constant. In such a model cohort effects are fixed over the lifetime of the cohort. The Lee-Carter with additive cohort effects model for the logarithm of death rates, $y_{ij} = \log m_{ij}$, can be written as follows:

$$y_{ij} = \alpha_i + \beta_i k_j + \gamma_{j-i} + \varepsilon_{ij} \tag{10}$$

The structure is invariant to the group transformations g defined by:

$$g : \begin{pmatrix} \alpha_i \\ \beta_i \\ k_j \\ \gamma_{j-i} \end{pmatrix} \mapsto \begin{Bmatrix} \alpha_i - d\beta_i - f \\ \dfrac{\beta_i}{c} \\ c(k_j + d) \\ \gamma_{j-i} + f \end{Bmatrix} \qquad (11)$$

where c, d and f are arbitrary constants with $c \neq 0$. To ensure the identifiability of the model, constraints such as defined by Cairns et al. (2009) for the Renshaw-Haberman model may be adopted:

$$\sum_{j=j_1}^{j_T} k_j = 0, \quad \sum_{i=i_1}^{i_I} \beta_i = 1, \quad \sum_{i=i_1}^{i_I} \sum_{j=j_1}^{j_T} \gamma_{j-i} = 0.$$

4.3. Data Description

In this study we use post-1950 data for the Portuguese male population. The data, number of deaths and exposures-to-risk, by ages at last birthday from 0 to 99 years old, were obtained from the Human Mortality Database, which is sponsored by the University of California, Berkeley, and the Max Planck Institute for Demographic Research (http://www.mortality.org). Data were tabulated by age and time of death using equally spaced intervals for age and calendar period. Age-specific death rates are calculated by the ratio of death counts and exposure-to-risk estimates in matched intervals of ages and time. Cohorts are found on the diagonals of the table. One should note that the diagonal cells in such a table, however, do not define non-overlapping cohorts. These correspond to "synthetic cohorts", they do not match to successive observations of the same birth generation but instead to successive age intervals in successive time intervals. If we consider one-year intervals for age and period, each cell represents the experience of two actual cohorts, or the number of deaths during a calendar year between two different birth cohorts. For example, individuals aged 50 who died during the calendar year 2000 were born sometime during the years 1949 and 1950, corresponding to two different birth generations. Similarly, individuals aged 51 belong to the 1948-1949 cohort, which overlaps the previous cohort.

The graphical analysis of the observed death rates provides a first step in better understanding the data (see figure 1.). However this figure may be interpreted with caution since not all cohorts are observed at all ages, and consequently not all contribute with the same information to the mean death rates. Older and younger cohorts are under-represented. The earliest birth cohort in our data corresponds to individuals born between 1st January 1850 and 31st December 1851, those deceased at the age of 99 years old in 1950, the first calendar year of observation in our data. According to the convention adopted in the study, the first cohort (1850-1851) will be designated as 1851 birth cohort, obtained by the difference between year of death and age at last birthday. For this birth cohort only one observation is available. As we move forward we are adding information on more cohorts to the dataset. For these initial cohorts we only have available data for very old ages. As we move forward across cohorts, till the birth year 1908, we are adding information on the death rates of younger ages, which are lower. It follows that the average of the cohort-specific death rates are falling as can be seen in Figure 1. At birth year 1909, we begin to lose the older ages, but keep adding cohorts with younger ages. Therefore, when analysing the mean death rate by cohort (Figure 1) we can detect a more pronounced decline after 1909, as older ages with high
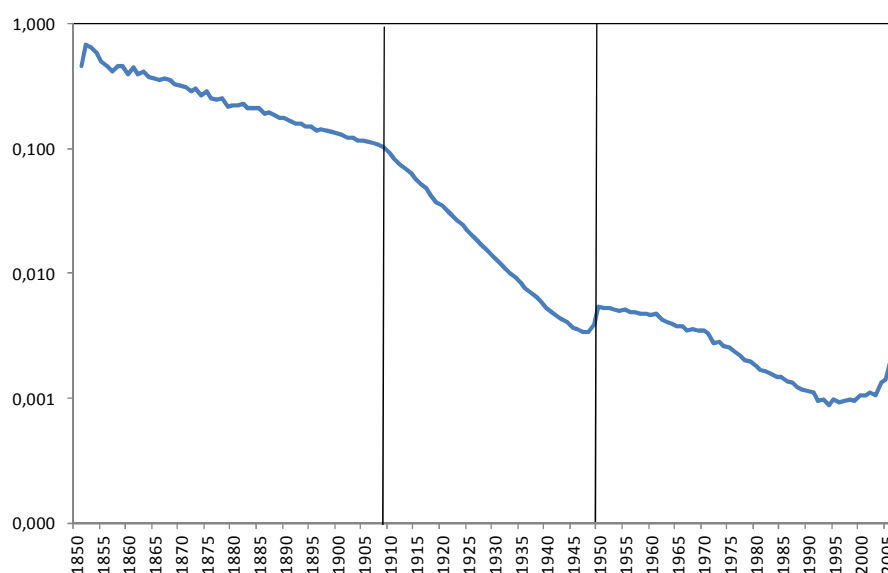

Figure 1. Mean male deaths rates by cohort year-of-birth (unbalanced sample).

death rates are being substituted by younger ages with lower death rates. For the birth years during the second half of 1940s we start adding to the data cohorts with ages under 5, at a time when infant mortality rates are still very high, which results in an increase of the mean death rate till 1950. From birth year 1950 onwards, all cohorts include ages 0 and above, but start losing older ages leading to a decline in the mean death rate. Finally, by the end of the sample, the mean death rate goes up as we have cohorts with data available only for the very young ages.

The graphical analysis of the observed age-specific death rates plotted by calendar year and year of birth offer a better approach for analyzing the effects of age, period and cohort (Carstensen, 2007; Kupper et al., 1985; Roberston and Boyle, 1998b). To facilitate the visual analysis, age-specific death rates for Portuguese males from 1950 onwards were plotted only for five-year age groups and five-year calendar periods.

An analysis of the cross-sectional age-specific deaths rates in Figures 2 and 3 shows that male mortality has improved at all ages in Portugal over the last 60 years, but not all improved at the same pace and a few ages even showed increases in some calendar periods.
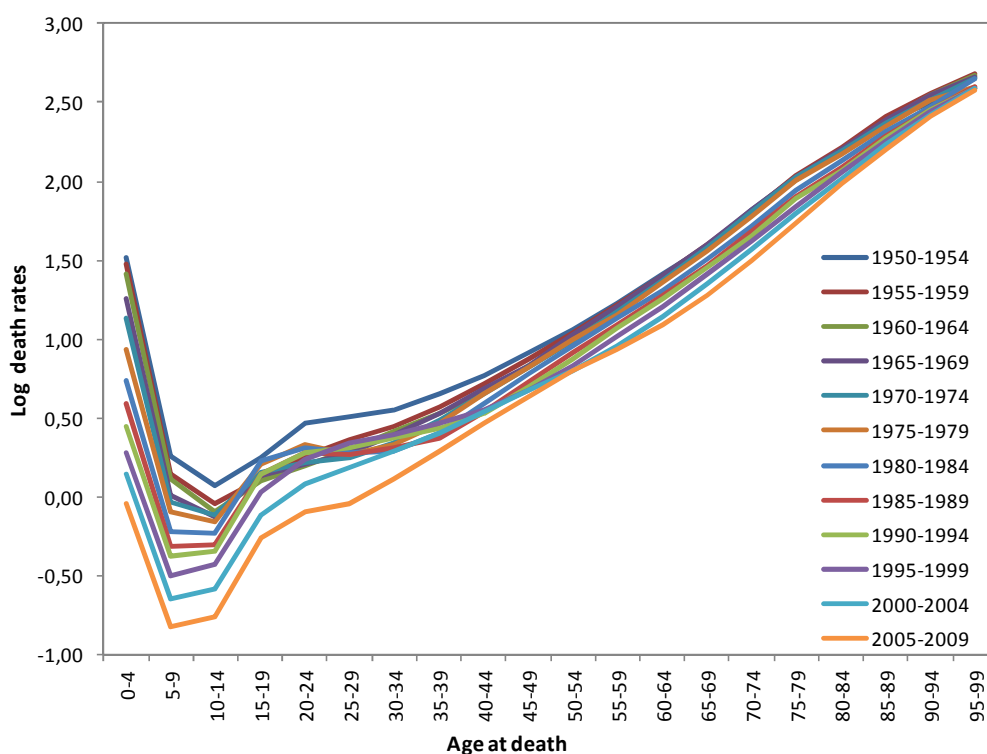

Figure 2. Male age-specific rates by period.

The decline in mortality begins earlier and is more accentuated among younger ages (Figure 3). Death rates from birth to 4 years old show an unprecedented decline and a consistent overall declining trend over time. Although at a slower pace, mortality between ages 5 and 14 also show the same consistent declining pattern over time. In the earlier years, an accentuated decline is showed by mortality between ages 15 and 50, but latter these improvements stabilize and these death rates even increased during some time, returning to a decreasing trend in more recent years. An example of such behavior is the decline in rates at ages 15-19 during the 1950's, after which these begin to increase attaining a peak in 1980-1984, when a decline starts again. The same behavior (of a clear decline and then an increase and then another decrease) can be observed regarding the rates at 20 to 30 years old. This pattern of behavior moves in time as people age, which is apparently an indication of a birth generation of individuals sharing higher deaths rates than previous generations.
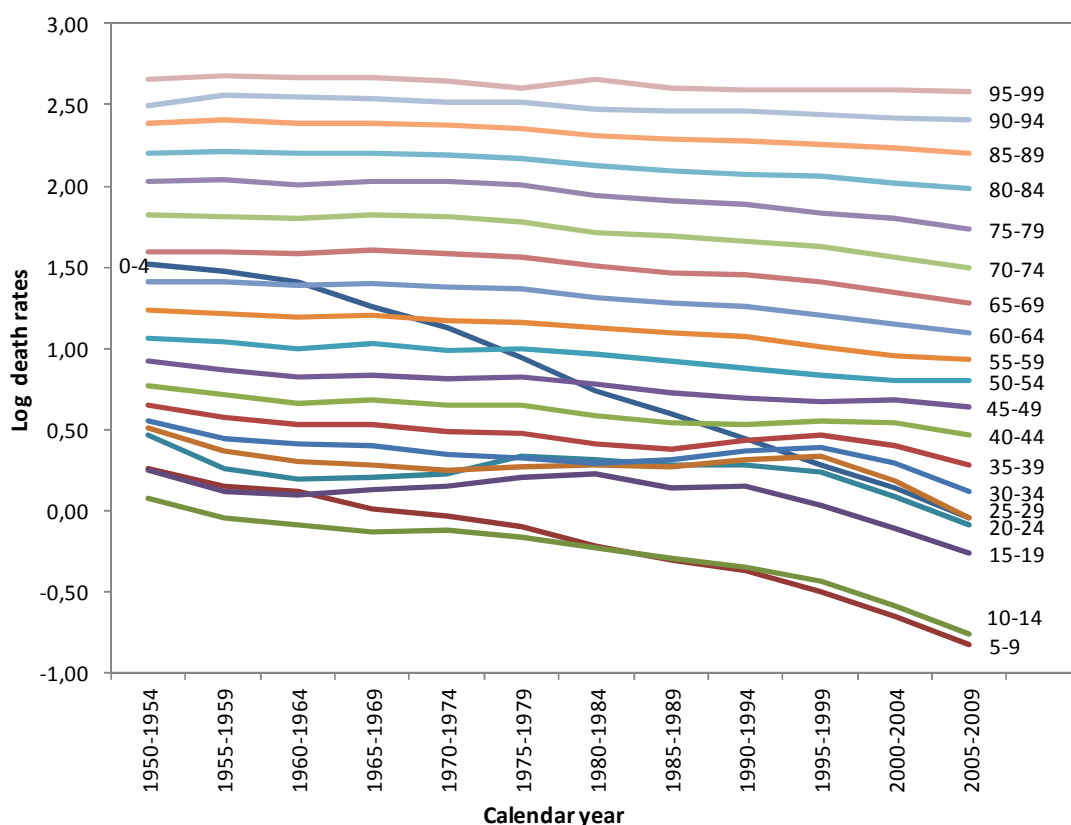

Figure 3. Male period-specific rates by age.

Age-specific death rates curves (Figure 2) move downwards over time, which is an evidence of the influence in mortality of the environmental conditions operating at the time of death; this is, of a period effect. The age distribution however changes from period to period. The shape of age distribution of death rates is affected both by varying age effects and by varying cohort effects, as these curves cut across a number of birth cohorts (Kupper, et al., 1985). After a high incidence of mortality in early childhood, the rates decrease dramatically until about 10-14 years of age when they begin to rise again. The incidence of mortality peaks around the age 20-24, stabilizing till age 35 in the early periods. In the more recent periods 1970-1974 to 1985-1989, after a peak at age 20-24, the incidence of mortality declined between those aged 25-29 and 30-34.

Cross-sectional age-specific death rates changed over time. These changes may reflect aging influences, period effects or it may be the reflection of the different characteristics of the cohorts that happen to be passing through those ages during those calendar years.
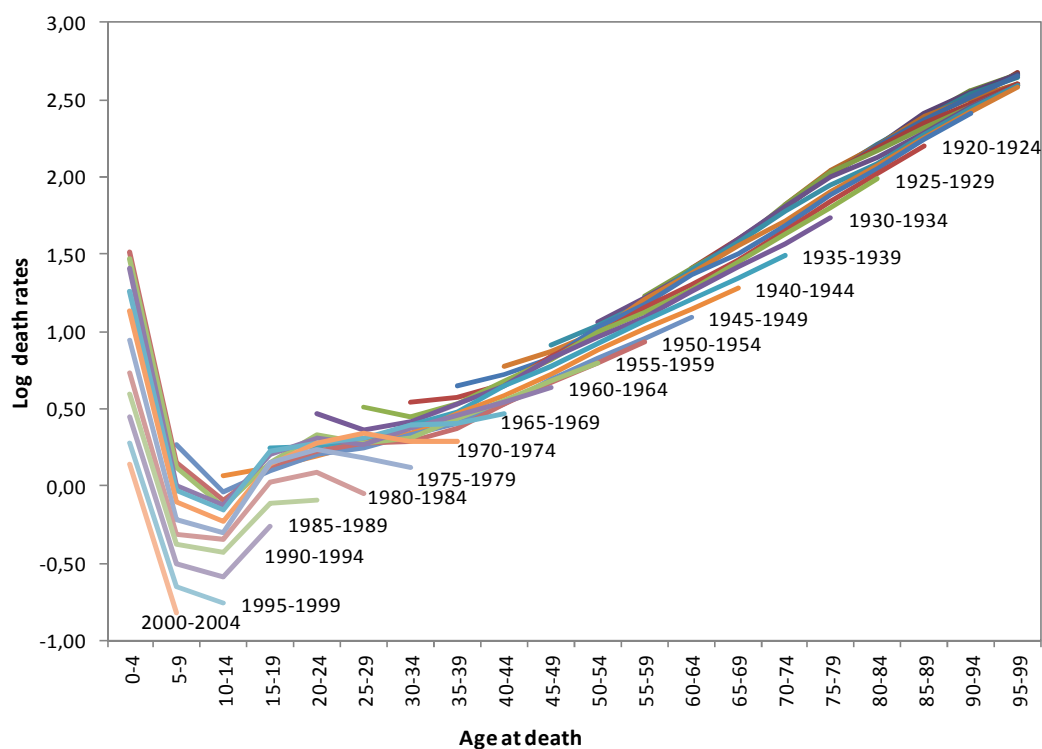

Figure 4. Male age-specific rates by date of birth.

Figure 4 displays the change in mortality of a population born within the same time interval over their lifetime. Following the trend from the early cohorts to the more recent ones, cohorts' life-course patterns show a general consistent behavior. A peak at early childhood is followed by a decrease till around age 10-14, and then by an increase in mortality which attains a peak at around age 20-24, followed by a decline of mortality in each cohort and then by an increase as each cohort ages. The age where the inflexion point between decline and increase happens, however, seems to move to the right. Some specific behavior is visible for the cohorts born between 1945 and 1974 (Figure 5), which comparatively with previous cohorts, show higher mortality at ages 15 – 19 to 35 – 39.
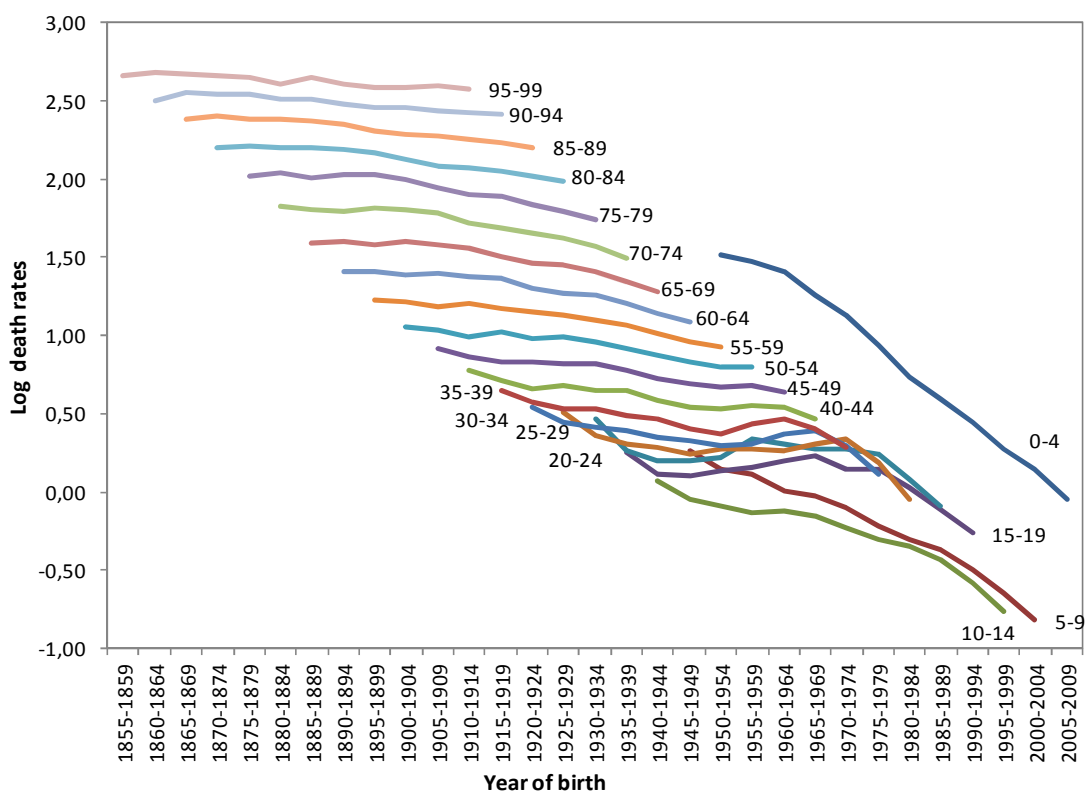

Figure 5. Male cohort-specific rates by age.

The evidence from the graphical analysis is that besides aging changing effects, period and cohort effects may both be present in the data. Since age, calendar time and cohort are fundamentally tied to each other (as time advances, a cohort ages), it is difficult to determine graphically if these three effects are simultaneously present.

However distortions of cross-sectional plot of age-specific death rates are a possible indication of the presence of cohort effects, while distortions in longitudinal plots of age-specific death rates point to the existence of period effects, and both are evident from the plots. Formalizing the age-period-cohort analysis may help to disentangle the influences of each type of effect.

4.4. Estimation Results

In this section we present the estimation results of the various models: the age-cohort (AC) model, the age-period (AP) model[5], the age-period-cohort (APC) model, the Lee-Carter (LC) model, de Lee-Carter with additive cohort effects (LCC), and the Renshaw-Haberman[6] (RH) model for the Portuguese male mortality data. For the models estimation we have considered one-year age and period intervals.

Maximum likelihood estimates for the parameters, for each of the above mentioned models, are obtained based on the following assumptions:

- The number of persons at risk of death at age i and calendar year j, E_{ij}, is a deterministic exposure measure;

- The number of deaths at age i and calendar year j, D_{ij}, are jointly independent random variables following a Poisson distribution with mean $E_{ij} \cdot \mu_{ij}$ or $D_{ij} \sim Poisson(E_{ij} \cdot \mu_{ij})$, where μ_{ij} denotes the age-specific forces of mortality, which are assumed to be constant within bands of age and time but allowed to change from one band to the next;

---

[5] The AC and AP models were fitted using the glm procedure implemented in the Stats R Package. The statistical software "R" is available free of charge from www.r-project.org.

[6] The APC, LC, LCC and RH models were fitted using an adaptation of the code, written by Professor Andrew Cairns in the "R" programming language, which is part of the LifeMetrics toolkit (Coughland, G., Epstein, D., Ong, A., Sinha, A., Hevia-Portocarrero, J., Gingrich, E., Khalaf-Allah, M., and Joseph, P., 2007).

- With a log-link function, the logarithm of the expected number of deaths is given by $\log E(D_{ij}) = \log E_{ij} + \log(\mu_{ij})$, where $\log(\mu_{ij})$ correspond to the parametric structure of $E(y_{ij})$ for each of the models considered.

4.4.1. The age-period and age-cohort models

As a first step in the analysis we fitted an age-period and an age-cohort models, which give a rather simplistic description of the actual impacts of risk factors on mortality. The deviance residuals from both models are plotted against age, calendar year and year of birth (Figure 6). The plot of deviance residuals versus age from both models shows large residuals at childhood, denoting the difficulties in capturing the trends of mortality at these ages. The dramatic decline of death rates at the youngest ages, from very high rates at the 1950s to very low rates at the first years of the twenty-first century, is not well captured in particular by the age-period model. The plot of deviance residuals against calendar year shows large positive deviance residuals for the early calendar years and large negative deviance residuals in more recent years.

The deviance residuals plotted against year of birth show a sort of ripple effect. It is more visible for the age-period model, showing larger deviance residuals for the generations born after 1940. This pattern may suggest the need of a cohort effect in explaining the mortality behavior. The ripple effect for the cohorts born between 1950 and around 1975 is also evident in the plot of deviance residuals from the age-cohort model. To check whether this apparent cohort effect is not caused by the particular structure of our dataset where higher infant death rates are first included in the year 1950, we re-estimated these models excluding ages 0-2. The results obtained still show a visible ripple effect in the residuals, symptom of the need to model a cohort effect.

The age-cohort model seems to capture the trends in mortality more effectively than the age-period model, reducing the deviance substantially comparatively to the age-period model. Neither, however, seem to be the appropriate models to describe the trajectory of male Portuguese mortality, since clear non-random patterns subsist in the residuals.
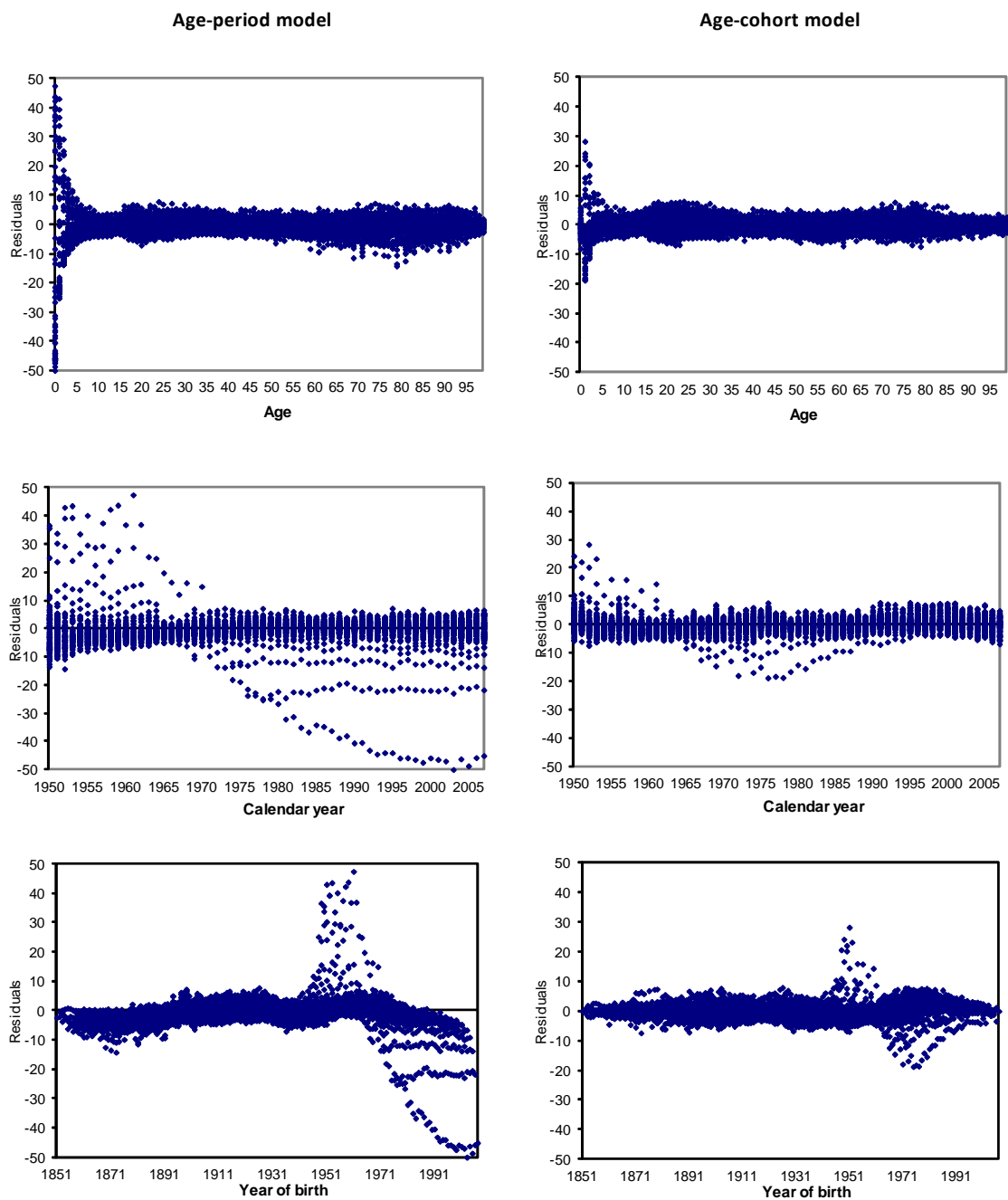
Age-period model

Age-cohort model

Figure 6. Age-period and age-cohort models deviance residuals plots against age, period and cohort year-of-birth, males, Portugal, 1950 – 2007.

4.4.2. The age-period-cohort model

A more realistic assumption will be that neither period nor cohort effects alone lead to an adequate description of the data. The age-period-cohort model, which combines the cohort and the period approaches, will allow a simultaneous analysis of both dynamics in the mortality trends. We fit the age-period-cohort model to mortality data using the approach proposed by Kuang et al. (2008a). The second differences of age, period and cohort are estimated directly.

Following the recommendations of Carstensen (2007) and Kuang et al (2008 a,b) on the parameterization of the effects, we look at the estimated parameters, and in conjunction with information external to the data on the mortality behavior patterns, we set the constant, the first age, cohort and period parameters equal to zero. Additionally we set an equality constraint on the first two calendar year parameters. The selected parameterization is based on the predominance of aging influences on all causes mortality (Hobcraft et al, 1982; Clayton and Schifflers, 1987), and on the declining time pattern of overall mortality in developed countries over the twentieth century associated to period influences, both important in explaining mortality decline in Portugal. The interpretation of the parameters is conditioned by the adopted parameterization.

Figure 7 shows the estimated age, period, and cohort parameters. The estimated age effects profile corresponds to the familiar shape, showing higher mortality levels at childhood and falling rapidly afterwards and then increases showing a clear accident hump, and then increasing as age advances. The period estimated parameters show a declining trend over time with several oscillations. The cohort parameters show an increase in mortality among generations born in the late 1940s and during the 1950s.
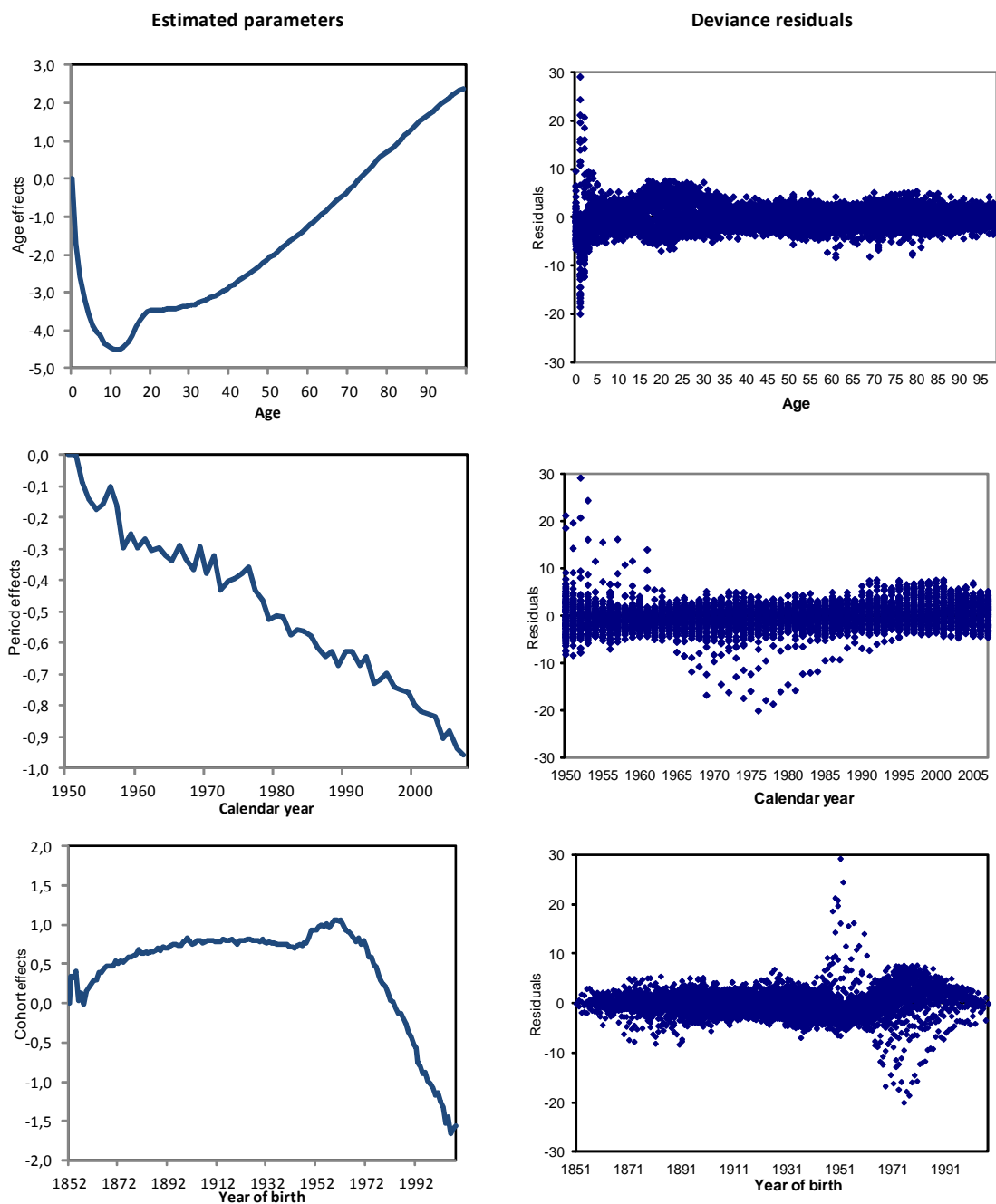
Figure 7. Age-period-cohort model: estimated parameters α_i, β_j, and γ_{j-i}, and deviance residual plots against age, period and cohort year-of-birth, males, Portugal, 1950 – 2007.

Although the fit of the age-period-cohort model improved over the age-period and age-cohort models (Table 1), the analysis of the residuals of the age-period-cohort model (Figure 7) still shows some difficulties in capturing the trend in mortality at childhood.

Table 1. *Summary Measures on the Estimated Models*

MODEL	AP	AC	APC	LC	LCC	RH
N.º of observations	5800	5800	5800	5800	5800	5800
N.º of parameters	157	256	312	256	412	511
Log-likelihood	-98252	-42929	-40714	-33423	-29304	-27547
BIC	197864	198722	84132	69064	63036	59522
Deviance	153583	42937	38509	23925	15687	12173

The age-period-cohort model provides a more adequate description of the data than a model that considers only period or cohort effects. However it is still just an approximation for explaining the influences of age, period and cohort risk factors on mortality. Being an additive model it is based on the assumption that the influence of age is the same in all time periods and for all cohorts; that the influence of period is the same for people of all ages; and that cohorts do not change over time. However, period and cohort influences in practice however have proven to operate differentially at different ages (Schulhofer-Wohl and Yang, 2011). The historical behavior of overall mortality tells a somewhat different story, namely that the influence of age changes over time and across cohorts. An obvious example is the huge decreases in infant mortality over the past century, with substantially different rhythms of decline over time, not only due to cohort effects but mainly influenced by environmental factors associated to the time of death.

4.4.3. The Lee-Carter model

The possibility of an interaction effect of age with period suggested by the graphical analysis may be verified by estimating the Lee-Carter model. Instead of considering period effects that have the same impact at all ages, as in the age-period model, the Lee-Carter model introduces an interaction term of age with period.

The results of the Lee-Carter model applied to Portuguese mortality are presented by Coelho and Nunes (2011). The age effects $\hat{\alpha}_i$ (Figure 8) show the general

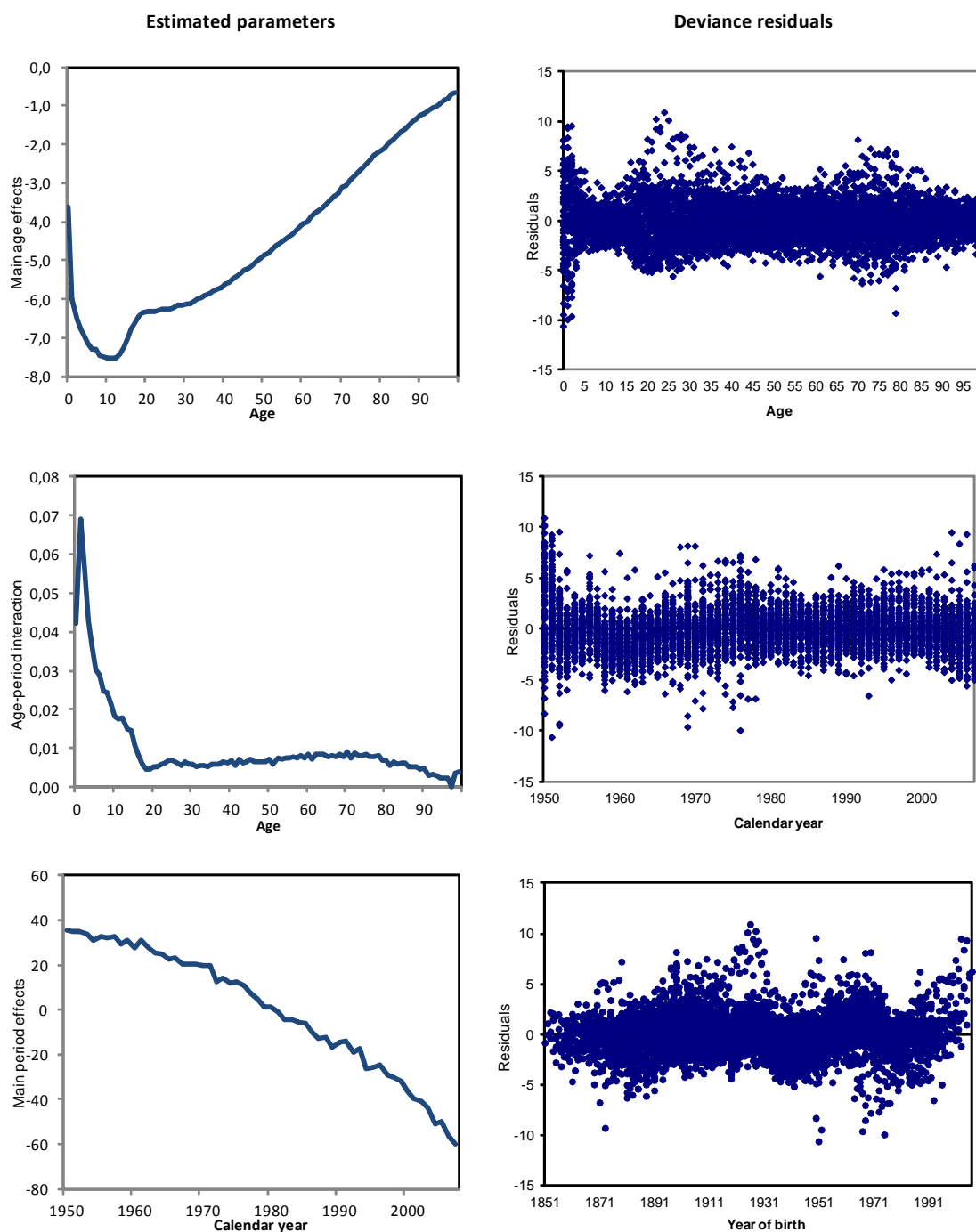**Estimated parameters**   **Deviance residuals**

*Figure 8.*Lee-Carter model: (1) estimated parameters (main age effects α_i, age-period interaction effects β_i, and main period effects k_j), and (2) deviance residual plots against age, period and cohort year-of-birth, males, Portugal, 1950 – 2007.

shape of the age specific death rates, with relatively high values at the infant and childhood ages, declining very rapidly with age till around ages 10 to 12, at which it attains its minimum. And then the $\hat{\alpha}_i$ increases to a local maximum around the late teens and early twenties (accident hump), after which it follow a nearly linear increase into the middle and older ages. The $\hat{k}_j$ (Figure 8) declines almost linearly over time, suggesting a more accentuated rate of decline after the mid-1970s, and driving all age-specific rates down, but not by the same amount due to the different values of $\hat{\beta}_i$ (Figure 8). Infant ages and childhood show relatively high $\hat{\beta}_i$ values, declining to a minimum value around age 20. A slight increase of $\hat{\beta}_i$ are visible till around age 70, when it declines again.

The introduction of an interaction term of age with period reduces the deviance residuals (Table 1), and the Lee-Carter model represents a fitting improvement over the age-period, age-cohort and age-period-cohort models. The residuals plotted against age (Figure 8) still shows larger residuals at infancy and around the accident hump ages; however the amplitude of variation is substantially inferior to the age-period model residuals. The deviance residuals plotted against year of birth however still show a clear ripple effect, suggesting a cohort effect. Comparatively to the age-period-cohort model, the Lee-Carter model provides a better description of the male Portuguese mortality but deviance residuals show an existing pattern, which may suggest that not only period effects differ between ages but also a cohort effect may be influencing the mortality pattern.

4.4.4. The Renshaw-Haberman model

The Renshaw and Haberman (2006) model includes interactions of age with period and with cohort effects. The main age effect plot $\hat{\alpha}_i$ (Figure 9) exhibits the familiar pattern. The main period effect $\hat{k}_j$ (Figure 9) declines almost linearly over time, although with some small oscillations. The age-by-period interaction term β_i^1
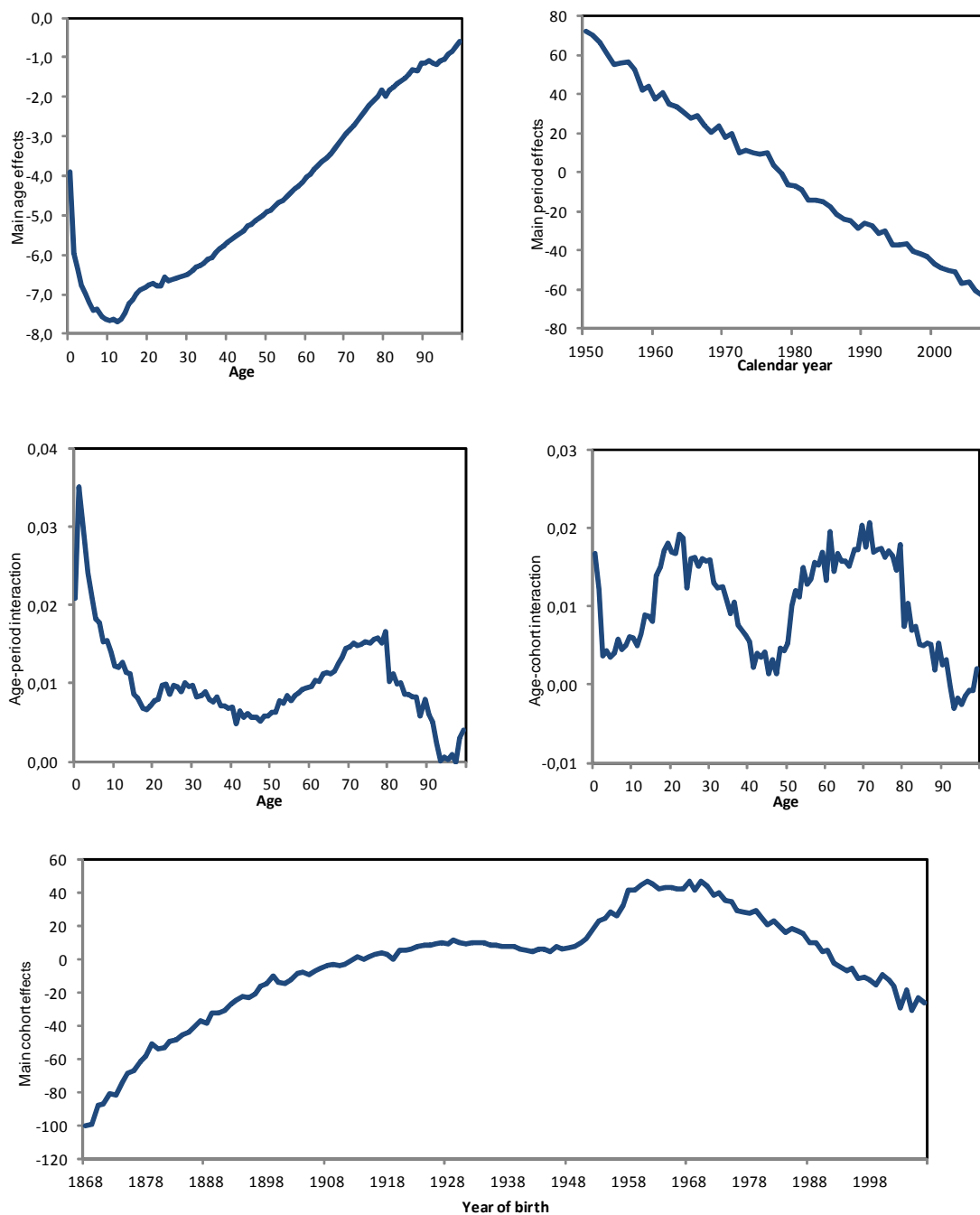
Figure 9. Renshaw-Haberman model estimated parameters: main age effects α_i, age-period interaction effects $\beta_i^{(1)}$, main period effects k_j, age-cohort interaction effects $\beta_i^{(2)}$, and main cohort effects γ_{j-i}, males, Portugal, 1950 – 2007

(Figure 9) pattern, which represents the rate of decline by period of age specific rates of mortality, show that even though all age-specific rates have declined with time, some

have had larger improvements, in particular those at childhood and among ages 65 to 80 years old.

The plot of the cohort effects (Figure 9) exhibit a pronounced increase in the $\hat{\gamma}_{j-i}$ profile for the generations born in the late 1940s, during the 1950s and the first years of the 1960s, remaining relatively stable over the 1960s. This behavior is consistent with the increases in death rates observed in the exploratory analysis for these cohorts. Generations born after 1970 seem to benefit from mortality improvements relatively to previous generations. The age-by-cohort interaction term $\hat{\beta}_i^2$ (Figure 9) pattern, which represents how the cohort effects change with age, show that the influence of cohort effects is more pronounced at infancy, between ages 15 to 35 years old and after 60 and below 80 years old.

The decline in mortality at older ages is represented by the combination of negative main cohort effects, $\hat{\gamma}_{j-i}$, in the oldest generations, which are highly sensitive to age (large positive values of $\hat{\beta}_i^2$). The estimated behavior suggests that the oldest cohorts have been able to improve mortality rates as they age.

Cohorts born after 1913 and before 1990 show positive main cohort effects, $\hat{\gamma}_{j-i}$, stronger, as already mentioned, among those born between 1950 and 1970. These are associated to relatively larger values of $\hat{\beta}_i^2$ around the accident hump ages, suggesting that these cohorts experienced worse mortality conditions than previous cohorts at the same ages.

As before, we also re-estimated the Renshaw-Haberman model excluding ages 0-2. The results obtained confirm the presence of cohort effects for the same generations.

The analysis of the residuals plots of the age-period-cohort model with age interactions (Figure 10) shows that the ripple effects in the year-of-birth residuals plots under the age-period and the Lee-Carter model are largely removed, suggesting the relatively success in capturing the mortality pattern and an improvement over the pure age-period-cohort model and the Lee-Carter model.
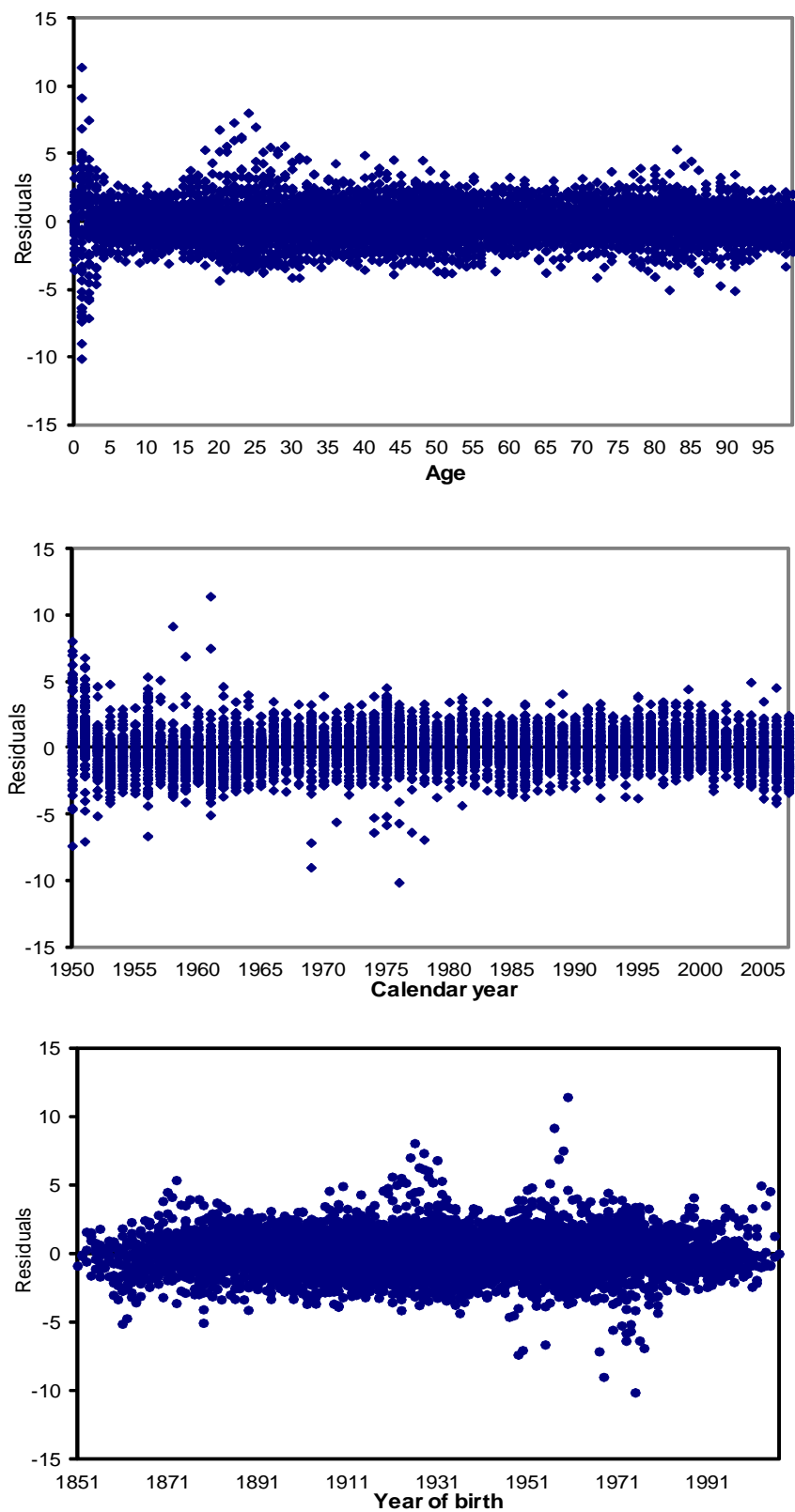
Figure 10. Renshaw-Haberman model deviance residual plots against: age, period, and cohort year-of-birth, males, Portugal, 1950 – 2007.

4.4.5. The Lee-Carter model with additive cohort effects

Finally we have also fitted the Lee-Carter model adding additive cohort effects. The estimated parameters are presented in Figure 11. The age effects $\hat{\alpha}_i$ show an identical profile to those estimated by LC and RH models. The $\hat{k}_j$ however is closer to the estimated main period effects from the RH model, denoting an almost linearly behavior over time. Also the age-period interaction profile, $\hat{\beta}_i$, show a profile closer to the one estimated by the RH model, with higher values at infant and childhood ages, but
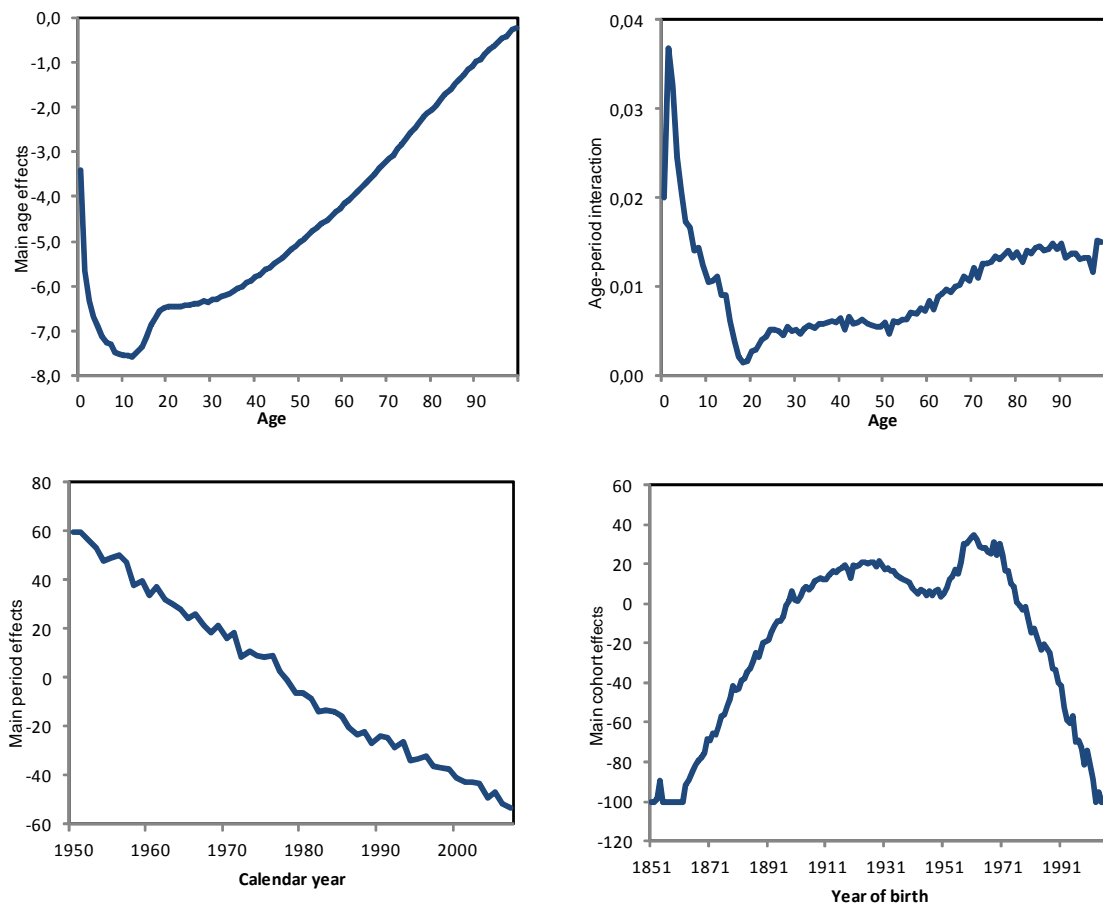

Figure 11. Lee-Carter with additive cohort effects model estimated parameters: main age effects α_i, age-period interaction effects β_i, main period effects k_j, and main cohort effects γ_{j-i}, males, Portugal, 1950 – 2007.

lower than the age-period interaction parameters estimated by the LC model. In the estimated cohort effects, $\hat{\gamma}_{j-i}$, are also evident worse mortality conditions for those born between 1950 and 1970, and comparatively with the RH model, a more accentuate decline for the more recent cohorts.

The introduction of additive cohort effect reduces the deviance residuals (Table 1), and the model represents a fitting improvement over the LC model. The plots of the deviance residuals show very similar patterns to those of the Renshaw-Haberman model, but with a higher variability.
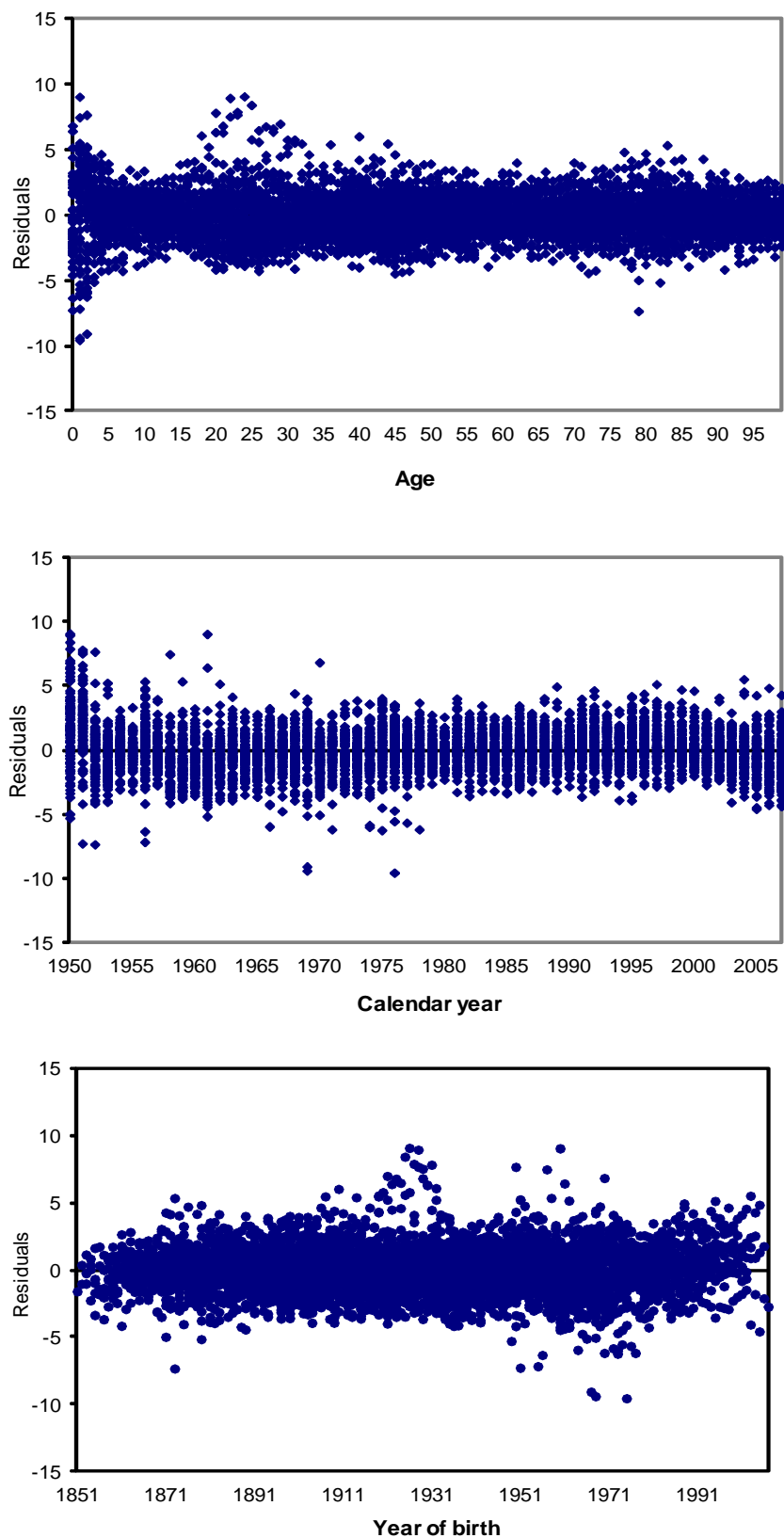
Figure 12. Lee-Carter with additive cohort effects model deviance residual plots against: age, period, and cohort year-of-birth, males, Portugal, 1950 – 2007.

4.4.6. Model selection

Although the analysis of the deviance residuals allows us to compare the fitting quality of the different models, it is natural for models with more parameters to fit the data better. The Bayes Information Criterion (BIC) is used to select the a parsimonious model that adequately describes the data (Cairns et al. 2009). BIC compares maximum likelihoods attained by each model but penalizes models that are overparameterized. The BIC is calculated as (Cameron and Triverdi, 1998) :

$$BIC = -2l(\hat{\phi}) + v \log n \qquad (11)$$

where $\hat{\phi}$ is the maximum likelihood estimate of the parameter vector, $l(\hat{\phi})$ is the maximum log likelihood, n is the number of observations and v is the effective number of parameters being estimated. Models can be ordered, with the most appropriate model having the lowest BIC.

Table 1 show some summary measures on the estimated models. According to the value of BIC, the Renshaw-Haberman model is the most suitable model to describe Portuguese male mortality patterns from 1950 to 2007. This result is coherent with our expectations from the exploratory analysis of male mortality pattern in Portugal over 1950 – 2007, showing different rates of improvement for different ages, and the fact that some birth cohorts showed particular poor mortality improvements over adulthood when compared to other generations.

4.5. Simulation Study

The evidence of the graphical analysis as well as the results from the estimated models allude to the existence of cohort effects influencing Portuguese male mortality data. The most adequate model to describe mortality patterns was selected based on BIC rankings, and we have selected a model with both age and cohort effects as well as the respective age interaction terms.

To make sure that one is choosing the appropriate model it is essential to determine whether BIC is actually able to select the best model. To do that we conducted a simulation study that should help us determine if BIC is choosing the

appropriate model for different data generating processes (DGP). Through the simulation study it is also possible to understand what happens when the fitted model does not correspond to the true model, that is, what are the consequences of an incorrect specification of the model.

Simulation settings are chosen such that they mimic real world situations, in particular Portugal and England and Wales[7].

Portugal is an obvious choice since we want to guarantee that we are choosing the most appropriate model to describe male Portuguese mortality. In England and Wales, on the other hand, there are important and well-document cohort effects (Willets, 1999; Willets, 2004; Renshaw and Haberman, 2006; Richards, Ellam, Hubbard, Lu, Makin and Miller, 2007). Details on mortality trends in England and Wales are discussed by Willets (1999; 2004). Although a decline in mortality have taken place in both countries over 1950 – 2007, one should note that, at the beginning of the period these countries were in quite different stages in mortality progress, with Portugal showing comparatively poor mortality conditions. Over the second half of the 20th century, however, we have observed a remarkable declining trend in overall Portuguese mortality levels, converging to similar mortality levels of England and Wales. The pattern of age-specific rate of decline of both countries is however very different. While Portugal shows greater improvements over time at children and younger adults, England and Wales show the greater improvements at adult and senior ages. These different patterns in mortality improvements are evident from the estimated parameters of the LC e RH models, shown in Figures 8 and 9 for Portugal and provided by Renshaw and Haberman (2006) for England and Wales.

Our simulation strategy is described by the following steps. In the first step, assuming a Poisson distribution for the number of deaths, D_{ij}, we estimate a model that is used to generate random samples in the second step.

In the second step, we simulate the number of deaths from a Poisson with a mean computed using the estimated parameters from the model estimated in step one. For each DGP we generate 1,000 simulated random data sets.

---

[7] Data for Spain is also used in some of the simulations as the demographic context of Spain is closer to the Portuguese.

In the third step, for each simulated data set, we estimated each of the above mentioned models (APC, LC and RH), obtain estimates of the respective parameters and the value of BIC, and select the best model.

In the fourth step, for each DGP, we compute the number of times each model is selected by BIC. The results are presented in Table 2.

We considered the following data generating processes (DGP): (1) Age-Period-Cohort (APC) model, (2) Lee-Carter (LC) model, and (3) Renshaw-Haberman (RH) model. Since in step 1 of the simulation we may be fitting a model that is not the most adequate to describe the data, we also considered the following alternative to determine the values of the parameters of the first two DGPs. In particular in step 1 we estimated the RH model, which includes all other models as special cases, imposing in step 2 the following constraints to generate the random samples: (1) for the APC model we impose the age-interaction terms equal to a constant, in particular $\beta_i^1 = \beta_i^2 = 0.01$ [8], (2) for the LC model we impose the cohort effect equal to zero, $\gamma_{j-i} = 0$. We also considered a fourth DGP setting the estimated age-cohort interaction term in the RH model equal to a constant, $\beta_i^2 = 0.01$, which corresponds to the LC model with additive main cohort effects.

Besides BIC, we have also analyzed the estimated parameters obtained from each simulation, in particular their mean and percentiles over the 1,000 replications.

4.5.1. Data generating process: Age-period-cohort model

Using as DGP the APC model, either using the parameters estimated directly by the APC or the RH model with restrictions, as expected the model most often selected by BIC is the APC model.

The RH model provides the second most adequate model, however the estimated age-period, $\hat{\beta}_i^1$, and age-cohort, $\hat{\beta}_i^2$, parameters do not show any clear pattern denoting

---

[8] We set the constant equal to 0.01 given the identification condition , $\sum_{i=i_1}^{i_I} \beta_i^{(2)} = 1$, and the fact that we have considered one hundred single ages.

a random behavior. This is a sign that the RH model is an over parameterized model to fit the data.

The LC model presents the weakest performance in terms of BIC. Looking at the estimated parameters of the LC from the simulation scenarios of Portugal and England and Wales, even though the DGPs do not include any age-period interaction effects, the estimated age-period interaction parameters, $\hat{\beta}_i^1$, (Figure 13) show a clear non-random profile with the greatest values at the younger ages and then showing a declining smooth trend till around age 40, when it increases slightly for the age interval 50 to 70 years old, returning to the declining trend after that. The pattern of $\hat{\beta}_i^1$ from the Lee-Carter model tries to capture the cohort effects in the simulated data.
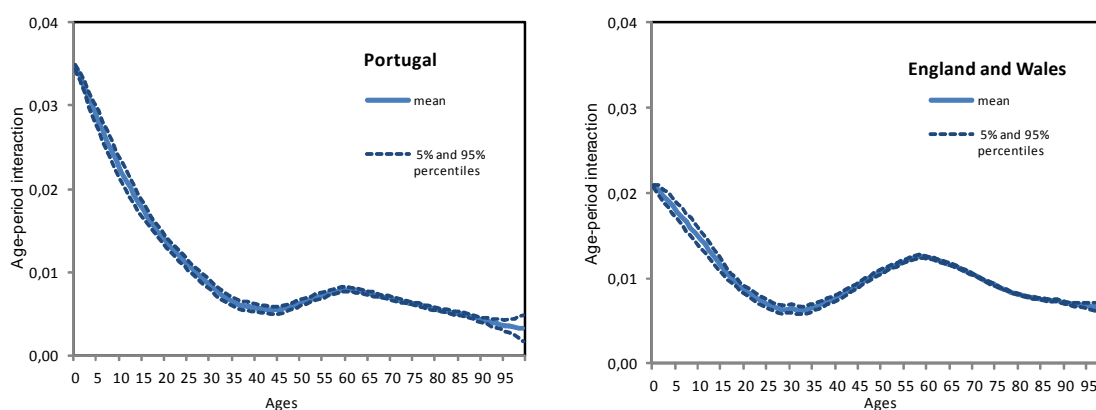
Figure 13. DGP APC model – LC model simulated age-period interaction effects, males, Portugal and England and Wales, 1950 – 2007.

The results suggest that if there are only additive cohort effects in the data, allowing for cohort effects interacting with age in the estimated model will result in an over parameterized model. On the other hand, if we estimate a model without cohort effects but with an age-period interaction term, as the LC model, the result will be a spurious age-period interaction term.

As a final point, we note that if we use as DGP an age-cohort model, if we fitted the LC model using the simulated data we get estimated age-period interaction effects with a similar profile to the ones estimated above for the APC model as DGP.

4.5.2. Data generating process: Lee-Carter model

Using as DGP the LC model, either using the parameters estimated directly by the LC model or the RH model with restrictions, as expected the Lee-Carter model is the model most often selected by BIC.

Due to the age-period interaction term, the RH model is the second most adequate model, and the estimated parameters, $\hat{\alpha}_i$, $\hat{\beta}_i^1$, and $\hat{k}_j$, show the typical pattern, while the estimated age-cohort interaction parameters $\hat{\beta}_i^2$ and the main cohort effect parameters $\hat{\gamma}_{j-i}$, both show a random behavior, suggesting that the RH model is over parameterized.

The APC model is the third choice according to BIC. The age and period effects showed the traditional patterns. However, the graphical profile of the estimated cohort effect suggests the presence of a cohort effect, which is not present in the simulated data. For the simulation scenario of Portugal, the cohort effect profile (Figure 14) shows a more rapid increase in mortality for the generations born between 1947 and 1952 when compared to previous generations.
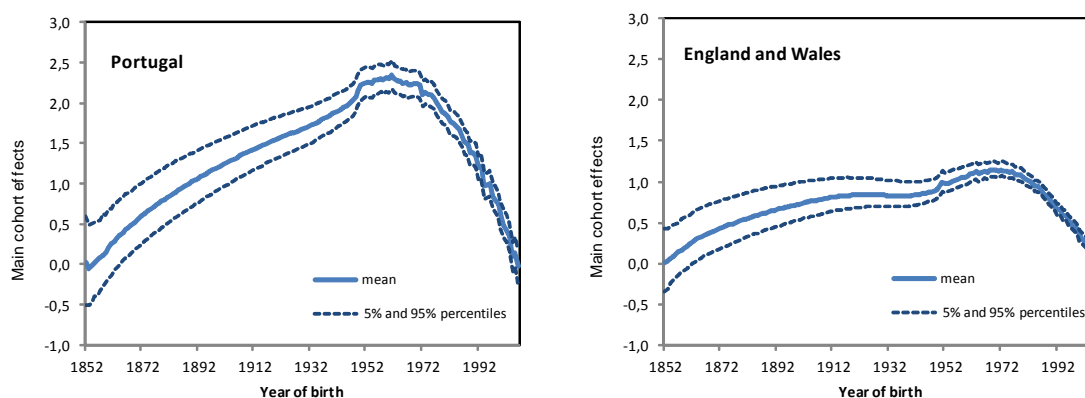

Figure 14. DGP Lee-Carter model – APC model simulated cohort effects, males, Portugal and England and Wales, 1950 – 2007.

Higher levels of mortality are shared by generations born during the 1950s. The top of the hump is located precisely when the data on mortality at youngest ages begin to be incorporated in the data, during the decades of 1940 and 1950, when generations under

study begin to comprise younger children and finally the generation born in 1950 include infants, ages where the death rates were particular high, and with a great sensitivity to the declining temporal trend. The simulated cohort effects for England and Wales shows a similar behaviour, although much smoother, which would be associated to the lesser sensitivity of age-specific mortality to the temporal trend.

In summary, the results of the simulation suggest that if we have data without a cohort effect but with an age-period interaction term, if we fit an APC model, the age-period interaction in the data is translated into an apparent cohort effect, and its magnitude depends on the strength of the age-period interaction term. Also, the location of the estimated cohort effect coincides with the period when the youngest cohorts enter the data. Introducing the young cohorts in the data, with a high sensitivity to time declining trend induces a spurious estimated cohort effect in the APC model. This conclusion is supported by the fact that if we use an AP model with no age-interaction (DGP) to generate the data, the APC model does not find a cohort effect, showing only the expected declining period effect.

4.5.3. Data generating process: Renshaw-Haberman model

Using as DGP the RH model, the simulated data for the two scenarios, Portugal and England and Wales, have main age, period and cohort effects, as well as age-period and age-cohort interaction terms, meaning that the impact of period effects and cohort effects are different for different ages. The simulated data for the two scenarios differ in the profile of age-interaction terms as well as in the strength of the main effects. The age-interaction effect is stronger for Portugal whereas the cohort effect is more pronounced in the case of England and Wales. According to the BIC criterion, the RH model is the model selected more often for both scenarios.

In the simulation scenario for Portugal, the LC model proves to be the second most adequate, followed by the APC model. In this scenario, there exists a strong age-period interaction, especially at younger ages, and the evidence of a rather weak cohort effect. So, between a model with no age-interactions and a cohort effect and a model

without cohort effect but with an age-period interaction, the LC model is chosen over the APC model[9].

In the simulated data for England and Wales, the APC model dominates the LC model with a lower BIC, a result that may be connected to the higher importance of cohort effects in England and Wales.

The relative importance of the age-period interaction and the cohort effects influences the model fitting results, and the choice of the most appropriate model.

4.5.4. Data generating process: Lee-Carter model with additive cohort effects

Finally, we consider as DGP the LC with additive cohort effects parameters obtained in the estimated RH model setting the age-cohort interaction parameters equal to a constant.

Table 2. *Simulation Summary Results*

DGP	Portugal			England and Wales		
	1st	2nd	3rd	1st	2nd	3rd
APC	APC	RH	LC	APC	RH	LC
APC (restricted RH)	APC	RH	LC	APC	RH	LC
LC	LC	RH	APC	LC	RH	APC
LC (restricted RH)	LC	RH	APC	LC	RH	APC
RH	RH	LC	APC	RH	APC	LC
LCC (restricted RH)	RH	LC	APC	RH	APC	LC

Note: In all cases the model the BIC selects more often (in the 1st column) is chosen in more than 99% of the times.

---

[9] If we use data for a mortality scenario close of the one of Portugal during the period 1950 – 2007, as is the case of Spain, we find the same results. The relative importance of the age-period interaction and period effects, stronger in Portugal and Spain (countries with similar demographic background) makes the LC model to capture the mortality pattern better than the APC. Cleries, Martínez, Valls, Pareja, Esteban, Gispert, Moreno, Ribes and Borràs (2009) discuss a cohort effect in Spain, and find a rising trend in mortality among the 1940 – 1970 cohorts.

According to the values of BIC, the model that is selected more often is the RH model. In the second place, for the scenario of Portugal, is the LC model, whereas for England and Wales's scenario is the APC model.

These results suggest that the chosen model depends on the relative importance of the cohort effect versus the age-period interaction effects.

4.6. Discussion of Results and Conclusions

This study focused on Portuguese male mortality post-1950. We conducted an APC analysis based on the combination of graphical descriptive analysis and formal modelling techniques.

The descriptive analysis of male mortality rates over age, year at death and year at birth show that overall mortality has improved substantially since 1950, but age-specific deaths rates have declined at different rhythms. In particular, male population born during the late 1940s and during the 1950s have experienced particular poor improvement in mortality, and even increased mortality rates, when compared to adjacent generations. Such type of pattern, in which some generations show specific conditions of mortality, either faster or slower improvements in mortality, when compared to adjacent generations is associated to the presence of cohort effects in mortality, which may be birth generation effects or due to the exposure to external causes over some period of their lives. Willets (1999, 2004) found cohort effects for people born in England and Wales between 1925 and 1945. Renshaw and Haberman (2006) have also demonstrated that an adequate modelling of past mortality patterns in England and Wales needs both period and cohort effects. Among others, cohort effects were also found in Japan, Austria and France (Willets, 2004; Vaupel and Andreev, 2005).

A modelling approach of mortality dynamics is used to separate the different sources of variation in deaths rates. We have fitted the classical APC model, using a recently developed method of estimation proposed by Kuang et al. (2008), and APC models allowing for age-period and age-cohort interactions proposed by Renshaw and Haberman (2006) and further explored by Cairns et al. (2009). The estimation results

show evidence of significant age-period interaction patterns and a declining time trend affecting overall mortality, but also evidence of the presence of cohort effects in male Portuguese mortality. According to the analysis of the residuals and the BIC criterion, the Renshaw-Haberman model, which includes age, period and cohorts as well as age-period and age-cohort interactions, was selected as the model that most adequately described male mortality patterns in Portugal since 1950. The profile of the estimated cohort effects shows a pronounced increase during the 1950s with particular incidence in the young adult ages, which is consistent with the results from the graphical analysis of the specific deaths rates in this study, and with the findings of Coelho and Nunes (2010), Lages (2007) and Morais (2002).

A simulation study is conducted to confirm if BIC is a satisfactory criterion to select a parsimonious model that most adequately fit the data. The simulation results demonstrate that BIC can be used to choose the correct model, and these also allow us to verify the impacts of an incorrect specification of the model on the estimated parameters. Specifically, if only additive cohort effects are present in the data, allowing for age interaction terms will result in an over parameterized model. On the other hand, if we omit cohort effects and estimate a model with an age-period interaction term, we obtain spurious age-period interaction terms. If only age-period effects are present in the data, allowing for age-cohort interaction term will result in an over parameterized model. On the other hand, the omission of an age-period interaction results in spurious cohort effects.

These findings highlight the potential importance of incorporating cohort effects for building better models of mortality temporal dynamics. Not only cohort effects may be important for clarifying recent mortality patterns, but also as a tool for forecasting future mortality patterns.

The results of the estimated models confirm the results obtained by the analysis of the evolution of death rates, suggesting a cohort effect in male Portuguese mortality, specifically that men born at the latest 1940s and during the 1950s seem to have experienced poorer improvements in mortality than generations born either side of this period. These findings give rise to questions on why these cohorts constitute a disadvantaged group of individuals before death, and if this phenomenon represents a

lasting characteristic of mortality patterns in such cohorts or otherwise it will fade with age.

In what concerns understanding and explanations for such cohort effects, as Hobcraft et al. (1982) argue, age, period and cohorts do not directly determine demographic outcomes, but instead are markers for underlying biological, epidemiological, social, and economic factors. Period is a proxy for contemporaneous influences, while cohort is a proxy for influences in the past. Factors such as early life conditions, improved nutrition, social and economic conditions, life styles and cultural behaviors are some of the factors likely to produce cohort effects. Cohorts may be differentiated from all others because different cohorts live through different experiences, or live through the same experiences at different ages, and these experiences change the cohort in long-term ways (Schulhofer-Wohl and Yang, 2009).

Although we have not found any explicit study on cohort effects on Portuguese mortality, Lages (2007), while studying the risk behaviors of young population in Portugal and their mortality, found an increase in mortality among young men after 1960-1961 that extend further in time to adults and later up to 50 years old. According to Lages (2007), the profile of mortality in 1960-1961 was still the result from natural conditions, showing a high infant mortality, a fairly regular growth of mortality with age, after preadolescence, and irregularities in the improvement rates at older ages. Mortality affected all cohorts in a predictable and undifferentiated way. In 1970-1971, however, the author identifies a slight increase of male mortality between ages 16 and 30. What was a small increase in 1970-1971 outspreads in intensity and extend to older ages. In 1980-81, this increase in mortality intensifies after age 16 and extends until around age 35. In 1990-1991, the same pattern of increase of mortality is found, but seems to extend much longer, up about 50 years old. One notes that the age groups referred by Lages (2007) over subsequent decades correspond to men born between the beginning of the 1940s and the middle 1960s.

Lages (2007) concludes that factors that led to the emergence of these worse conditions of mortality, that keep affecting men into the middle-age, can only be associated to behaviors that do not fall under natural causes or random mortality, but instead associated to factors arising from group cultures that increasingly extended further in time. Also, he argues that the increase of mortality among young individuals

that extend further in time are not circumstantial, but instead have structural features, since it subsists in the cohorts that experienced poorer mortality conditions in more recent years.

Morais (2002) also finds a break in the declining trend of mortality among males aged 15 to 19 years old after 1960, which in the 1980s and 1990s extends to adult ages. According to Morais (2002), those increases in mortality are associated to traffic accidents and other violent causes that are largely determined by social pathologies, such as alcoholism, lifestyle and the cumulative incidence of such behaviour.

The explanations provided by Lages (2007) for the phenomenon are based on the effects of the major transformations that occurred in Portugal since 1960. Among them are: the effects of emigration to Europe and the years of war overseas on the structure and changing behaviors of the population during the 1970s; the impact of the political revolution and of the return of around half a million Portuguese citizens from former African colonies, particularly visible during the 1980s, the consolidation of the democracy, the entry of Portugal in the European Union, and the establishment of a prominent middle class in the following years. These were fifty years of major structural changes, imposing huge pressures on the Portuguese society and economic structure, accompanied by cultural transformations of similar weight, which are reflected in the differentiated risk exposure of the population.

Cohort effects in male Portuguese mortality are likely to have been caused by a combination of all those factors, which might have produced a different imprint in generations born a few years apart. The eventual determination of the exact origin of such effects would require further multidisciplinary research, for instance related to epidemiological developments and to social and economic conditions that could have an enduring impact on rates of mortality improvement.

The results of this study show a cohort effect for men born between 1948 and 1959. These generations have now attained, or are close to attaining ages 53 – 64. If these men keep having higher mortality rates over their whole life-time experience, the cohort effect will shift mortality rates for the elderly men upwards, changing the lifespan expectations.

The importance of considering cohort effects in forecasting future mortality patterns depends however on the enduring or temporary nature of those effects, which

still remains to establish. If we expect that the poor mortality conditions are to be kept as the cohort ages, one should consider such cohort effects in forecasting future mortality patterns. Otherwise, if cohort effects are expected to fade with age, their impacts in forecasts are negligible and should be excluded. Clues on future mortality expectations for such cohort may be provided by the study of their causes of death, as well as by the explanatory factors that contributed for the presence of a cohort effect in male mortality.

We leave for future research the detailed analysis of the factors that may explain the presence of cohort effects in Portugal, which also would be helpful in accessing the nature of such cohort effects and the importance of incorporating cohort effects into mortality forecasting models. Also, further research on how to incorporate cohort effects into forecasting models and what are the impacts of their consideration in future mortality patterns, and consequently on future life expectancy.

References

Alho, J.M., 2000. Discussion of Lee (2000). *North American Actuarial Journal* 4, 91–93.

Andreev, K. and Vaupel, J. (2005). Patterns of mortality improvements over age and time in developed countries: estimation, presentation and implications for mortality forecasting. (http://paa2005.princeton.edu/download.aspx?submissionId=51061)

Brillinger, D. (1986). The natural variability of vital rates and associated statistics. *Biometrics* 42, 693 – 734.

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

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and 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.

Cameron, A. C., and Trivedi, P. K. (1998). Regression Analysis of Count Data. Cambridge University Press.

Case, R.A.M. (1956). Cohort analysis of mortality rates as an historical or narrative technique. *British Journal of Preventive and Social Medicine* 10, 159 – 171.

Carstensen, B. (2007). Age-period-cohort models for the Lexis diagram. *Statistics in Medicine 2007*; 26: 3018-3045.

Clayton, D. and Schifflers, E. (1987). Models for temporal variation in cancer rates II: Age-Period-Cohort Models. *Statistics in Medicine*, Vol. 6, 469-481 (1987).

Cleries, R., Martínez, J. M., Valls, J., Pareja, L., Esteban, L., Gispert, R., Moreno, V., Ribes, J., and Borràs, J. M. (2009). Life expectancy and age-period-cohort effects: analysis and projections of mortality in Spain between 1977 and 2016. *Public Health*, 123 (2): 156-62.

Coelho, E., and Nunes, L.C. (2010). Patterns of mortality decline in Portugal since 1950. Manuscript.

Coelho, E. and Nunes, L. C. (2011). Forecasting mortality in the event of a structural change. *Journal of the Royal Statistical Society A* 174, Part 3, pp 713-736.

Coughland, G., Epstein, D., Ong, A., Sinha, A., Hevia-Portocarrero, J., Gingrich, E., Khalaf-Allah, M., and Joseph, P. (2007), LifeMetrics: A toolkit for measuring and managing longevity and mortality risks. Technical document. Available at www.lifemetrics.com.)

Collins, J. J. (1982). The contribution of medical measures to the decline of mortality from respiratory tuberculosis: an age-period-cohort model. *Demography*, vol. 19, n. º 3, 409-427.

Fienberg, S. E. and Mason, W. M. (1982). Specification and implementation of age, period and cohort Models. Technical Report n.º 235, Carnegie-Mellon University.

Fienberg, S. E. and Mason, W. M. (1979). Identification and estimation of age-period-cohort models in the analysis of discrete archival data. *Sociological Methodology*, Vol. 10, pp. 1-67.

Fu, W. J. (2000). Ridge estimator in singular design with application to age-period-cohort analysis of disease rates. *Communications in Statistics – Theory and Method* 29: 263-278.

Glenn, N. D. 1976. "Cohort analysts' futile quest: statistical attempts to separate age, period, and cohort effects." *American Sociological Review*, 41:900–905.

Greenberg, B. G., Wright, J. J. , and Sheps, C. G. (1950). A technique for analysing some factors affecting the incidence of syphilis. *Journal of the American Statistical Association*, 45(251):373-99.

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

Hobcraft, J., Menken, J. and Preston, S. (1982). Age, period, and cohort effects in demography: a review. *Population Index*, Vol. 48, n.º 1, pp. 4-43.

Holford, T. R. (1983). The estimation of age, period and cohort effects for vital rates. *Biometrics* 39, 311-324.

Holford, T. R. (1991). Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annual Rev. Publ. Health*, 12:425-57.

Holford, T.R. (2005). Age-Period-Cohort Analysis. Encycl. *Biostatistics* 2. Ed.: 105-123.

Holford, T. R. (2006). Approaches to fitting age-period-cohort models with unequal intervals. *Statistics in medicine,* 2006; 25: 977-933.

Human Mortality Database (2007). Methods protocol for the Human Mortality Database. Human Mortality Database, University of California, Berkeley. (Available from www.mortality.org/Public/Docs/MethodsProtocol.pdf.)

James, I. R. and Segal, M. R. (1982). On a Method of Mortality Analysis Incorporating Age-Year Interaction, with Application to Prostate Cancer Mortality. *Biometrics*, Vol. 38, No. 2 (Jun., 1982), pp. 433-443.

Kuang, D., Nielsen, B. and Nielsen, J.P. (2008a). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika*, 95, 4, pp.979 – 986.

Kuang, D., Nielsen, B. and Nielsen, J.P. (2008b). Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika*, 95, 4, pp.997 – 991.

Kupper, L. L., and Janis, J. M. (1980). The multiple classification model in age-period-cohort analysis: Theoretical considerations. Institute of Statistics, University of North Carolina, Mimeo Series n.º 1311, 1980.

Kupper, L. L., Janis, J. M., Karmous, A., and Greenberg, B. G. (1985). Statistical age-period-cohort analysis: A review and critique. *Journal of Chronic Disease* 38, n.º 10, pp. 811-830.

Kupper, L.L., Janis, J. M., Salama, I. A., Yoshizawa, C. N., and Greenberg, B. G. (1983). Age-Period-Cohort Analyses: An Illustration of the Problems in Assessing Interaction in One Observation per Cell Data. Communications in Statistics - Theory and Methods, Volume 12, Issue 23.

Lages, M. F. (2007). "Os comportamentos de risco dos jovens portugueses e a sua mortalidade*", Análise Social*, vol. XLII (183), pp. 395-418.

Lee, R. D. and Carter, L. R. (1992). Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, Vol. 87, No. 419 (Sep., 1992), pp. 659-671.

Mason, K. O., Mason, W. M., Winsborough, H. H. and Poole, W. K. (1973). Some Methodological Issues in Cohort Analysis of Archival Data. *American Sociological Review*, Vol. 38, n. º2, pp. 242-258.

Moolgavkar, S. H., Stevens, R. G. and Lee, J. A. H. (1979). Effect of age on incidence of breast cancer in females. *Journal of the National Cancer Institute* 62, 493-501.

Morais, M. G. (2002). "Causas de morte no século XX: Transição e estruturas de mortalidade em Portugal Continental, Lisboa, Edições Colibri.

Murphy, M. (2010). Reexamining the dominance of birth cohorts effects on mortality. *Population and Development Review* 36 (2): 365 -390.

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

Richards, S. J. (2008). Detecting year-of-birth mortality patterns with limited data, *Journal of the Royal Statistical Society A* 171, Part 1, pp 279-298.

Richards, S. J., Ellam, J. R., Hubbard, J., Lu, J. L. C., Makin, S. J. and Miller, K. A. (2007). Two-dimensional mortality data: patterns and projections. British Actuarial Journal, Volume 13, Part III, No. 59, pages 479-555.

Robertson, C. and Boyle, P. (1998a). Age-Period-Cohort Analysis of Chronic Disease Rates I: Modelling Approach. *Statistics in Medicine*, 17, 1305-1323 (1998).

Robertson, C. and Boyle, P. (1998b). Age-Period-Cohort Analysis of Chronic Disease Rates II: Graphical approaches. *Statistics in Medicine*, 17, 1325-1340 (1998).

Robertson, C., Gandini, S., and Boyle, P. (1999). Age-period-cohort models: A comparative study of available methodologies. *Journal Clinical Epidemiology*, Vol. 52, N° 5, pp. 569-583.

Rodgers, W.L. (1982). Estimable functions of age, period, and cohort effects. *American Sociological Review*, Vol. 47, n.°6, pp. 774 – 787.

Ryder, N. B. (1965). The cohort as a concept in the study of social change. *American Sociological Review*, Vol. 30, N° 6, 843-861.

Schulhofer-Wohl, S. And Yang, Y., 2011. Modelling the evolution of age and cohort effects in social research," Staff Report 461, Federal Reserve Bank of Minneapolis.

Willekens, F., and Scherbov, S. (1991). Age-period-cohort analysis of mortality with applications to soviet data. WP-91-42, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Willets, R.C. (1999). Mortality in the next millennium. Staple Inn Actuarial Society Metting 7 December 1999. Available at: http://www.sias.org.uk/.

Willets, R.C. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, 10, 833-877.

Wilmoth, J. R. (2001). Les modèles âge-période-cohorte en démographie. In: Graziella Caselli, Jacques Vallin, and Guillaume Wunsch (eds.), Démographie: Analyse et Synthèse. I. La Dynamique des Populations, Paris: INED, pp. 379-397.

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

Wilmoth, J. R. (1990). Variation in vital rates by age, period and cohort. *Sociological Methodology*, Vol. 20, pp. 295-335.

Yang, Y., Fu, W. J. and Land, K. C. (2004). A Methodological Comparison of Age-Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models. *Sociological Methodology*, 34: 75-110.