On the use of multi-population mortality models for deprivation subgroups in a population

Kam Kuen Kenny Mok

A thesis submitted for the degree of Master of Research

Department of Applied Finance and Actuarial Studies

Faculty of Business and Economics

Macquarie University

October 2017

Declaration

I certify that the work contained in this thesis has not been submitted for a higher degree to any other university or institution. All sources of information and literature utilised within have been appropriately disclosed and acknowledged.

Kam Kuen Kenny Mok

Summary

Actuaries and demographers have shown a keen interest in the longevity improvement in the past decades since human longevity significantly affects society, especially influencing the creation of government policy for pensioners and life insurance products. Studies have shown that mortality rates have converged across populations, in which academics have been developing multi-population models to capture this phenomenon.

Multi-population models can be used to assess longevity basis risk, which is the mismatch of the longevity outcome between two different populations. Longevity basis risk arises from hedging the sub-population using mortality hedging instruments based on the reference mortality rate. Good multi-population models are able to capture and reduce the basis risk along with hedging instrument.

In this thesis, we study a variety of multi-population models from the literature. These models are fitted for the UK population data and the deprivation subgroups in England. Goodness-of-fit tests, and the examination of forecasting accuracy and hedging effectiveness are used to compare the multi-population models.

Acknowledgements

I would like to thank my supervisor, Associate Professor Jackie Li, for his guidance and invaluable support. He has provided many insights and feedback for my thesis. I also want to thank Chong It Tan, Kenneth Wong and Colin Zhang for improvement of the writings and ideas for the thesis.

I also want to thank Lin Bai, Shauna Ferris and Angela Chow for the faculty- and department-level administrative support. Finally, I want to thank my family and friends for supporting me throughout the year.

Table of Contents

Chapter 1

Introduction

Rapid increase in life expectancy over the last few decades has drawn much attention from actuaries and demographers. The study of human longevity plays an integral part in social science research, as it provides information that is useful to various decision makers such as the government and insurance companies. Some examples include calculating an optimal retirement age for individuals, designing a pension policy for retirees, and pricing insurance and annuity products for prospective customers. Improvements in mortality model projections can potentially reduce the risk of miscalculation faced by insurance companies and assist the government in future budget planning. In particular, it would be informative to investigate the differences in mortality experience between different groups in a population, which usually have a variety of socioeconomic status. Accordingly, the aim of this thesis is to examine the performances of several multi-population mortality projection models, and a few possible modifications of them, on the deprivation subgroups in the UK population.

Mortality models have been proposed since a few centuries ago. The classical mortality models are often deterministic in nature, such as the Gompertz Law proposed by Benjamin Gompertz in 1825 and the Makeham Law developed by William Makeham in 1860. One major limitation in using such deterministic models is that these models only take a snapshot of the mortality scenario at a particular point of time. Stochastic models, on the other hand, cater for temporal developments in mortality and can be used to generate a forecast probability distribution of future mortality rates.

Lee and Carter (1992) have pioneered in modelling mortality data and forecasting future mortality rates. Their original model has the benefit of being a stochastic model and makes a projection using only the usual mortality data but without the need of experts' opinion on future mortality improvement. The model structure itself is straightforward and incorporates the age and time effects explicitly. However, there are several limitations underlying this model. For example, it fails to capture the cohort effect, which explains the common mortality experience for all those lives born in the same

year. Since then, many extensions and modifications have been proposed to improve the Lee-Carter model's performance. For instance, the age-period-cohort model developed by Renshaw and Haberman (2006) and Currie (2006) caters for the cohort effect. A detailed review of single-population mortality projection models can be found in Cairns et al. (2009) and Li et al. (2012).

The main disadvantage of single-population mortality projection models is that they only tackle one particular population at a time and cannot handle multiple populations simultaneously. As shown by Wilson (2001, 2011), mortality rates tend to converge on a global scale due to many countries having quite similar socioeconomic conditions and development trends. Mortality projections using a single-population model independently for each population without allowing for the potential co-movements between populations could result in unreasonable divergence in the projected mortality rates between those populations. In contrast, a number of multi-population mortality projection models can ensure the long-term convergence between populations or countries and can possibility provide better mortality projections than otherwise.

Mortality convergence can also be found amongst sub-populations, e.g. Villegas and Haberman (2014) showed that the mortality rates of the deprivation subgroups in the UK tend to converge over time. The deprivation level is a measure of the socioeconomic status within the country. Accordingly, it is crucial to consider multi-population mortality modelling, in order to deal with the underlying relationships between different groups of lives. This thesis will extend the study carried out by Villegas and Haberman by examining a more extensive list of multi-population mortality projection models for the deprivation subgroups in the UK. The data used in the thesis cover a very recent period; more details will be provided in Chapter 3.

Multi-population models can be used to assess longevity basis risk – the mismatch of the longevity outcomes between two different populations. Standard mortality hedging instruments have been developed in practice in recent years to hedge longevity risk, but the mortality rates built into the instruments are based on the reference population. Longevity basis risk arises from hedging the sub-population using mortality hedging instruments. Hedging effectiveness under each model will also be examined in the study. Very recently, Villegas et al (2017) performed a comparative study of two-population models for assessing longevity basis risk. The thesis will perform a more extensive study on hedging effectiveness, using a wider range of models, ages, and cohorts.

The remainder of the thesis is structured as follows. Chapter 2 provides a literature review on various mortality projection models. Chapter 3 describes the deprivation subgroups dataset and the methodology for fitting the selected models to the data. Chapter 4 compares the fitting results between different models. Chapter 5 projects the future mortality rates based on each model, compares the forecasting performances via an out-of-sample analysis, and examines the hedging effectiveness under each model assumption. Chapter 6 discusses potential future research and concludes the thesis.

Chapter 2

Literature review

2.1 Single population mortality models

Booth and Tickle (2008) classified mortality forecasting methods under three categories: expectation, explanation and extrapolation. Expectation methods require experts' opinion on future mortality improvement. Explanation methods use structural or epidemiological methods to link the cause of death with the mortality rate. Extrapolative methods assume that future mortality trends continue from the past.

There is a main disadvantage of using expectation methods - it is subject to personal bias. Studies have also shown that expectation methods tend to underestimate the actual mortality improvement (Lee and Carter, 1992; Lee and Miller,2001). Explanation models often use a regression structure with explanatory variables and lagged risk factors, which result in difficulties and lower accuracy in long term projection (Booth, 2006; Booth and Tickle, 2008). Extrapolative methods only require assuming the continuation of the historical trend; this is the key advantage but can also be considered as a limitation. Out of the three methods, the extrapolative methods are widely used by the academics and practitioners (Booth, 2006).

The remainder of this section will focus on extrapolative models. A natural starting point is the Lee-Carter model, which pioneered the use of extrapolative models in mortality projection analysis. The Lee-Carter model is expressed as:

$$\ln m_{x,t} = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}$$

where $m_{x,t}$ refers to the central death rate at age x in year t, α_x represents the base mortality rate at age x, κ_t represents the overall mortality changes over time t, and β_x is an age-specific sensitivity measure to κ_t. The term $\varepsilon_{x,t}$ refers to a homoscedastic error term. Singular value decomposition (SVD) is used to estimate the parameters of the Lee-Carter model. The time varying factor usually exhibits a linear trend; as such, a random walk with drift is used to project the future value of κ_t.

The main advantage of the Lee-Carter model is simplicity since it does not require involve a complex structure and parameter setting. The mortality projection also does not require any expert or subjective judgement as the projection is driven purely by the historical trend. These features made the Lee-Carter model very popular for use in mortality projection. However, the Lee-Carter model has some limitations that academics have been trying to improve. Booth et al. (2002) included additional time varying factors with respective additional sensitivity measures to improve the fit. Different parameter estimations other than the SVD method have been used for the Lee-Carter model, such as Wilmoth (1993) who used Weighted Least Squares. Brouhns et al (2002) assumed that the number of deaths follows a Poisson distribution and applied a Poisson log-bilinear regression model proposed by Brillinger (1986) to estimate the parameters in the Lee-Carter model; these assumptions allow the use of standard statistical techniques such as maximum likelihood estimation.

Another extension of the Lee-Carter model is to incorporate the cohort effect, where the cohort refers to the groups of people who were born in the same year. This class of models is referred to as age-period-cohort models. Willets (2004) showed that cohorts in the UK undergo similar mortality experience. Renshaw and Haberman (2006) included the cohort effect to extend the Lee-Carter model, which can be expressed as:

$$\ln m_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t + \beta_x^{(2)} \iota_{t-x} + \varepsilon_{x,t}$$

where $\beta_x^{(i)}$ refers to the sensitivity measure, ι_{t-x} refers to the cohort year, and $t - x$ refers to the year of birth. Currie (2006) further simplified the model by setting the sensitivity measure $\beta_x^{(1)} = \beta_x^{(2)} = 1$.

Lee-Carter models have a poor performance for fitting old age groups pointed out by Cairns et al. (2009). Cairns et al. (2006) proposed the Cairns-Blake-Dowd (CBD) model specifically for older ages and assumed the logit death rates of these ages vary linearly across age. The model includes two time varying factors and it is expressed as:

$$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \varepsilon_{x,t}$$

where $q_{x,t}$ represents the probability that an individual aged x years old at time t will die within a year, $\bar{x}$ represents the average age of the data, $\kappa_t^{(1)}$ is the first time varying factor representing the base mortality level across the population, and $\kappa_t^{(2)}$ is the second time varying factor representing the slope to separate the mortality for each age. Usually $\kappa_t^{(1)}$ has a downward trend across time as mortality improves. A downward trend for $\kappa_t^{(2)}$ represents younger age groups experiencing better mortality improvement compared to the older age groups. This model is also known as the M5 model in Cairns et al. (2009), where the model is further expanded to the M6 model by including an additional cohort effect, which can be expressed as:

$$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \iota_{t-x} + \varepsilon_{x,t}$$

where ι_{t-x} represents the cohort effect. Another variation of the M5 model, namely M7, includes a quadratic term to improve the fitting, which can be expressed as:

$$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \iota_{t-x} + \varepsilon_{x,t}$$

where $\kappa_t^{(3)}$ is the third time varying factor and $\hat{\sigma}_x^2$ is the simple average of $(x - \bar{x})^2$.

2.2 Convergence of the mortality rates

White (2002) found that mortality rates converge between high income countries and stated the dangers of mortality projection from using only one population. Wilson (2001) studied 184 entities from the United Nations and showed that mortality rates converge on a global scale with many countries exhibiting similar socioeconomic conditions. The global convergence also reduces the mortality difference between developing and developed countries (Wilson, 2011).

Moreover, mortality convergence can occur within a single country, such as between both sexes. Historically, males and females are known to have different mortality experiences - females have always been expected to live longer than males since the life table method was established (Luy, 2003). The gap in mortality rates between the two sexes, however, has fallen since the 1980s (Li, 2013). Li and Lee (2005) argued that the original Lee-Carter model would perform reasonably well for a single sex or aggregated sexes. However, when projecting mortality independently between the two sexes, the resulting projected mortality rates could diverge gradually over time, which would be inconsistent with the general views in the literature and the patterns shown in historical data. There is a need for improvement on the single population models. Thus, in recent years, academics have developed several multi-population models to capture the mortality convergence to provide better fitting and mortality projection.

2.3 Multi-population mortality models

Many multi-population models are built based on extensions of the single population models. In the following, the models are grouped by whether they share the same key features. 11 multi-population mortality models have been tested for the studies. The studies have applied the Poisson death count assumptions for the 11 models which are similar to the assumptions that Brouhns et al (2002) have used for the Lee-Carter model. Further details for the Poisson framework are shown in section 3.3. The additional subscripts i in this section refers to which population are fitted, where $i = 1$ refers to the reference population and $i = 2$ refers to the subpopulation.

Associated mortality indices

Under this category, a single population model is fitted to each population separately. The estimated temporal parameters of both populations are then modelled jointly with a suitable time series process to ensure the convergences in the projected mortality rates. The mortality structure of the first three models are basically the same, in terms of using the Lee and Carter (1992) model structure. The main differences between applying the three models lie in the time series processes being adopted for projection.

Model (1):

$$\ln m_{x,t,i} = \alpha_{x,i} + \beta_{x,i}\kappa_{t,i}; \qquad \text{(Lee-Carter model)}$$

$$\mathrm{K}_t = \Theta + \mathrm{K}_{t-1} + \Delta_t, \qquad \text{(bivariate random walk with drift)}$$

In model (1) (Lee-Carter model), $\alpha_{x,i}$ represents the base mortality rate at age x, $\kappa_{t,i}$ represents the overall mortality changes over time t, and $\beta_{x,i}$ is an age-specific sensitivity measure to $\kappa_{t,i}$. A bivariate random walk with drift is fitted to the time varying factors where $\mathrm{K}_t = \left(\kappa_{t,1}, \kappa_{t,2}\right)'$, Θ is the vector term for the drift, and $\Delta_t = \left(\varepsilon_{t,1}, \varepsilon_{t,2}\right)'$ is the vector error term, in which the error terms $\varepsilon_{t,1}$ and $\varepsilon_{t,2}$ follow a bivariate normal distribution. The identifiability constraints for the Lee-Carter models are taken as $\sum_x \beta_{x,i} = 1$ and $\sum_t \kappa_{t,i} = 0$. One shortfall of applying this time series process is the possibility that the projected mortality rates of the two populations may diverge indefinitely over time.

Model (2):

$$\ln m_{x,t,i} = \alpha_{x,i} + \beta_{x,i}\kappa_{t,i}; \qquad \text{(Lee-Carter model)}$$

$$\kappa_{t,1} = \theta + \kappa_{t-1,1} + \delta_t; \qquad \text{(random walk with drift)}$$

$$\kappa_{t,2} = a_0 + a_1\kappa_{t,1} + \omega_t, \qquad \text{(co-integrated process)}$$

Li and Hardy (2011) improved the time series modelling by using a co-integrated process for the time varying factors (model (2)) and intended to allow for some more connection between the two time series. The extent of possible divergence between the simulated mortality rates of the two populations may then be reduced. A co-integrated process can be applied if there exists a linear combination of $\kappa_{t,1}$ and $\kappa_{t,2}$ which is a stationary process. The terms θ, a_0, and a_1 are the coefficients of the co-integrated process. The error terms δ_t and ω_t are independent and normally distributed.

Model (3):

$$\ln m_{x,t,i} = \alpha_{x,i} + \beta_{x,i}\kappa_{t,i}; \qquad \text{(Lee-Carter model)}$$

$$\mathrm{K}_t - \mathrm{K}_{t-1} = \Theta + \Pi\mathrm{K}_{t-1} + \Gamma\,(\mathrm{K}_{t-1} - \mathrm{K}_{t-2}) + \Delta_t, \qquad \text{(vector error correction model)}$$

The vector error correction model adopted in model (3) for the time varying factors, proposed by Yang and Wang (2013), has a similar advantage to model (2) in reducing the extent of possible divergence in the simulated mortality rates between the two populations. Under certain technical condition, it can also ensure the long-run equilibrium in the projected mortality rates between the two populations whilst allowing certain deviations in the short term. Note that $\mathrm{K}_t = \left(\kappa_{t,1}, \kappa_{t,2}\right)'$, Θ is the vector term, Π and Γ are 2 by 2 coefficient matrices, and $\Delta_t = \left(\varepsilon_{t,1}, \varepsilon_{t,2}\right)'$ is the vector error term, where the error terms $\varepsilon_{t,1}$ and $\varepsilon_{t,2}$ are bivariate normal random variables.

Model (4):

$$\ln m_{x,t,i} = \alpha_{x,i} + \beta_x\kappa_{t,i}; \qquad \text{(joint Lee-Carter model)}$$

$$\kappa_{t,1} = \theta + \kappa_{t-1,1} + \delta_t; \qquad \text{(random walk with drift)}$$

$$\kappa_{t,1} - \kappa_{t,2} = b_0 + b_1\left(\kappa_{t-1,1} - \kappa_{t-2,2}\right) + \omega_t, \qquad \text{(AR (1) process)}$$

Regarding model (4), Zhou, Li, and Tan (2013) used a common sensitivity measure for the age-specific factor in the Lee-Carter model structure for both populations such that $\beta_x = \beta_{x,1} = \beta_{x,2}$. In addition, the difference between the two time varying factors $\kappa_{t,1}$ and $\kappa_{t,2}$ is modelled with an weakly stationary AR(1) process to ensure that the mortality rates of the two populations are projected along the same direction over the long term. The terms θ, b_0, and b_1 are the coefficients of the time series processes. The error terms δ_t and ω_t follow a bivariate normal distribution. The constraints for this model are set as $\sum_x \beta_x = 1$ and $\sum_t \kappa_{t,i} = 0$.

Model (5):

$$\ln m_{x,t,i} = \alpha_{x,i} + n_\alpha^{-1}\kappa_{t,1} + n_\alpha^{-1}\iota_{t-x,i}; \qquad \text{(age-period-cohort model)}$$

$$\kappa_{t,1} = \theta + \kappa_{t-1,1} + \delta_t; \qquad \text{(random walk with drift)}$$

$$\kappa_{t,1} - \kappa_{t,2} = b_0 + b_1\left(\kappa_{t-1,1} - \kappa_{t-1,2}\right) + \omega_t; \qquad \text{(AR (1) process)}$$

$$\tilde{\iota}_{h,1} = c_{0,1} + c_{1,1}\iota_{h-1,1} + c_{2,1}\tilde{\iota}_{h-2,1} + \varepsilon_{h,1}; \qquad \text{(AR (2) process)}$$

$$\iota_{h,1} - \iota_{h,2} = c_{0,2} + c_{1,2}\left(\iota_{h-1,1} - \iota_{h-1,2}\right) + c_{2,2}\left(\iota_{h-2,1} - \iota_{h-2,2}\right) + \varepsilon_{h,2}, \qquad \text{(AR (2) process)}$$

In models (5) and (6), each population is fitted individually with the age-period-cohort model discussed by Renshaw and Haberman (2006) and Currie (2006). Cairns, Blake, Dowd, Coughlan, and Khalaf-Allah (2011) (model (5)) modelled the two sets of time series jointly. The parameters $\alpha_{x,i}$ represents the age effect, $\kappa_{t,1}$ represents the period effect, and $\iota_{t-x,i}$ represents the cohort effect, where $h = t - x$ is the year of birth. The term n_α^{-1} is the reciprocal of the number of ages. The terms $\tilde{\iota}_{h,1} = \iota_{h,1} - \rho_0 - \rho_1(h - \bar{h})$, where $\bar{h}$ represents the average cohort year plus one, and ρ_0 and ρ_1 represent the linear trend for $\iota_{h,1}$. The term$\tilde{\iota}_{h,1}$ is assumed to follow a stationary process after removing the linear trend. The terms θ, b_0, b_1, $c_{0,1}$, $c_{1,1}$, $c_{2,1}$, $c_{0,2}$, $c_{1,2}$, and $c_{2,2}$ are the coefficients of the time series processes. The error terms δ_t and ω_t are bivariate normal random variables, while the error terms $\varepsilon_{h,1}$ and $\varepsilon_{h,2}$ also follow a bivariate normal distribution but are independent of δ_t and ω_t. The difference between the period effects of the two populations is assumed to follow a stationary process to ensure that the mortality rates of the two populations are projected in the same direction over the long run. The same argument applies for the difference between the cohort effects of the two populations. The constraints for the age-period-cohort model used in the study are set as $\sum_t \kappa_{t,i} = 0$ and $\sum_h \iota_{h,i} = 0$.

Model (6):

$$\ln m_{x,t,i} = \alpha_{x,i} + n_{\alpha}^{-1}\kappa_{t,1} + n_{\alpha}^{-1}\iota_{t-x,i}; \qquad \text{(age-period-cohort model)}$$

$$\kappa_{t,1} = \theta + \kappa_{t-1,1} + \delta_t; \qquad \text{(random walk with drift)}$$

$$\kappa_{t,2} - \kappa_{t-1,2} = b_0 + b_1\left(\kappa_{t-1,1} - \kappa_{t-1,2}\right) + \omega_t; \qquad \text{(gravity model)}$$

$$\iota_{h,1} - \iota_{h-1,1} = c_{0,1} + c_{1,1}\left(\iota_{h-1,1} - \iota_{h-2,1}\right) + \varepsilon_{h,1}; \qquad \text{(AR (1) process)}$$

$$\iota_{h,2} - \iota_{h-1,2} = c_{0,2} + c_{1,2}\left(\iota_{h-1,2} - \iota_{h-2,2}\right) + c_{2,2}\left(\iota_{h-1,1} - \iota_{h-1,2}\right) + \varepsilon_{h,2}, \qquad \text{(gravity model)}$$

On the other hand, Dowd, Cairns, Blake, Coughlan, and Khalaf (2011) (model (6)) adopted the gravity model such that the subpopulation's (population 2) death rates move back closer to those of the reference population, the pace of which depends on the two gravity parameters. The error term structure of model (6) is the same as that of model (5).

Model (7):

$$\text{logit}\left(q_{x,t,i}\right) = \kappa_{t,i}^{(1)} + \kappa_{t,i}^{(2)}(x - \bar{x}) + \kappa_{t,i}^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2); \qquad \text{(3-factor CBD model)}$$

$$\mathrm{Z}_t = \Theta + \sum_{i=1}^{r} \Phi_l \mathrm{Z}_{t-l} + \sum_{u=1}^{s} \Lambda_u \Delta_{t-u} + \Delta_t, \qquad \text{(VARMA } (p, q) \text{ model)}$$

Tan, Li, Li, and Balasooriya (2014) extended the Cairns, Blake, and Dowd (2006) model (model (7)) for individual populations and used a vector autoregressive moving average process (VARMA) for the time series of the computed parameters. The terms $\kappa_{t,i}^{(1)}$, $\kappa_{t,i}^{(2)}$, and $\kappa_{t,i}^{(3)}$ are the time varying factors of the CBD model, $\bar{x}$ represents the average age of the data, and $\hat{\sigma}_x^2$ is the simple average of $(x - \bar{x})^2$. The vector $\mathrm{Z}_t = (\Delta^{(d)}\kappa_{t,1}^{(1)}, \Delta^{(d)}\kappa_{t,1}^{(2)}, \Delta^{(\mathrm{d})}\kappa_{t,1}^{(3)}, \Delta^{(d)}\kappa_{t,2}^{(1)}, \Delta^{(d)}\kappa_{t,2}^{(2)}, \Delta^{(d)}\kappa_{t,2}^{(3)})$, where $\Delta^{(d)}\kappa_{t,i}^{(j)}$ represents the dth difference in $\kappa_{t,i}^{(j)}$. The terms Φ_l and Λ_u are the VARMA coefficient matrices. The vector error term Δ_t follows a multivariate normal distribution.

Moreover, Li et al. (2015) performed an extensive study of two-population models by extending the models[1] in Cairns et al. (2009). These extensions are constructed by considering a common time

[1] Cairns et al. (2009) covered a variety of models such as the Lee-Carter model, age-period-cohort model, and CBD models.

varying factor between populations and utilizing multivariate time series to model the time varying factors jointly.

Ratio of death rates

The ratio of death rates models are another set of models which can ensure the convergence of the two populations' projected ratio of mortality rates at each age. This class of models adopts the Jarner and Kryger (2011) approach of multi-population mortality modelling, in which the single population model is fitted to the reference population, and then the ratio of the two populations' death rates is modelled with another structure. They applied a frailty model to the reference population and modelled the log ratio of death rates with a function of period and age factors. One difficulty of applying this model is that the choice of the frailty is subjective and is based on the population data.

Model (8):

$$\text{logit}(q_{x,t,i}) = \kappa_{t,i}^{(1)} + \kappa_{t,i}^{(2)}(x - \bar{x}); \qquad \text{(CBD model – M5)}$$

$$\frac{q_{x,i,2}}{q_{x,i,1}} = 1 + \psi_x(1)\gamma_t(1) + \psi_x(2)\gamma_t(2) + \cdots + \psi_x(n)\gamma_t(n); \qquad \text{(age and period factors)}$$

$$\begin{pmatrix} \kappa_t^{(1)} \\ \kappa_t^{(2)} \end{pmatrix} = \begin{pmatrix} \theta^{(1)} \\ \theta^{(2)} \end{pmatrix} + \begin{pmatrix} \kappa_{t-1}^{(1)} \\ \kappa_{t-1}^{(2)} \end{pmatrix} + \begin{pmatrix} \delta_t^{(1)} \\ \delta_t^{(2)} \end{pmatrix}; \qquad \text{(bivariate random walk with drift)}$$

$$\gamma_t(j) = c_0(j) + c_1(j)\gamma_{t-1}(j) + \varepsilon_t(j), \qquad \text{(AR (1) process)}$$

Plat (2009) used the CBD (2006) M5 model for the reference population and modelled the ratio of death rates between the two populations as a function of age and period (model (8)). The ratio between the two populations will converge if the corresponding time series follow stationary AR(1) processes. The terms $\kappa_{t,i}^{(1)}$ and $\kappa_{t,i}^{(2)}$ are the time varying factors of the CBD model, and $\bar{x}$ represents the average age of the data. The terms $\psi_x(j)$ refers to the age factor and $\gamma_t(j)$ is the period factor. Moreover, $\theta^{(1)}$ and $\theta^{(2)}$ are the drift terms, $c_0(j)$ and $c_1(j)$ are the coefficients of the AR (1) process, the error terms $\delta_t^{(1)}$ and $\delta_t^{(2)}$ follow a bivariate normal distribution, and the error terms $\varepsilon_t(j)$'s are independent and normally distributed. The number of age and period factors are determined by comparing the BIC values as described in Section 4.

Model (9):

$$\ln m_{x,t,1} = \alpha_{x,} + \beta_{x,}\kappa_t + \iota_{t-x};$$ (Lee-Carter model with cohort factor)

$$\ln \frac{m_{x,t,2}}{m_{x,t,1}} = a_x + b_x k_t\,;$$ (age and period factors)

$$\kappa_t = \theta + \kappa_{t-1} + \delta_t;$$ (random walk with drift)

$$k_t = c + k_{t-1} + \varepsilon_t;$$ (random walk with drift)

$$\iota_h - \iota_{h-1} = c_{0,1} + c_{1,1}(\iota_{h-1} - \iota_{h-2}) + \omega_h,$$ (AR (1) process)

Villegas and Haberman (2014) applied the Lee-Carter model with a cohort factor proposed by Renshaw and Haberman (2006) to the reference population, and then modelled the ratio of death rates with the Lee-Carter bilinear age and period structure (model (9)). The main assumption of this model is that subpopulations exhibit similar cohort experience as the reference population, so this common cohort effect feature limits the possibility of multi-country modelling. The parameters $\alpha_{x,i}$ represents the base mortality rate at age x, $\kappa_{t,i}$ describes the overall mortality changes over time t, $\beta_{x,i}$ is an age-specific sensitivity measure of $\kappa_{t,i}$, and ι_{t-x} depicts the cohort effect. The terms a_x and b_x are the age factors and k_t is the time factor. The terms θ, c, $c_{0,1}$, and $c_{1,1}$ are the coefficients of the time series processes, the error terms ε_t and δ_t are independent and normally distributed, and the error term ω_h is also independent of δ_t and ε_t and normally distributed. The constraints for this model are taken as $\sum_x \beta_x = 1$, $\sum_t \kappa_t = 0$, $\sum_h \iota_h = 0$, $\sum_x b_x = 1$ and $\sum_t k_t = 0$.

<u>Modelling the difference of the logit death rates between two populations</u>

Haberman et al. (2014) considered modelling the difference of the logit death rates between a book and a reference population in an investigation about longevity basis risk. It follows the Jarner and Kryger (2011) multi-population modelling approach. They compared a variety of models under the relative modelling framework, where the results indicated that the so-called M7-M5 model and the CAE + Cohorts model perform the best in terms of fitting and calculating reasonable hedging effectiveness. This class of models provides more flexibility for dealing with different data lengths between both populations. The time series processes for the difference between the two populations are selected to ensure that the long-term difference in the logit death rates converges if the stationary condition is satisfied.

Model (10):

$$\text{logit}(q_{x,t,1}) = \kappa_{t,1}^{(1)} + \kappa_{t,1}^{(2)}(x-\bar{x}) + \kappa_{t,1}^{(3)}((x-\bar{x})^2 - \hat{\sigma}_x^2) + \iota_{t-x};$$ (CBD model – M7)

$$\text{logit}(q_{x,t,1}) - \text{logit}(q_{x,t,2}) = \kappa_{t,2}^{(1)} + \kappa_{t,2}^{(2)}(x-\bar{x});$$ (CBD model – M5)

$$\begin{pmatrix}\kappa_{t,1}^{(1)}\\ \kappa_{t,1}^{(2)}\\ \kappa_{t,1}^{(3)}\end{pmatrix} = \begin{pmatrix}\theta_1\\ \theta_2\\ \theta_3\end{pmatrix} + \begin{pmatrix}\kappa_{t-1,1}^{(1)}\\ \kappa_{t-1,1}^{(2)}\\ \kappa_{t-1,1}^{(3)}\end{pmatrix} + \begin{pmatrix}\varepsilon_t^{(1)}\\ \varepsilon_t^{(2)}\\ \varepsilon_t^{(3)}\end{pmatrix};$$ (multivariate random walk with drift)

$$\iota_h - \iota_{h-1} = b_0 + b_1(\iota_{h-1} - \iota_{h-2}) + \omega_h;$$ AR(1) process

$$\begin{pmatrix}\kappa_{t,2}^{(1)}\\ \kappa_{t,2}^{(2)}\end{pmatrix} = \begin{pmatrix}\varphi_{1,0}\\ \varphi_{2,0}\end{pmatrix} + \begin{bmatrix}\varphi_{1,1} & \varphi_{1,2}\\ \varphi_{2,1} & \varphi_{2,2}\end{bmatrix}\begin{pmatrix}\kappa_{t-1,2}^{(1)}\\ \kappa_{t-1,2}^{(2)}\end{pmatrix} + \begin{pmatrix}\delta_t^{(1)}\\ \delta_t^{(2)}\end{pmatrix},$$ bivariate VAR (1) process

The M7-M5 model in the IFoA/LLMA Phase 1 Report written by Haberman et al. (2014) (model (10)) extends the Cairns, Blake, and Dowd (2006) model. The reference population is fitted with the M7 model, and the difference between the two populations is fitted with the M5 model. The time varying factors are the same as in models (7) and (8) regarding the CBD counterpart, with the addition of a cohort effect for the reference population ι_h. These time varying factors of the reference population are noted as $(\kappa_{t,1}^{(1)}, \kappa_{t,1}^{(2)}, \kappa_{t,1}^{(3)})'$, in which $(\theta_1, \theta_2, \theta_3)'$ is the vector of the drift terms and $\left(\varepsilon_t^{(1)}, \varepsilon_t^{(2)}, \varepsilon_t^{(3)}\right)'$ are the multivariate normal error terms. The time varying factors of the subpopulation are $(\kappa_{t,2}^{(1)}, \kappa_{t,2}^{(2)})'$, where $\varphi_{1,0}$, $\varphi_{1,1}$, $\varphi_{1,2}$, $\varphi_{2,0}$, $\varphi_{2,1}$, and $\varphi_{2,2}$ are the coefficients of the bivariate VAR(1) process and $(\delta_t^{(1)}, \delta_t^{(2)})$ are the bivariate normal error terms which are independent of $\left(\varepsilon_t^{(1)}, \varepsilon_t^{(2)}, \varepsilon_t^{(3)}\right)'$. The following constraints are needed for the M7-M5 model: $\sum_h \iota_h = 0$, $\sum_h h\, \iota_h = 0$ and $\sum_h h^2 \iota_h = 0$.

Model (11):

$$\text{logit}(q_{x,t,1}) = \alpha_{x,1} + \beta_{x,1}\kappa_{t,1} + \iota_{t-x};$$ (CAE + cohort)

$$\text{logit}(q_{x,t,1}) - \text{logit}(q_{x,t,2}) = \alpha_{x,2} + \beta_{x,1}\kappa_{t,2};$$ (CAE)

$$\kappa_{t,1} = \theta + \kappa_{t-1,1} + \varepsilon_t;$$ (random walk with drift)

$$\iota_h - \iota_{h-1} = b_0 + b_1(\iota_{h-1} - \iota_{h-2}) + \omega_h;$$ (ARIMA(1,1,0) process)

$$\kappa_{t,2} = \varphi_0 + \varphi_1\kappa_{t-1,2} + \delta_t;$$ (AR(1) process)

The common age effect with cohort (CAE+Cohorts) model in the IFoA phase 1 Report (model (11)) combines the common age effect model proposed by Kleinow (2015), where the sensitivity measure is identical across the two populations, with the Jarner and Kryger (2011) approach, where the reference sensitivity measure is used as the common sensitivity measure. The common age effect model with cohort effect is used for the reference population, and the difference between the two populations is fitted with another common age Effect model. The parameters $\alpha_{x,i}$ represents the base mortality rate at age x, $\kappa_{t,i}$ describes the overall mortality changes over time t, and $\beta_{x,1}$ is the sensitivity measure for the reference population, which is also used as the sensitivity measure for the difference between the two populations. The terms ι_h depicts the cohort effect, θ is the drift term, b_0, b_1, φ_0, and φ_1 are the coefficients of the time series processes, and the error terms ε_t, ω_h, and δ_t are independent and normally distributed. The following constraints are needed for the CAE+Cohorts model: $\sum_x \beta_{x,1} = 1, \sum_t \kappa_{t,1} = 0,\ \sum_h \iota_h = 0$ and $\sum_t \kappa_{t,2} = 0$.

Common and specific factors

For this group of models, the two populations are modelled together by a common factor explaining the long-term trend for both populations and an additional factor allowing for short-term deviations from the main trend of each population. This model structure addresses the possible divergence issue from the Lee-Carter model but the short term forecast uncertainty is widened, since there are multiple trends imposed by the model. This class of models requires both populations to have the same data period and age range, otherwise the common period effect only reflects on one population when the other population data is missing. The data for this study involve different periods for the reference population and book population. Direct fitting is not applicable in this study and further details concerning to the model fitting for this class of models are detailed in section 3.3.

Model (A1)

$\ln m_{x,t,i} = \alpha_{x,i} + B_x K_t + \beta_{x,i}\kappa_{t,i};$ (augmented common factor model)

$K_t = \Theta + K_{t-1} + \delta_t;$ (random walk with drift)

$\kappa_{t,1} = c_{0,i} + c_{1,i}\kappa_{t-1,1} + \varepsilon_{t,i},$ (AR (1) process)

Li and Lee (2005) proposed the augmented common factor model, named as model (A1) here. One of the features of the augmented common factor model is that the projected ratio of death rates at each age converges to a constant if the fitted AR(1) processes are stationary for both populations' additional

factors. The terms $\alpha_{x,i}$ represents the base mortality rate at age x, B_x refers to the common factor regarding the age effect, and K_t refers to the common factor regarding the period effect. The other terms $\beta_{x,i}$ and $\kappa_{t,i}$ correspond to the additional factor for age and period respectively for population i. The terms Θ, $c_{0,i}$, and $c_{1,i}$ are the coefficients of the time series processes, and the error terms δ_t, $\varepsilon_{t,1}$, and $\varepsilon_{t,2}$ are independent normal random variables.

Model (A2)

$$\ln m_{x,t,i} = \alpha_{x,i} + B_x K_t + \sum_{j=1}^{n} \beta_{x,i,j} \kappa_{t,i,j} \; ; \qquad \text{(generalised common factor model)}$$

$$K_t = \Theta + K_{t-1} + \delta_t; \qquad \text{(random walk with drift)}$$

$$\kappa_{t,i,j} = c_{0,i,j} + c_{1,i,j}\kappa_{t-1,i,j} + \cdots + c_{r,i,j}\kappa_{t-r,i,j} + \varepsilon_{t,i,j}, \qquad \text{(AR(}r\text{) process)}$$

Li (2013) proposed the Poisson common factor model (model (A2)), which is an extension of the augmented common factor model. It incorporates more than one factor for the short-term deviations of each individual population from the common trend. The model provides better fitting and adopts standard statistical technique for parameter estimation. One of the limitations is the possibility of overfitting, which can be avoided by choosing the optimal number of additional factors. The subscript j represents the jth additional factor for the short-term deviations. The BIC values determine the desired number of additional factors n. The order for the autoregressive process r is chosen by examining the partial autocorrelation functions for $\kappa_{t,i,j}$ and whether the projected ratio of death rates converges.

Wong (2016) further extended the Poisson common factor model with variable sex-specific factors (PCFM-VSF) to improve fitting performance. In this model the number of additional factors for each sex can differ to reduce the overfitting issue. Yang et al. (2016) incorporated a cohort parameter into the Poisson common factor model to improve the fitting and to allow for the existence of cohort effects in a multiple-population setting.

Other models with joint fitting

Kleinow (2015) proposed the common age effect model to allow the populations to share a common sensitive measure in the Lee-Carter structure, where each population has a population-specific time

varying factors. Similar to Booth et al. (2002), the model can also be extended to incorporate additional bilinear factors if it improves the fitting performance. The model outperforms the augmented common factor model in old age groups for developed countries as these populations exhibit similar age effect, as shown in Kleinow (2015).

The joint-k model (Lee and Carter,1992; Li and Lee 2011) uses common time varying factor for the populations. It has the parsimonious model advantage of just using one time-varying factor for two populations, but it also brings about a major limitation where the two populations are perfectly correlated, which is difficult to cater for any short-term deviations, leading to an underestimation of the demographic basis risk.

To allow for the structure of mortality dependency, Wang et al. (2015) extended the error term assumption for the mortality index from the Lee-Carter Model using non-Gaussian innovations such as the student's t distribution. The authors also modelled the multi-population mortality dependency using time-varying copula and showed that non-Gaussian copulas provide a better fit since there is tail dependency for the multi-population mortality index.

Other multi-population model

Hyndman et al. (2013) utilized a different approach by proposing a product-ratio model to fit the product and ratio of mortality rates of the two populations instead of the original mortality rate, motivated by fact that the product and ratio are roughly uncorrelated if the two populations have similar variances. This model results in similar accuracy as the single population model while also ensures that the long-term forecast ratio will converge. The limitation of this model is that the convergence ratio is sensitive to the choice of the fitting period.

Danesi et al. (2015) adapted the mortality improvement definition from Haberman and Renshaw (2012) and modelled the mortality improvement rate jointly using the Lee-Carter structure and their variants. The results showed that common time varying factors provide the best trade-off between parsimony and goodness of fit. Other contributions to the multi-population modelling include Chen et al. (2015) who adopted a factor copula approach after fitting the mortality improvement rates with an ARMA-GARCH model in order to explain the mortality co-movements for multiple populations.

Chapter 3

Data and methodology

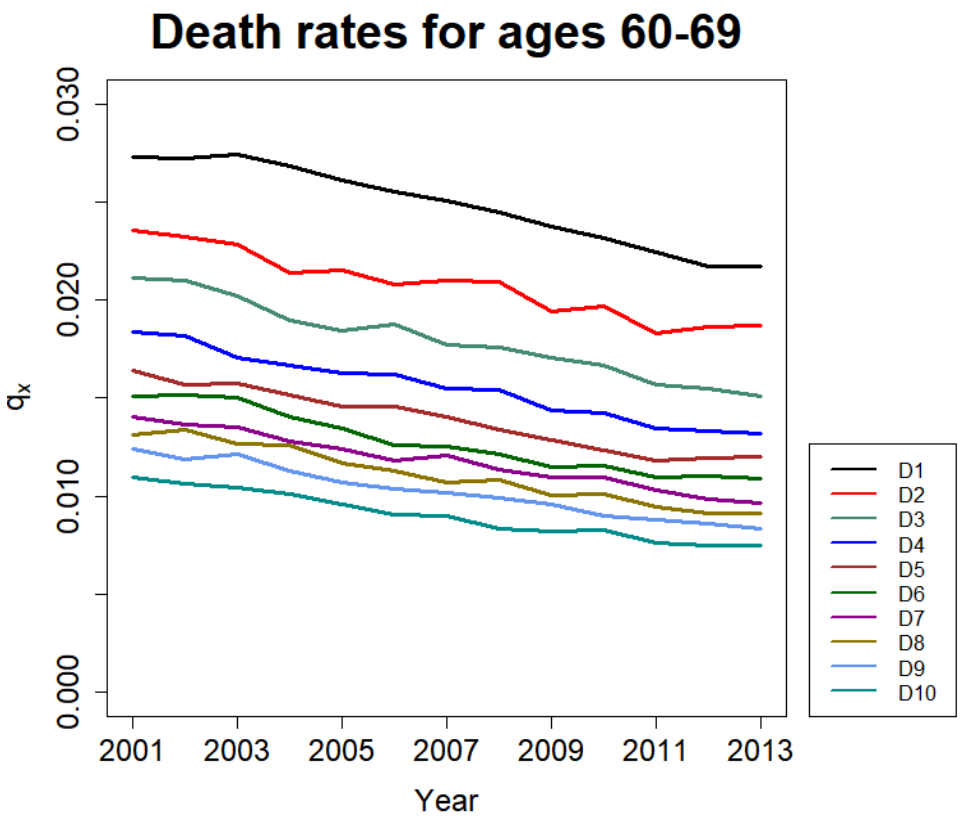
3.1 Data description

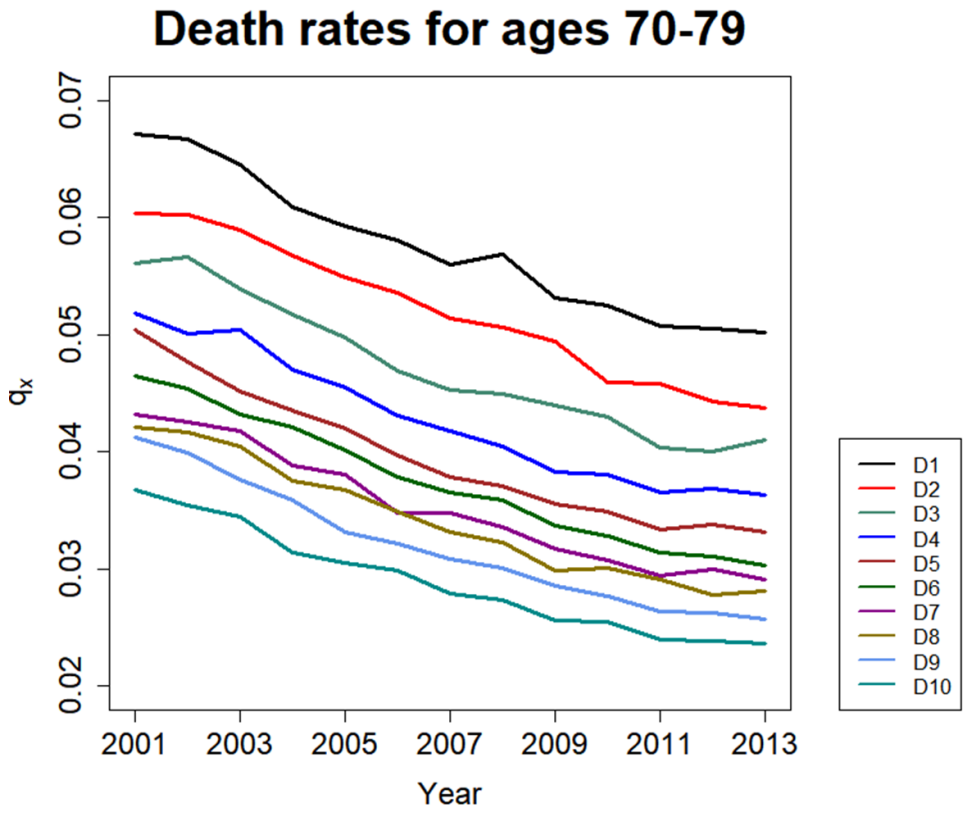
The mortality data used in this thesis are collected from the Office for National Statistics (ONS) in the UK. This dataset is split by sex, age, year, and decile group based on the Index of Multiple Deprivation (IMD). The welfare status of a subgroup is determined by the deprivation level, which is a measure for the socioeconomic status. It is calculated according to the following aspects: income, employment, education, skills, and training, health and disability, crime, barriers to housing and services, and living environment. Different weights are assigned to these seven categories to determine the final deprivation score. The deprivation level is measured for groups of approximately 1,500 people who live in the same area (known as the Lower Layer Super Output Area (LSOA)). These areas are first ranked based on their deprivation levels amongst all areas. Then the ranked areas are split into ten subgroups of equal sizes, each of which forms a decile. Further information on the classifications can be found in the English Indices of Deprivation 2015 Statistical Release. The dataset consists of population exposures and death counts in the UK from 2001 to 2013, for individual ages from 0 to 89 and 90+. Because the mortality models in this thesis require data for each age and each year, those data of aged 90+ are not used. The younger age groups are also omitted, as the number of deaths of younger people in each sub-population is generally very small, which would cause rather inaccurate model fitting and projections, and as the major recent concern is the longevity risk of retirees. Accordingly, the focus is on the age range of 60-89.

The whole UK population is chosen as the reference population. The dataset is obtained from the Human Mortality Database (HMD 2017). The reference population data covers a wider period of 1980-2013. This longer time period of data could lead to more accurate estimation of the age and cohort effects. The age range is also set as 60-89 for the reference population, in line with the treatment of the ONS data.

3.2 Preliminary analysis for the deprivation subgroup data

The average death rates for ages 60-69, 70-79, and 80-89 over time for each decile are plotted below with D1 representing the most deprived decile subgroup, D2 referring to the second most deprived decile subgroup, and so on, up to D10 which stands for the least deprived decile subgroup. Using age groups rather than individual ages for demonstration here makes the mortality trends more distinct between different deciles. However, as shown in Figure 3.2.1, some overlapping in mortality rates is still observed for the older age groups 70-79 and 80-89. Further grouping the deciles into quintiles would help reduce the extent of overlapping, the average death rates of which are then plotted in Figure 3.2.2.





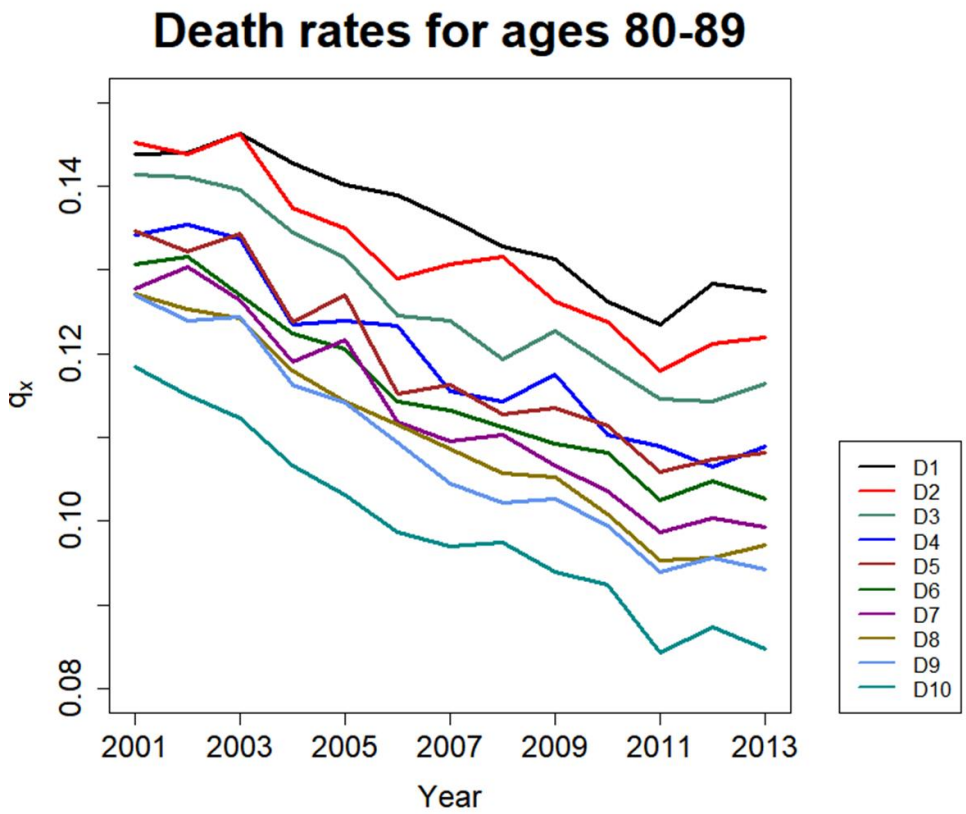
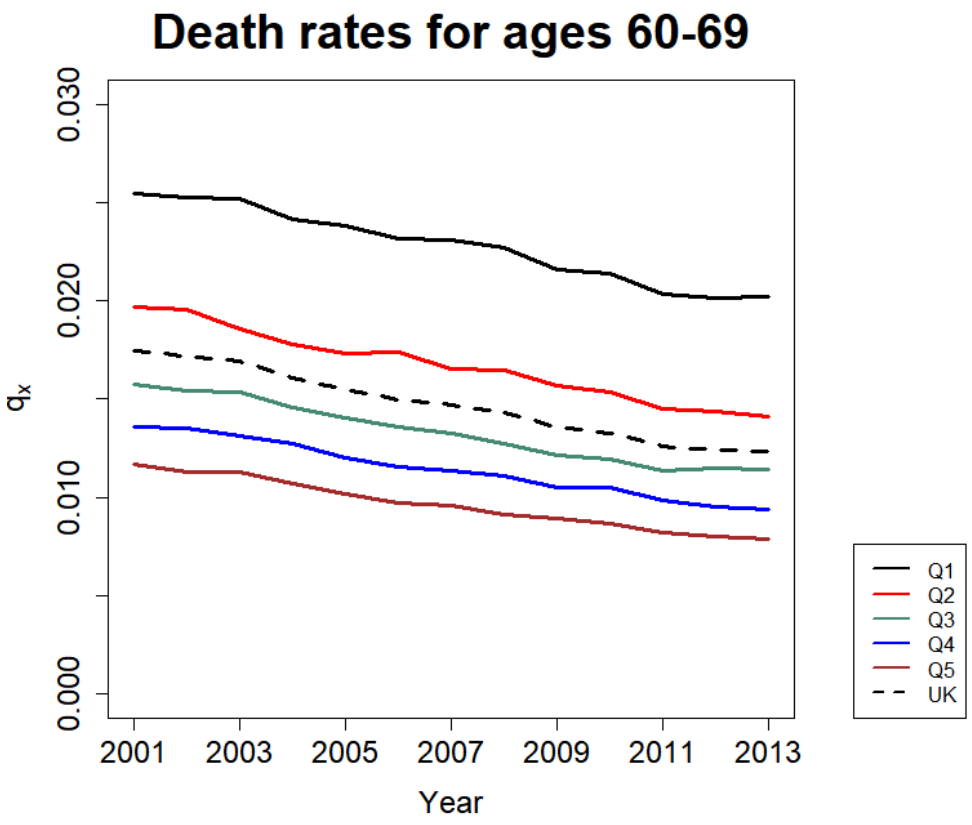
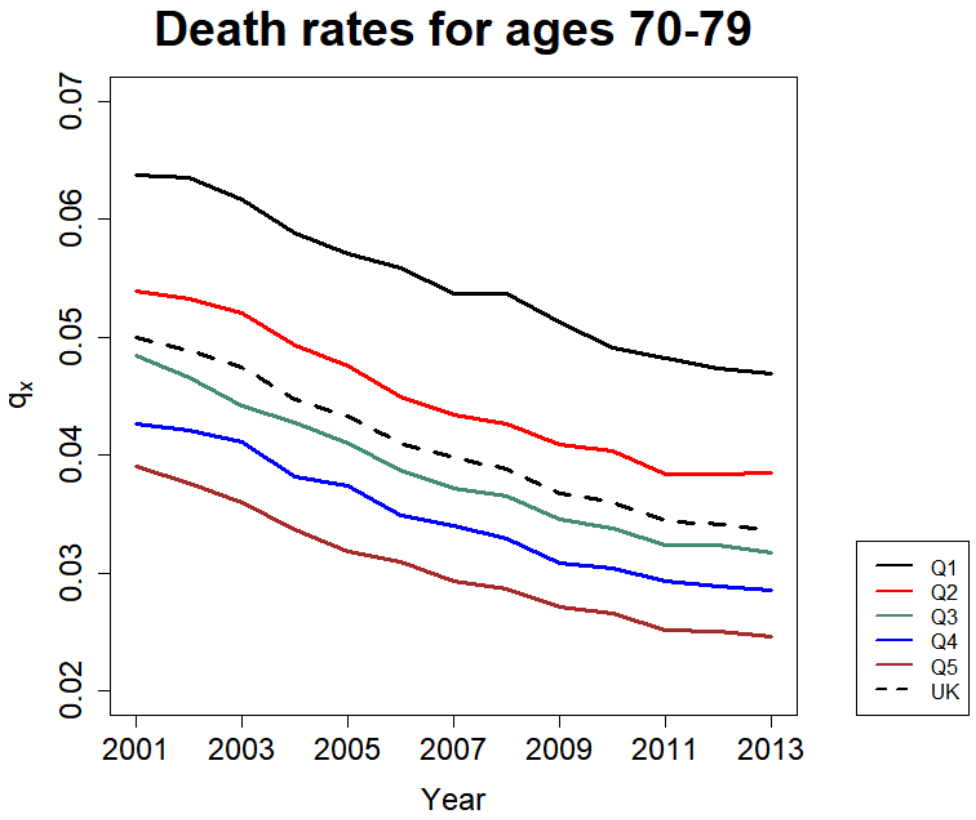


Figure 3.2.1: Average mortality rates for ages 60-69, 70-79, and 80-89 from 2001 to 2013 for different decile groups in the UK.

In Figure 3.2.2, the whole UK population's average death rates of the three age groups are also included for comparison. There is no clear overlapping between the quintiles for the three age groups. The UK population's average death rates lie between those of the second most deprived quintile and the third most deprived quintile over all years for the three age groups, with an exception for ages 80-89 in years 2009-2013, when the UK population's average death rates lie between the third most deprived quintile and the fourth most deprived quintile instead. In order to identify those subgroups with distinct mortality patterns and trends for modelling purposes, this research will group the deciles into quintiles in the remaining analysis. In addition, using the quintiles would increase the exposures of each subgroup and so reduce sampling variability.





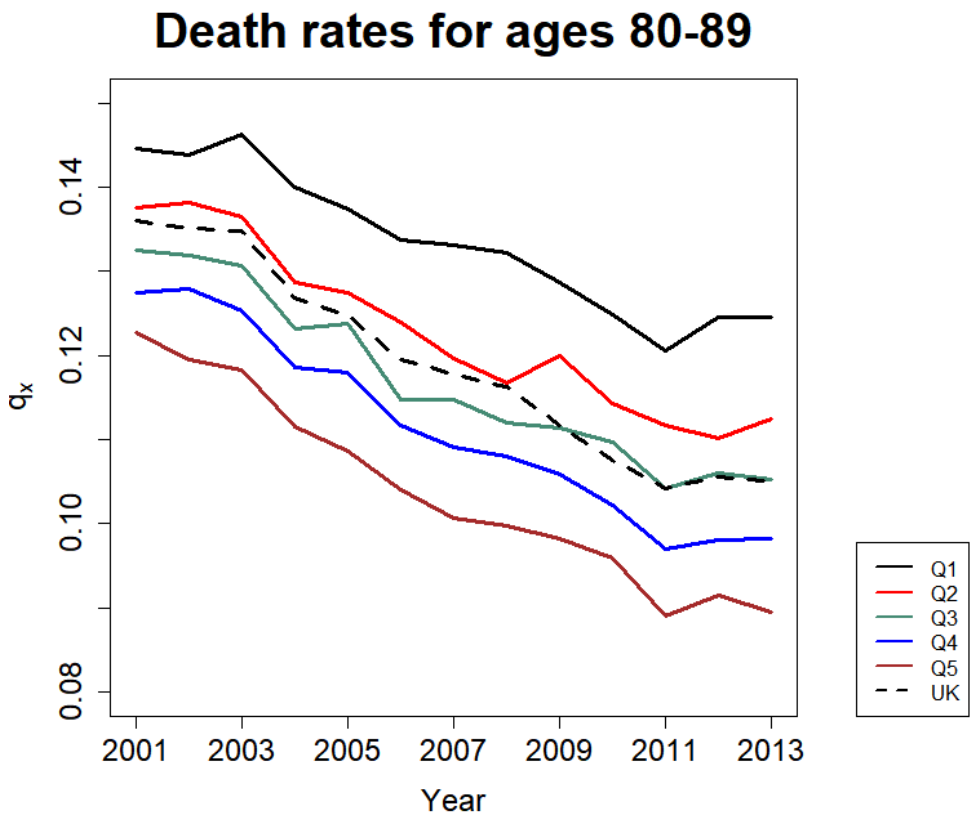


Figure 3.2.2: Average mortality rates for ages 60-69, 70-79, and 80-89 from 2001 to 2013 for different quintile groups in the UK.

3.3 Model fitting procedure

Multi-population mortality modelling is a relatively new research area. In this thesis, several models from the literature have been tested on the UK population and the England and Wales deprivation subgroups. A general framework of model fitting is applied for all the models being considered. The basic assumptions underlying the general framework are listed as follows:

1) The force of mortality $\mu_{x,t,i}$ is constant between integer ages and calendar years for population i. Then $\mu_{x,t,i} = m_{x,t,i}$, where $m_{x,t,i}$ is the central death rate at age x in year t for population i.

2) $q_{x,t,i} = 1 - \exp(-m_{x,t,i})$ where $q_{x,t,i}$ represents the probability that an individual aged x years old at time t from population i will die within a year.

(The two assumptions above are often used when comparing models using the central death rate $m_{x,t,i}$ and the one-year probability death rate $q_{x,t,i}$, for example, those in Cairns (2009) and Li et al (2015).)

3) $D_{x,t,i} \sim Poi(E_{x,t,i}, m_{x,t,i})$ where $D_{x,t,i}$ represents the random number of deaths at age x in year t for population i and $E_{x,t,i}$ represents the corresponding (known) exposure to risk.

The Poisson distribution for modelling the number of deaths has been a common choice in the literature, such as Li (2013) and Li et al (2014). As mentioned by Li (2013), the Poisson assumption provides a rigorous statistical framework for mortality data analysis. The parameters in the model can then be estimated by maximum likelihood estimation. Model selection criteria such as the Akaike information criterion (AIC) or Bayesian information criterion (BIC) can be used to assess the goodness-of-fit and the parsimony of the fitted model.

The parameters of the models described below are estimated using the maximum likelihood method. The log likelihood function under the Poisson death count assumption is expressed as:

$$l = \sum_{x,t,i} [d_{x,t,i} \ln(E_{x,t,i}\, \widehat{m}_{x,t,i}) - E_{x,t,i}\, \widehat{m}_{x,t,i} - \ln(d_{x,t,i})!]$$

where $d_{x,t,i}$ is the actual number of deaths at age x in year t for population i, and $\widehat{m}_{x,t,i}$ is the fitted central death rate based on the model applied. The parameters are updated repetitively using the Newton-Raphson method: $\theta^* = \theta - \frac{\partial l/\partial \theta}{\partial^2 l/\partial \theta^2}$, where l refers to the log likelihood and θ^* refers to the updated parameter. As it is possible for some models to encounter the identification problem[2], additional constraints are required for those models in order to produce a unique solution.

[2] Identification problem exist when there is no unique solution for the parameter being estimated.

The temporal parameters are then modelled with specific time series processes, the details of which are provided in Section 5. When co-modelling these parameter time series between the UK population and the deprivation subgroup, there is a data constraint that the fitting periods of the two populations are not equal. One possible approximation is to take the year range 2001-2013 for the subgroup population and use the same range for the reference population when fitting the time series process. This approximate approach is adopted only for those models which involve modelling two distinct parameter time series jointly.

Further modification for model (7) required in this study. Since the subpopulation's data only cover a period of 13 years, it is impossible to fit a VARMA process with the exception for $p + q = 1$. For the $p + q = 2$ case, it requires $13 \times 6 = 78$ parameters to be estimated in the model, but the sample size is only $12 \times 6 = 72$ for the case when $d = 1$. With the only choice being the VARMA(1,0) process and $d = 1$, the coefficient matrix is unstable due to the small sample size. An approximation method may be applied here: an AR(1) process is used for the first difference for each of the 6 time varying factors, while the error terms remain to be distributed as multivariate normal.

For model(A1) and model (A2), the period covered by the reference population data is not equal to the period of the sub-population data. To fit these models, 13 years of reference population data from 2001-2013 will be used instead, but then the corresponding results cannot be directly compared with those of the previous 11 models due to the different periods of reference data being tested. The fitting results for models (A1) & (A2) are shown in the Appendix.

3.4 Forecasting method

In the literature, there are several methods for simulating future death rates. The simplest approach is to simulate the random error $\varepsilon_{x,t,i}$, which is normally distributed as specified above regarding the time series models. There is a main drawback for this method as it only considers the process error while omitting the parameter uncertainty from the model fitting. Li (2014) performed a comprehensive analysis in comparing several simulation strategies. One simulation method stated in the paper which also includes the parameter error is residuals bootstrapping, proposed by Koissi et al. (2006). In this bootstrapping process, the first step is to resample the deviance residuals $r_{x,t,i}$ with replacement. As the number of deaths $d_{x,t,i}$ follows a Poisson distribution, the next step is to apply the inverse formula to the resampled residuals to obtain the pseudo data sample of the number of deaths. Then the model parameters are estimated based on the pseudo data sample, the temporal parameters of which are projected over time incorporating the random error.

Liu and Braun (2010) extended the residuals bootstrapping to the block residuals bootstrapping to cater for the serial correlations between the residuals within the population. In this block bootstrapping process, one first needs to split the residuals into overlapping blocks of equal sizes. Then random samples of the blocks, instead of individual residuals, are drawn with replacement. In this research, a block size of 5 (see Li and Haberman, 2015) is used to preserve the autocorrelations between the residuals in the population. Note that the deprivation subgroup data only span over 13 years and so the block size cannot be too large.

The block residuals bootstrapping used in this study need to be further modified, since the UK population and the deprivation subgroups are related demographically in some way. The relationship between the two populations would be left out if the block residuals bootstrapping is performed independently for each population. To preserve the association between the populations, the residuals blocks of the two populations are matched in pairs for each age-time cell before the resampling. In this way, the block residuals bootstrapping can be first carried out for the UK population, then the matching blocks of the residuals of the deprivation subgroup are selected accordingly.

Chapter 4

Fitting results and performance

4.1. Model Fitting: Parameter Estimation

The fitting results for the 11 models are shown in this section. All computations are conducted by using the R statistical software package (R Core Team, 2017). The parameter estimates of the models are plotted in Figures 4.1.1 to 4.1.8.

For the models with associated mortality indices (models (1) - (7)), the patterns of the parameter estimates of the deprivation subgroups are similar to those of the UK population, except for the sensitivity measure $\beta_{x,i}$ for models (1) – (3) and the cohort effect for models (5) & (6). Only this group of models allows for a direct comparison between the UK population and individual deprivation subgroups since the single population model is fitted to each population.

In relation to the deprivation subgroups in England under models (1) – (6), the base mortality rate $\alpha_{x,2}$ for the more deprived subgroups is higher compared to the less deprived subgroups. As age increases, the gap in the base mortality rates between the quintiles narrows, which is consistent with the preliminary analysis in Chapter 3. The steeper slope of the period effect $\kappa_{t,2}$ for the less deprived subgroups may imply that they have larger mortality improvements compared to the more deprived subgroups, though strictly speaking, the mortality indices of different model fits cannot be compared directly. Also, the parameter $\beta_{x,2}$ under models (1) - (3) has similar patterns across all deprived subgroups. The common sensitivity measure β_x under model (4) is highly similar across all deprived subgroups as the larger data size of the UK population dominates the parameter estimation. The plots of $\iota_{h,2}$ under models (5) & (6) show a similar profile of cohort effect across all deprived subgroups with a spike around cohort year 1920.

Under model (7), $\kappa_{t,2}^{(1)}$ distinguishes the base mortality level for each deprivation subgroups and reflects the overall mortality decline over time. The slope parameter $\kappa_{t,2}^{(2)}$ experiences a slightly upward trend for all quintiles. The quadratic age term used to describe the curvature of the logit rates across age, $\kappa_{t,2}^{(3)}$, also has an upward trend for all quintiles.

For model (8), up to five age and period factors are fitted to the log ratio of the death rate. The BIC values show that only one age and period factor are needed since this choice leads to the lowest BIC values as shown in Table 4.1.1. The age factor $\psi_x(1)$ exhibits a downward slope for all deprivation subgroups because the mortality gap decreases as age increases. Additionally, it is worth noticing that the magnitude for the age factor in the third most deprived quintile is different from other quintiles since the corresponding period factor values are closer to zero. The slope for the period factor $\gamma_t(1)$ is similar in all but the most deprived quintile. These results show that the UK population experiences better mortality improvement compared to the most deprived quintile.

For model (9), the age factor a_x shows the age effect of the log ratio between the deprivation subgroup and the UK population. The gap in this age effect between different quintiles narrows for the older lives. For the most deprived quintile, the period factor k_t moves in the opposite direction to those of the other quintiles to offset the higher mortality improvement in the overall UK population. Note that the scale of the sensitivity measure b_x used in the fourth most deprived quintile is different to those of the other quintiles because the corresponding value of k_t is very close to zero.

For model (10), $\kappa_{t,2}^{(1)}$ represents the base level of difference in logit mortality between the UK population and the deprivation subgroups. Unlike other quintiles, the most deprived quintile depicts an upward trend of $\kappa_{t,2}^{(1)}$ because of similar arguments as for models (8) and (9). The slope of $\kappa_{t,2}^{(2)}$ fluctuates around its own level for all quintiles.

Under model (11), $\alpha_{x,2}$ represents the base level of difference in logit mortality between the UK population and the deprivation subgroup. The gap in the differences between the quintiles narrows at the older ages. Moreover, $\kappa_{t,2}$ of the most deprived quintile moves in the opposite direction compared to those of the other quintiles.

Looking at the parameter estimates of the UK population, there is one interesting finding on the cohort effect for models (5), (6), and (9) to (11). While the general trends may differ among the models, they all exhibit a spike around the cohort year 1920. The cohort effect will be discussed in more detail in the next section.

Model (1)-(3):

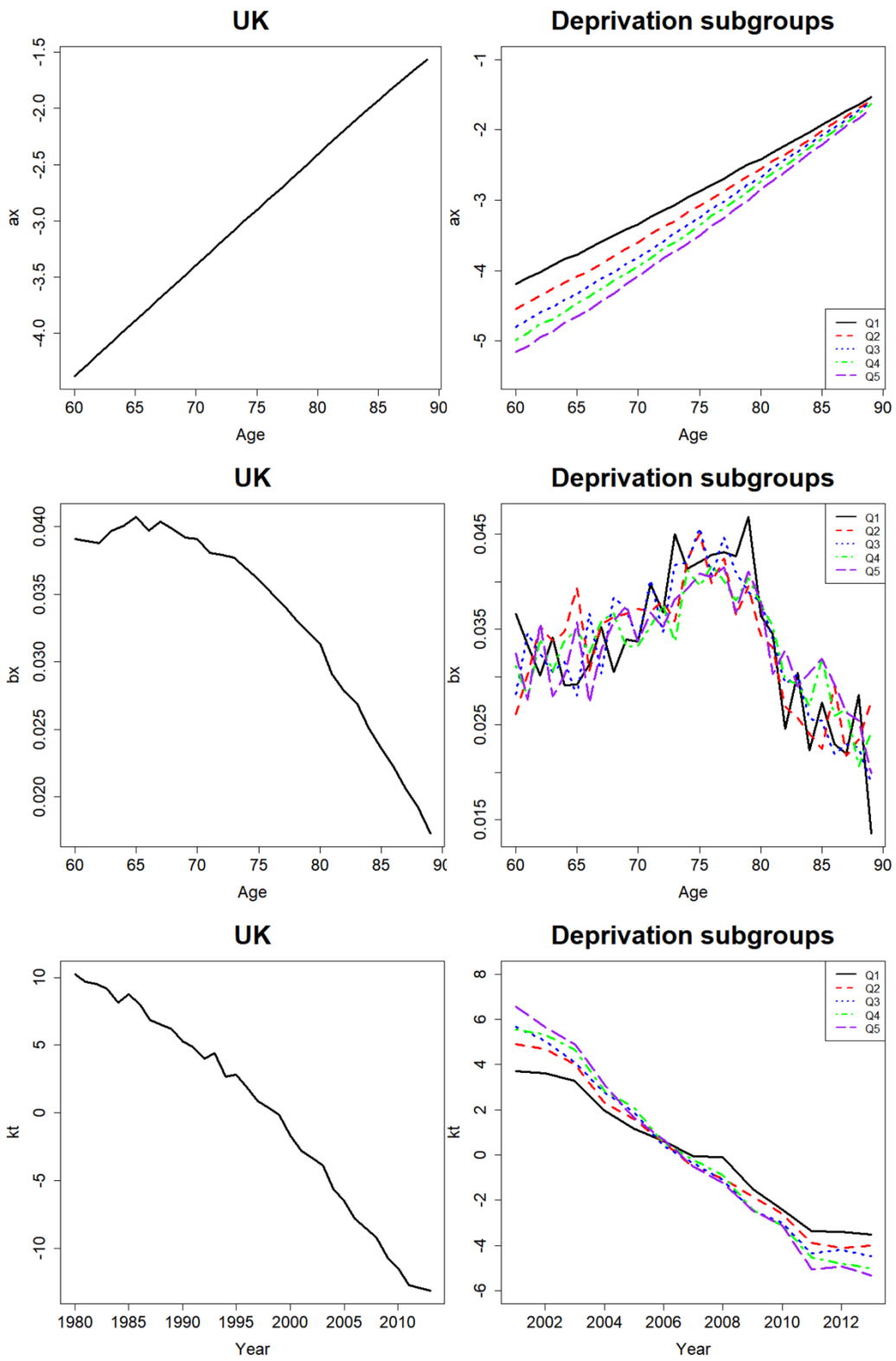


Figure 4.1.1: Parameter estimates of the Lee-Carter model for UK population (left) and the deprivation subgroups in England (right).

Model (4):

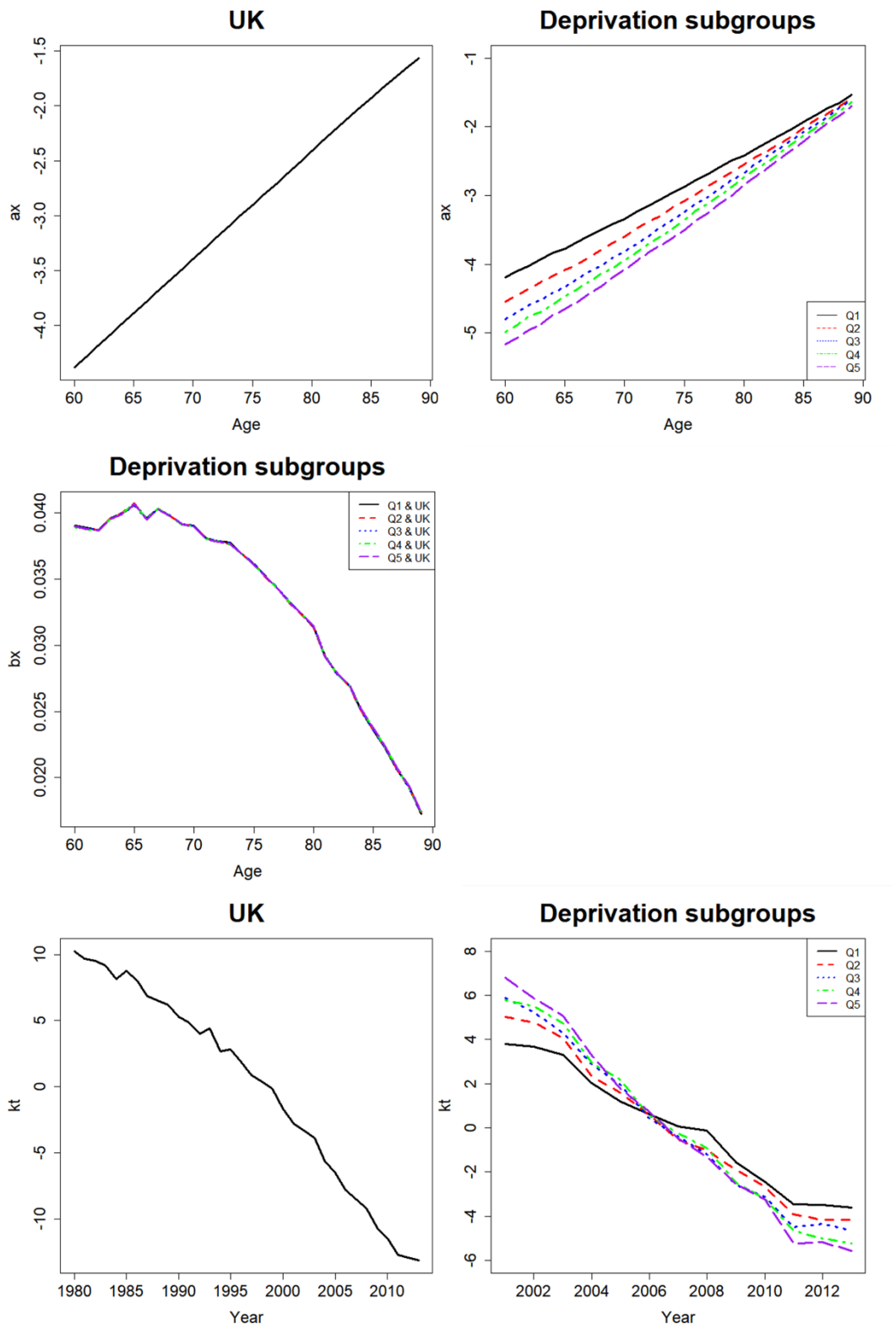


Figure 4.1.2: Parameter estimates of the Lee-Carter model with common sensitivity measure for UK population (left) and the deprivation subgroups in England (right).

Model (5) & (6):

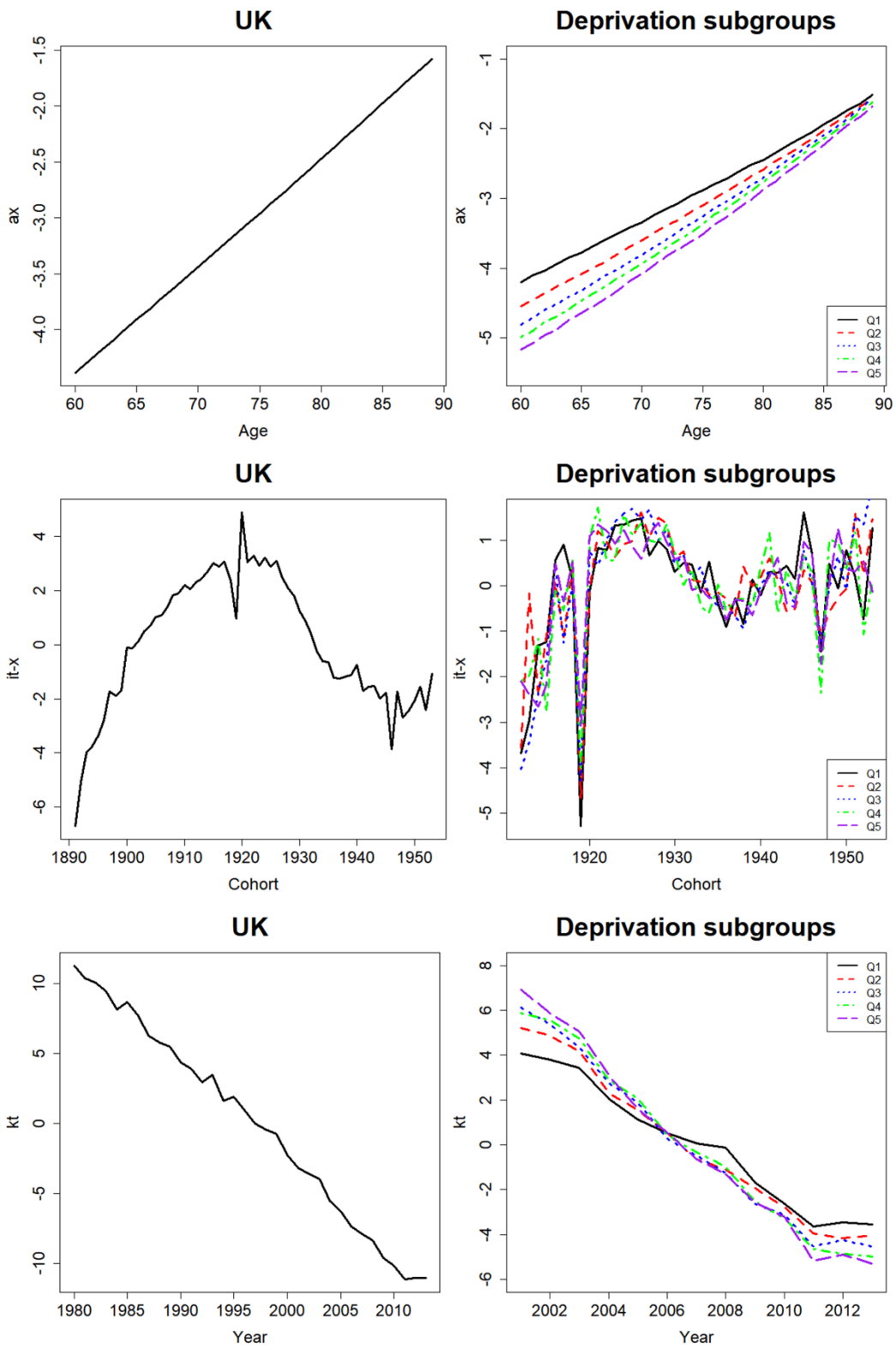


Figure 4.1.3: Parameter estimates of the age-period-cohort model for UK population (left) and the deprivation subgroups in England (right).

Model (7):

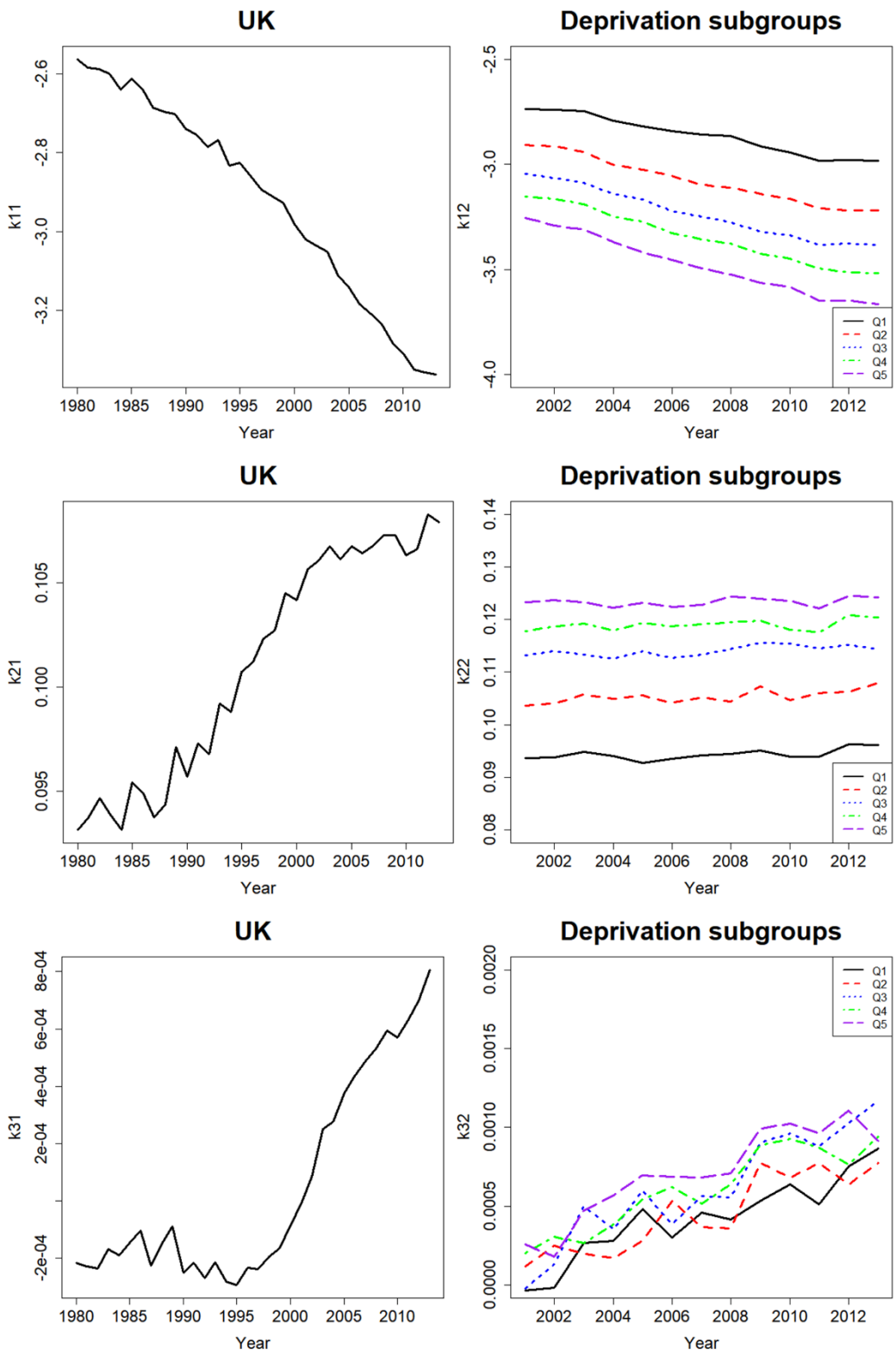


Figure 4.1.4: Parameter estimates of the CBD M7 model for UK population (left) and the deprivation subgroups in England (right).

Model (8):

Table 4.1.1 BIC values for using different numbers of age and period factors for the ratio of mortality rates between the deprivation subgroup in England and UK population under model (8).

Number of factors	Most deprived quintile	2nd most deprived quintile	3rd most deprived quintile	4th most deprived quintile	Least deprived quintile
1	**4517**	**4354**	**4252**	**4283**	**4193**
2	4611	4467	4395	4401	4325
3	4729	4614	4537	4543	4477
4	4875	4769	4731	4702	4654
5	5042	4948	4914	4891	4846

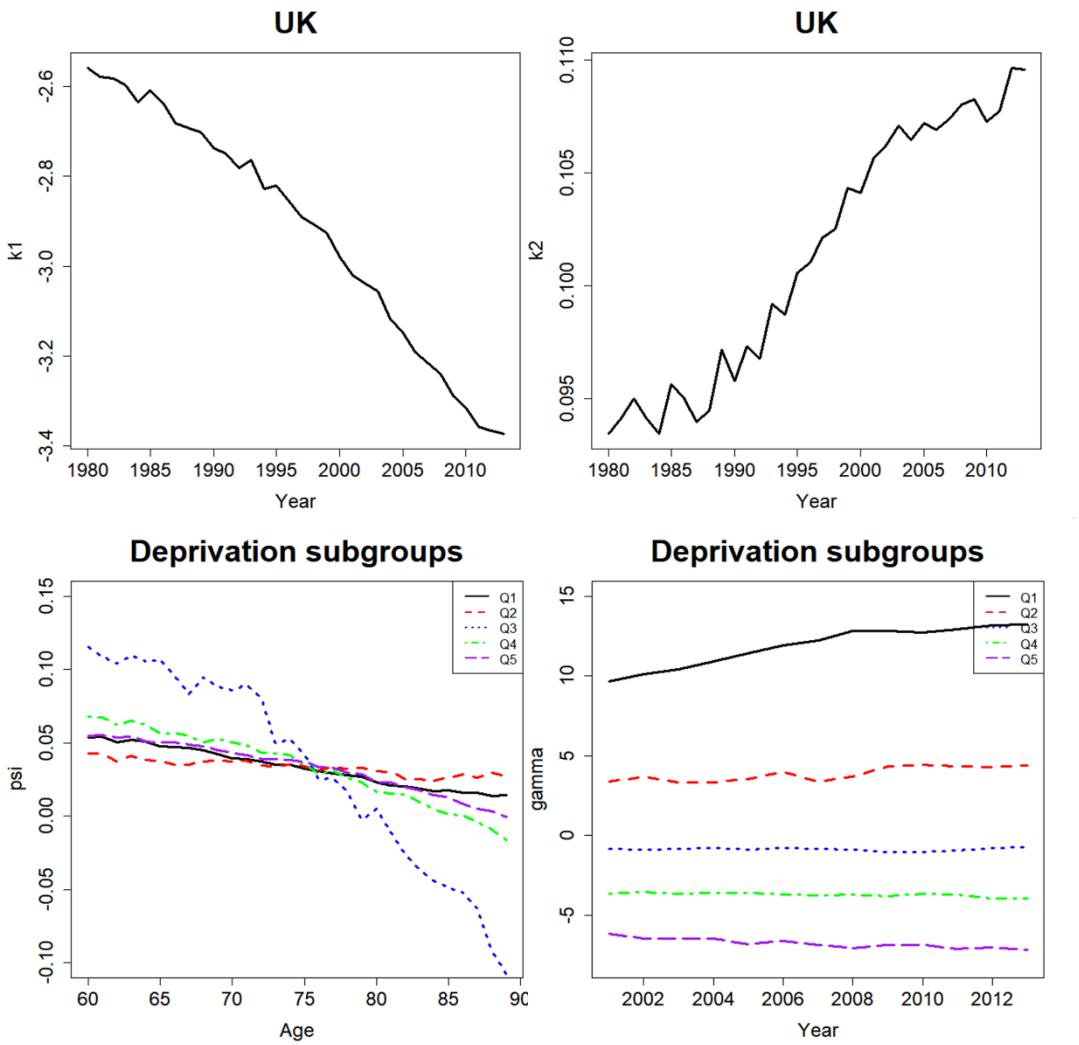


Figure 4.1.5: Parameter estimates of the CBD M5 model for UK population (top) and the age and period factor estimates for the ratio of mortality rates between the deprivation subgroup in England and UK population (bottom).

Model (9):

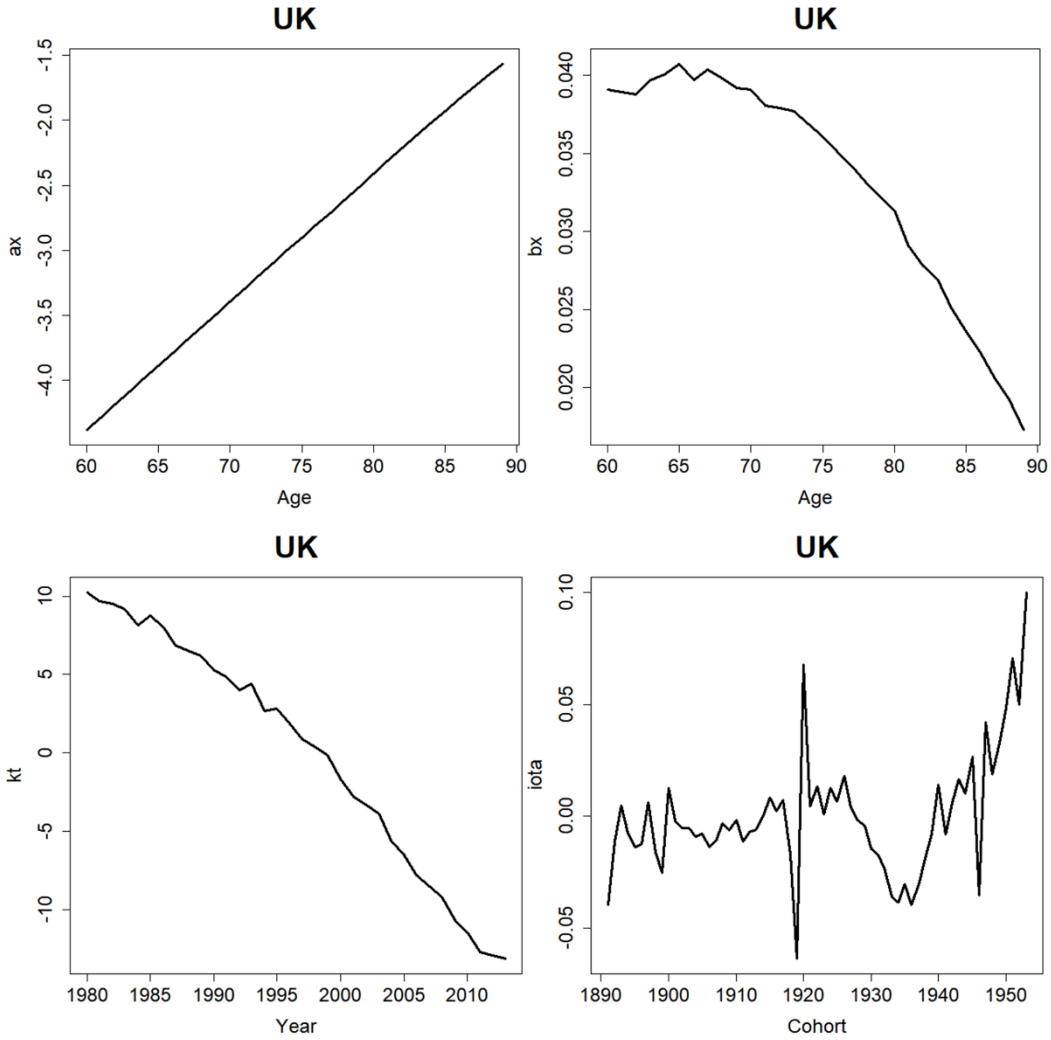


Figure 4.1.6a: Parameter estimates of the Lee-Carter model with the cohort effect for UK population

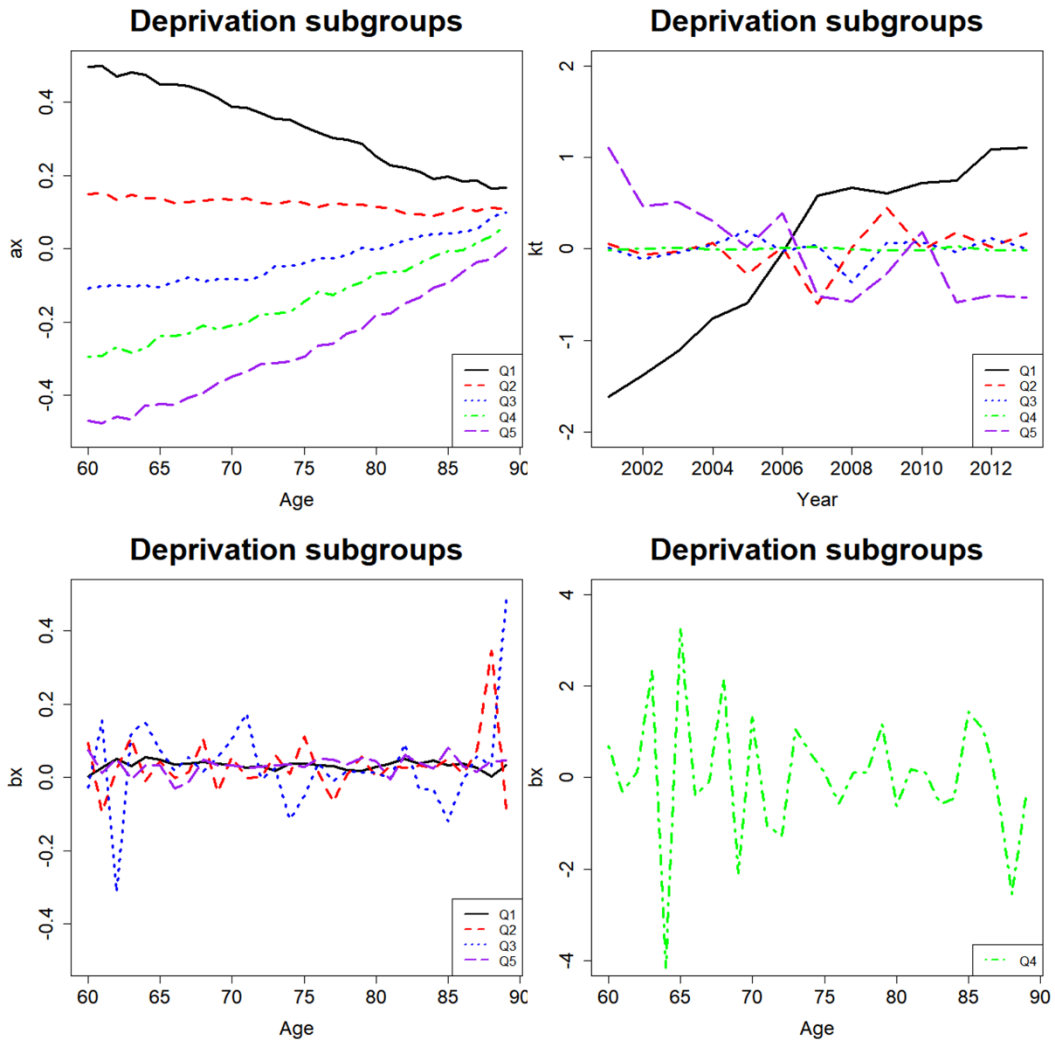


Figure 4.1.6b: The age and period factor estimates for the log ratio of mortality rates between the deprivation subgroup in England and UK population.

Model (10):

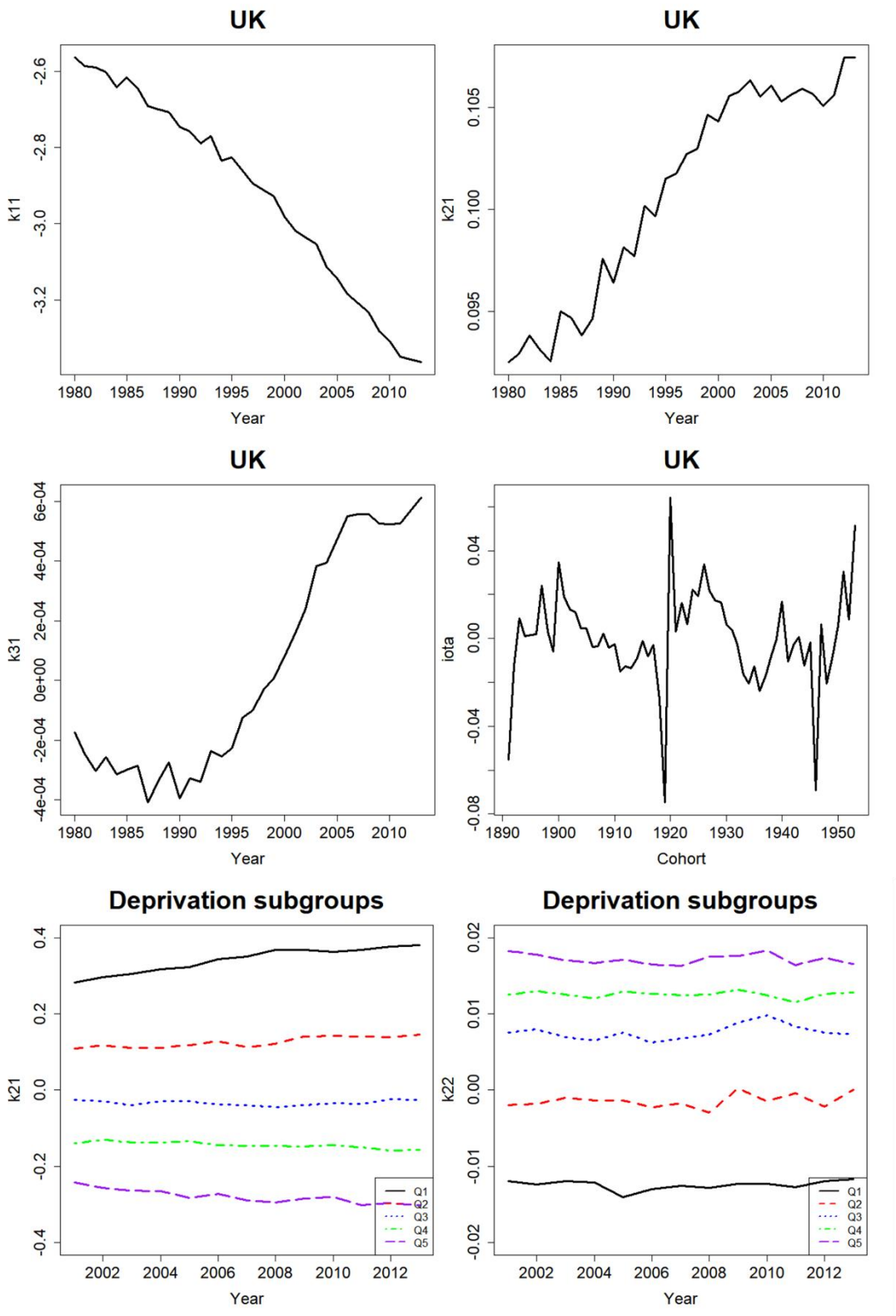


Figure 4.1.7: Parameter estimates of the CBD M7 model for UK population (top and middle) and parameter estimates of CBD M5 model for the difference between the UK population and deprivation subgroup in England (bottom).

Model (11):

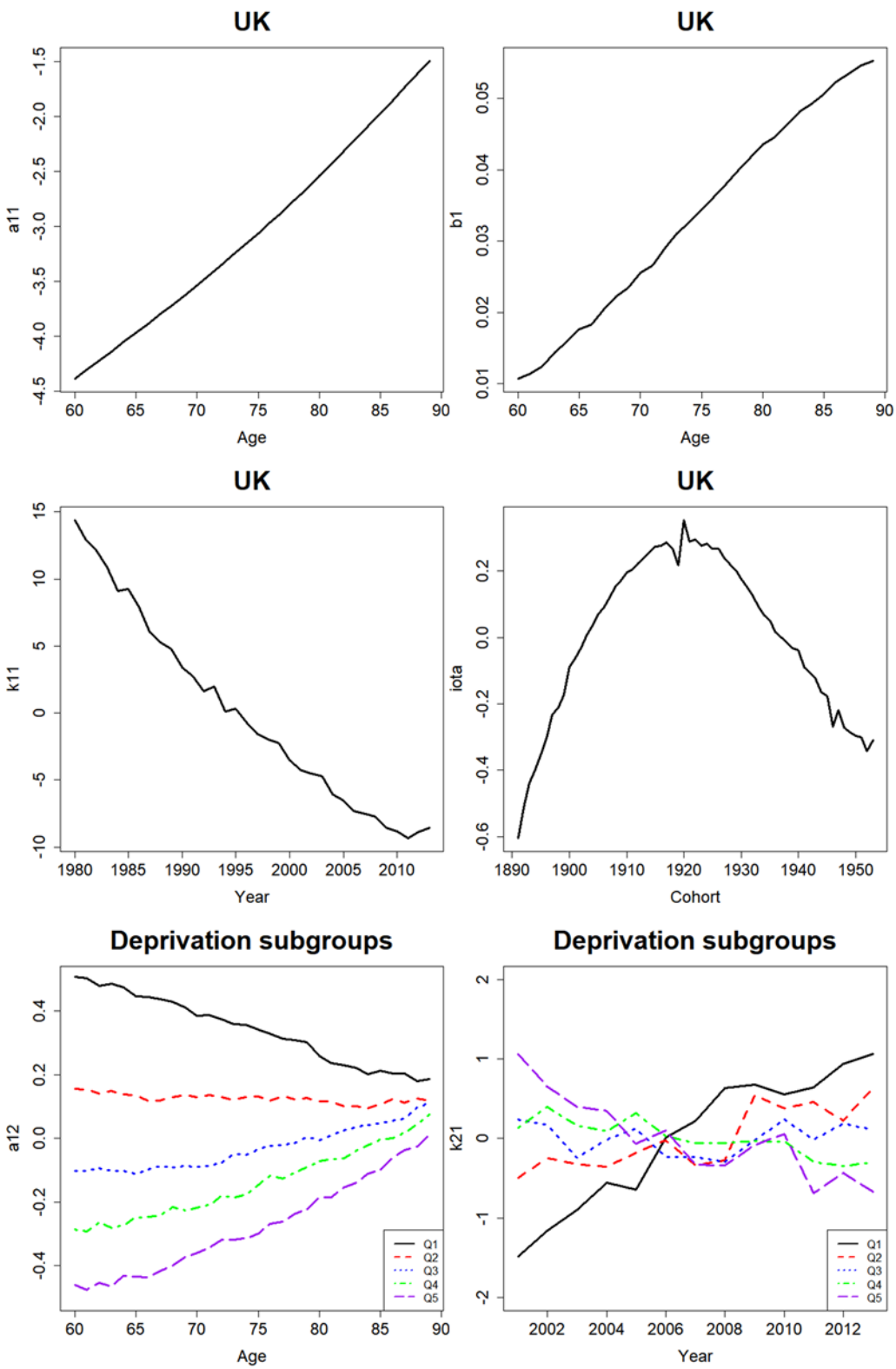


Figure 4.1.8: Parameter estimates of the CAE + cohort model for UK population (top and middle) and parameter estimates of CAE model for the difference between the UK population and deprivation subgroup in England (bottom).

4.2 Model Fitting performance

4.2.1 Bayesian Information Criterion

The fitting performance for the 11 models is discussed in this section. As mentioned previously in Chapter 3, the BIC (Bayesian Information Criterion) can be used to access both the goodness-of-fit and the parsimony of the fitted model. A parsimonious model is a model that does not use too many parameters whilst still explaining an adequate level of the data, since the log-likelihood will always increase with more estimated parameters. To overcome the problem of overfitting, a common approach is to penalise the use of excessive model parameters. Note that there is a larger penalty for including an extra parameter under the BIC than the AIC (Akaike Information Criterion). The BIC is calculated as:

$$\mathrm{BIC} = -2\hat{l} + n_p \ln(n_d)$$

where $\hat{l}$ is the log-likelihood, n_p is the number of effective parameters to be estimated for the model, and n_d is the number of observations.

Table 4.2.1 BIC values for UK & deprivation subgroups in England

BIC	Most deprived quintile	2nd most deprived quintile	3rd most deprived quintile	4th most deprived quintile	Least deprived quintile
(1) (2) (3)	22386 [7]	22249 [7]	22169 [7]	22199 [7]	22021 [7]
(4)	22322 [6]	22156 [6]	22131 [6]	22156 [6]	22100 [6]
(5) (6)	18935 [4]	18887 [4]	18829 [4]	18800 [4]	18791 [4]
(7)	20637 [5]	20475 [5]	20416 [5]	20446 [5]	20267 [5]
(8)	23389 [8]	23226 [8]	23124 [8]	23155 [8]	23064 [8]
(9)	18232 [3]	18080 [3]	17974 [3]	17975 [3]	17899 [3]
(10)	17700 [1]	17505 [1]	17466 [1]	17483 [1]	17466 [1]
(11)	17921 [2]	17703 [2]	17576 [2]	17607 [2]	17520 [2]

The BIC values of the fitted models are shown in Table 4.2.1. The numbers in the square brackets rank the BIC values from the smallest to largest. The M7-M5 model (model (10)) has the lowest BIC value across all quintiles. It is worth noticing that the models that incorporate the cohort effect (i.e. models (5), (6), (9), (10), and (11)) produce the four smallest BIC values for all quintiles.

4.2.2 Mean Absolute Percentage Error

Another goodness-of-fit measure is the mean absolute percentage error (MAPE), which is used to examine the fitting accuracy of death rates. The MAPE is computed as:

$$\frac{1}{n_d}\sum_{x,t,i}\frac{\ln \widehat{m}_{x,t,i} - \ln\left(d_{x,t,i}/E_{x,t,i}\right)}{\ln\left(d_{x,t,i}/E_{x,t,i}\right)}$$

where n_d represents the number of observations, $d_{x,t,i}$ is the actual number of deaths at age x in year t for population i, $E_{x,t,i}$ is the corresponding exposed to risk, and $\widehat{m}_{x,t,i}$ is the fitted central death rate.

In Tables 4.2.2 to 4.2.6, the MAPE values of the log central death rate is shown for each quintile, where the MAPE values of the central death rate is shown in the bracket for comparison. All MAPE values of the log central death rates are small, indicating that all models have satisfactory model fitting. The same finding also holds true for the MAPE values of the central death rates. Similar to the BIC analysis above, the four models which allow for the cohort effect produce smaller MAPE values than the other models. The small differences in the MAPE values between the models that incorporate the cohort effect make it difficult to conclude which model provides the best fit due to the variations in the model rankings for different quintiles.

Table 4.2.2 MAPE values for UK & most deprived quintile

Model	UK	Most deprived quintile	Total	Rank
(1) (2) (3)	0.76% (2.19%)	1.16% (2.97%)	0.87% (2.40%)	[6]
(4)	0.76% (2.18%)	1.31% (3.39%)	0.91% (2.52%)	[7]
(5) (6)	0.58% (1.51%)	0.91% (2.39%)	0.67% (1.75%)	[3]
(7)	0.60% (1.66%)	1.21% (3.12%)	0.77% (2.07%)	[5]
(8)	0.86% (2.46%)	1.15% (2.95%)	0.94% (2.60%)	[8]
(9)	0.52% (1.13%)	1.11% (2.84%)	0.68% (1.60%)	[4]
(10)	0.36% (0.96%)	1.22% (3.17%)	0.60% (1.55%)	[1]
(11)	0.38% (1.04%)	1.25% (3.28%)	0.62% (1.66%)	[2]

Table 4.2.3 MAPE values for UK & second most deprived quintile

Model	UK	2nd most deprived quintile	Total	Rank
(1) (2) (3)	0.76% (2.19%)	1.06% (2.87%)	0.84% (2.37%)	[6]
(4)	0.76% (2.18%)	1.17% (3.22%)	0.87% (2.47%)	[7]
(5) (6)	0.58% (1.51%)	0.85% (2.43%)	0.65% (1.76%)	[4]
(7)	0.60% (1.66%)	1.09% (2.98%)	0.73% (2.03%)	[5]
(8)	0.86% (2.46%)	1.07% (2.88%)	0.92% (2.57%)	[8]
(9)	0.52% (1.13%)	0.99% (2.72%)	0.65% (1.57%)	[3]
(10)	0.36% (0.96%)	1.06% (2.86%)	0.55% (1.47%)	[1]
(11)	0.38% (1.04%)	1.11% (2.99%)	0.58% (1.58%)	[2]

Table 4.2.4 MAPE values for UK & third most deprived quintile

Model	UK	3rd most deprived quintile	Total	Rank
(1) (2) (3)	0.76% (2.19%)	0.91% (2.63%)	0.80% (2.31%)	[6]
(4)	0.76% (2.18%)	1.07% (3.18%)	0.84% (2.46%)	[7]
(5) (6)	0.58% (1.51%)	0.74% (2.21%)	0.62% (1.70%)	[4]
(7)	0.60% (1.66%)	0.98% (2.89%)	0.70% (2.00%)	[5]
(8)	0.86% (2.46%)	0.92% (2.70%)	0.87% (2.52%)	[8]
(9)	0.52% (1.13%)	0.82% (2.45%)	0.60% (1.49%)	[3]
(10)	0.36% (0.96%)	0.99% (2.95%)	0.53% (1.49%)	[2]
(11)	0.38% (1.04%)	0.90% (2.69%)	0.52% (1.49%)	[1]

Table 4.2.5 MAPE values for UK & fourth most deprived quintile

Model	UK	4th most deprived quintile	Total	Rank
(1) (2) (3)	0.76% (2.19%)	0.99% (3.03%)	0.82% (2.42%)	[6]
(4)	0.76% (2.18%)	1.17% (3.58%)	0.87% (2.57%)	[7]
(5) (6)	0.58% (1.51%)	0.73% (2.34%)	0.62% (1.74%)	[3]
(7)	0.60% (1.66%)	1.04% (3.18%)	0.72% (2.08%)	[5]
(8)	0.86% (2.46%)	0.97% (2.96%)	0.89% (2.60%)	[8]
(9)	0.52% (1.13%)	0.89% (2.64%)	0.62% (1.55%)	[4]
(10)	0.36% (0.96%)	1.01% (3.14%)	0.54% (1.55%)	[2]
(11)	0.38% (1.04%)	0.95% (2.96%)	0.54% (1.57%)	[1]

Table 4.2.6 MAPE values for UK & least deprived quintile

Model	UK	Least deprived quintile	Total	Rank
(1) (2) (3)	0.76% (2.19%)	0.93% (2.85%)	0.81% (2.37%)	[6]
(4)	0.76% (2.18%)	1.12% (3.54%)	0.86% (2.56%)	[7]
(5) (6)	0.58% (1.51%)	0.80% (2.58%)	0.64% (1.81%)	[4]
(7)	0.60% (1.66%)	0.98% (3.07%)	0.70% (2.05%)	[5]
(8)	0.86% (2.46%)	0.99% (3.09%)	0.90% (2.63%)	[8]
(9)	0.52% (1.13%)	0.92% (2.88%)	0.63% (1.61%)	[3]
(10)	0.36% (0.96%)	1.09% (3.46%)	0.56% (1.64%)	[2]
(11)	0.38% (1.04%)	0.97% (3.10%)	0.54% (1.61%)	[1]

4.2.3 Residual Plots

The last goodness-of-fit test used to compare the models in this study is the examination of the standardised deviance residuals (the residual calculations are detailed in the Appendix). Figures 4.2.1 to 4.2.8 demonstrate the standardised deviance residuals for the UK population and deprivation subgroups in England. There are no clear systematic patterns in the residual plots against age and calendar year. However, there is a clear systematic pattern in the residual plots against the cohort year for the UK population under those models without the cohort effect. This finding is consistent with the residual cohort effect found in the English and Wales population by Villegas and Haberman (2014), suggesting the importance of allowing for a cohort parameter in the model structure. The spike in 1920 was caused by the influenza pandemic during 1918-1919. Additionally, Willets (2004) has shown that the UK cohorts between 1925 and 1945 have experienced a rapid improvement in mortality. There are no systematic patterns in the residual plots against the cohort year for the deprivation subgroups as the data period is shorter and there are not many cohorts available before year 1945.

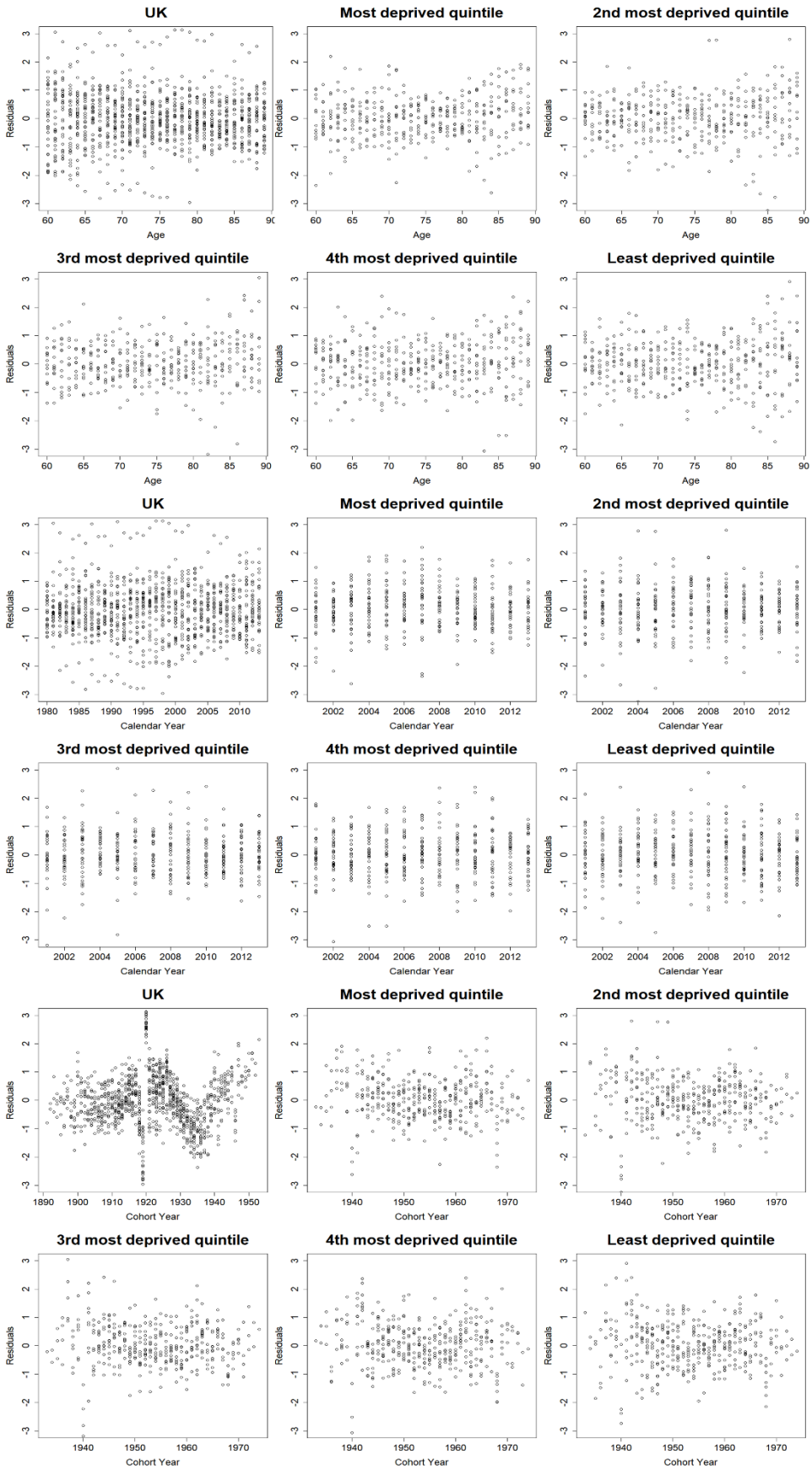


Figure 4.2.1: Standardised deviance residuals for the Lee-Carter model, UK and the deprivation subgroups. The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

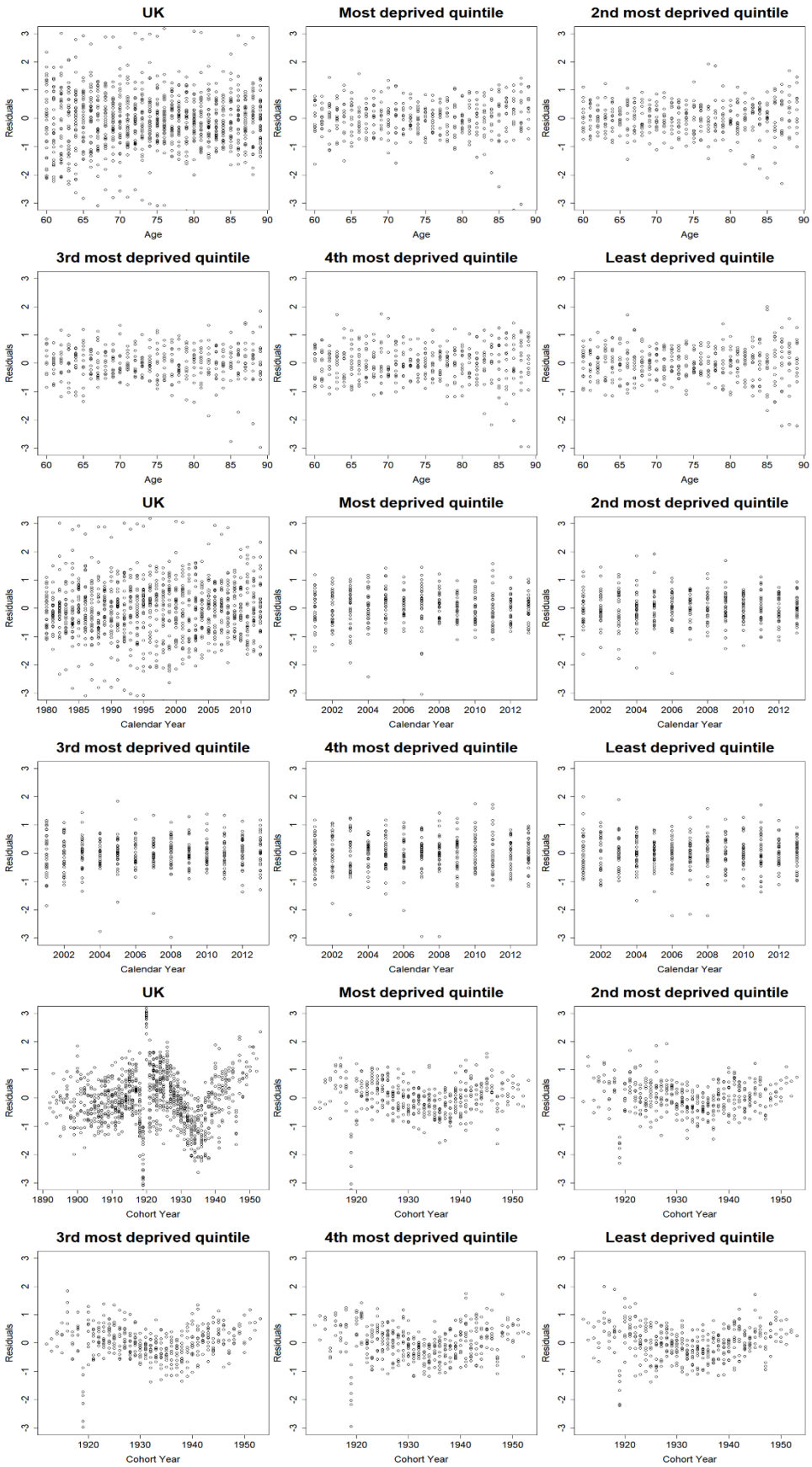


Figure 4.2.2: Standardised deviance residuals for the Lee-Carter (common sensitivity) model, UK and the deprivation subgroups. The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

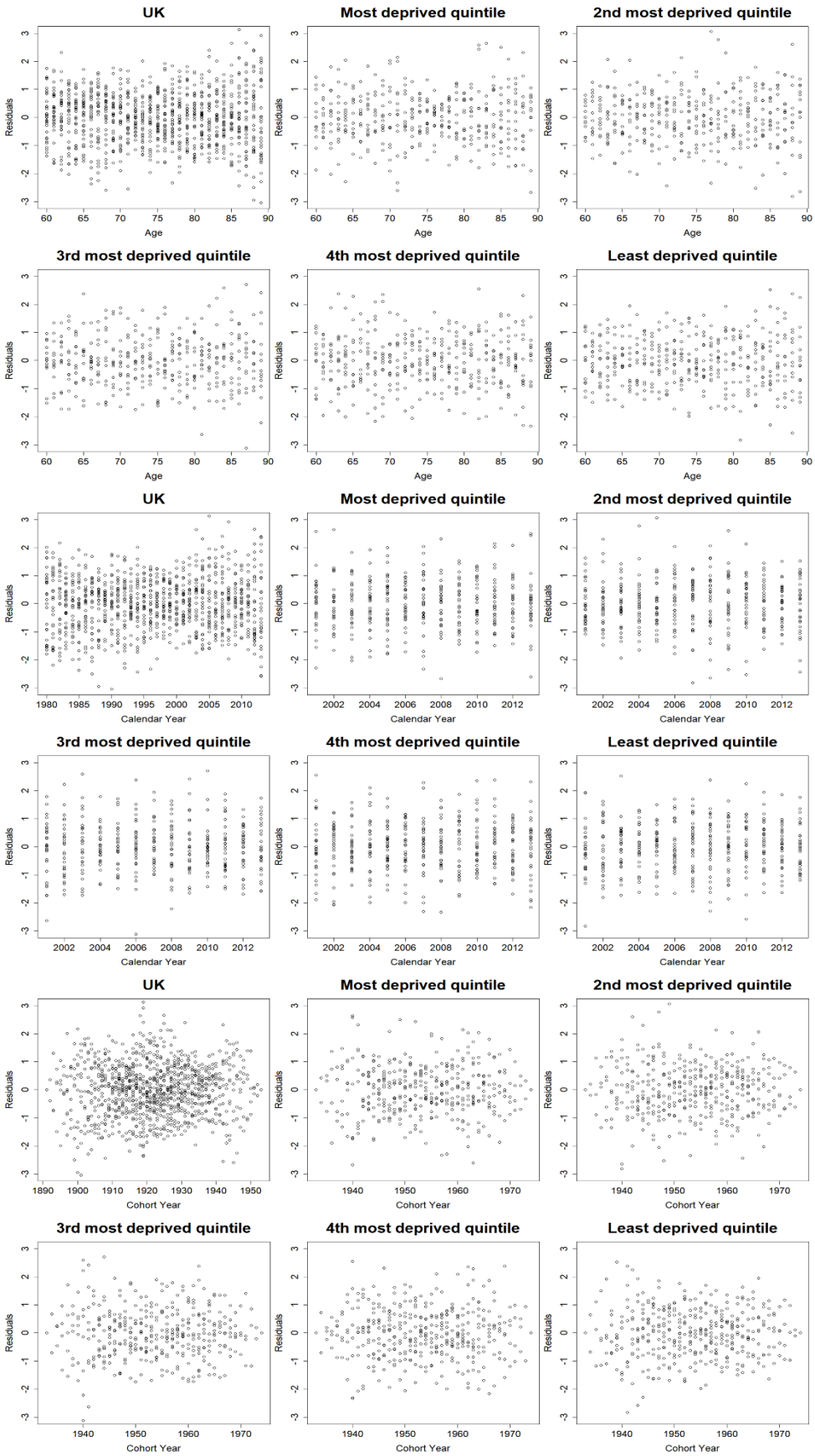


Figure 4.2.3: Standardised deviance residuals for the age-period-cohort model, UK and the deprivation subgroups. The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

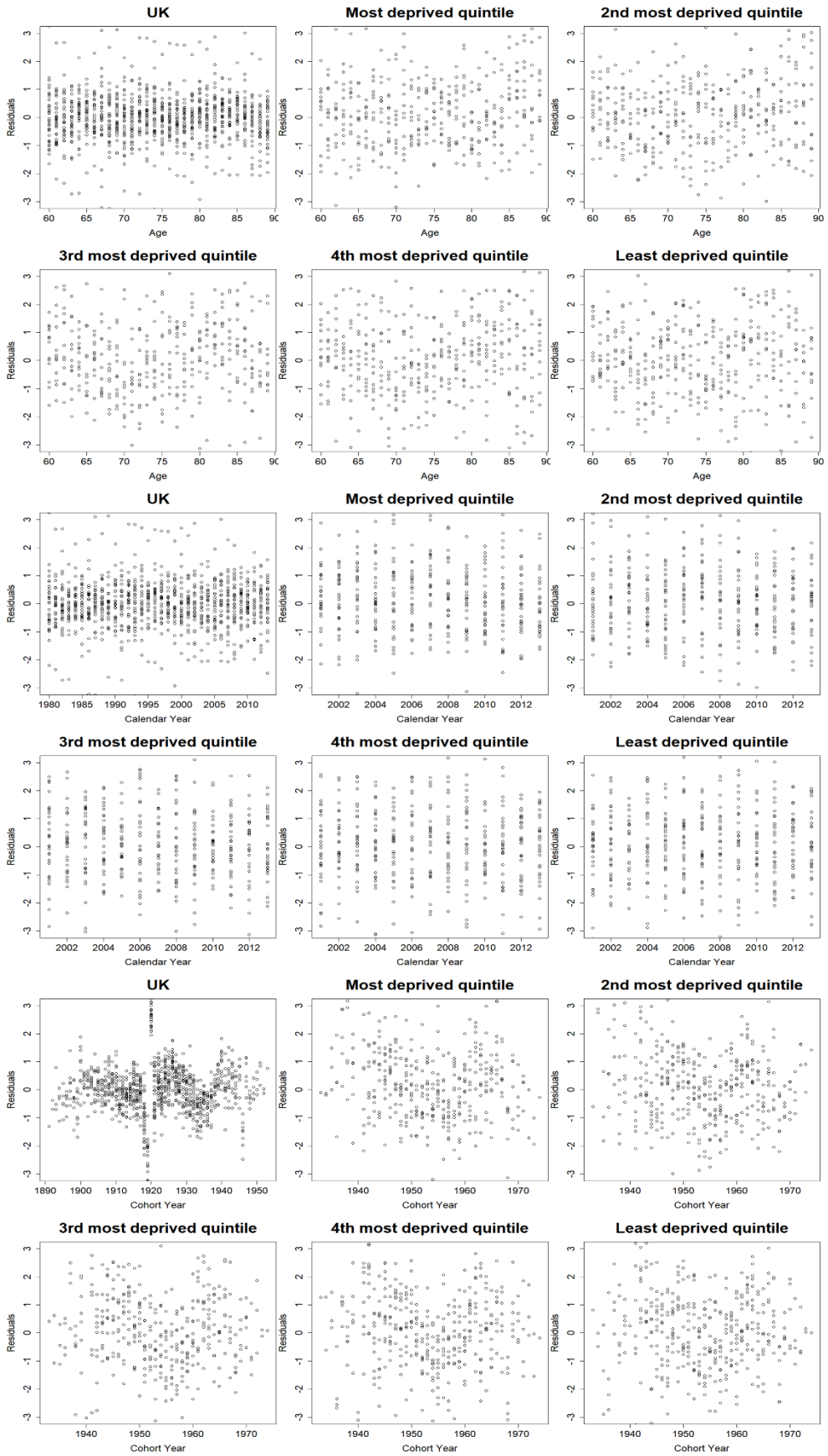


Figure 4.2.4: Standardised deviance residuals for the CBD model, UK and the deprivation subgroups. The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

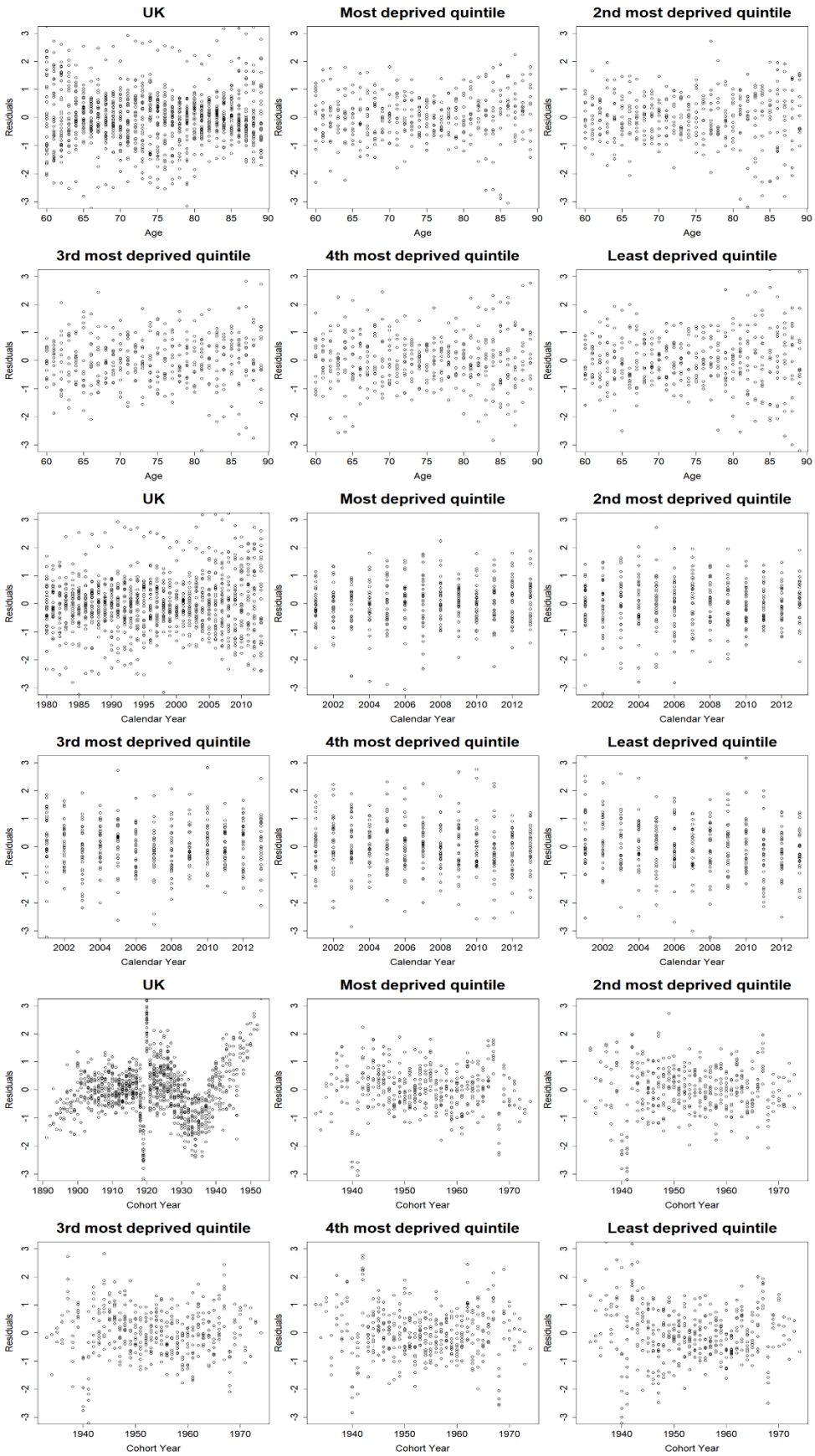


Figure 4.2.5: Standardised deviance residuals for the CBD model (reference) and age and period factor model (book). The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

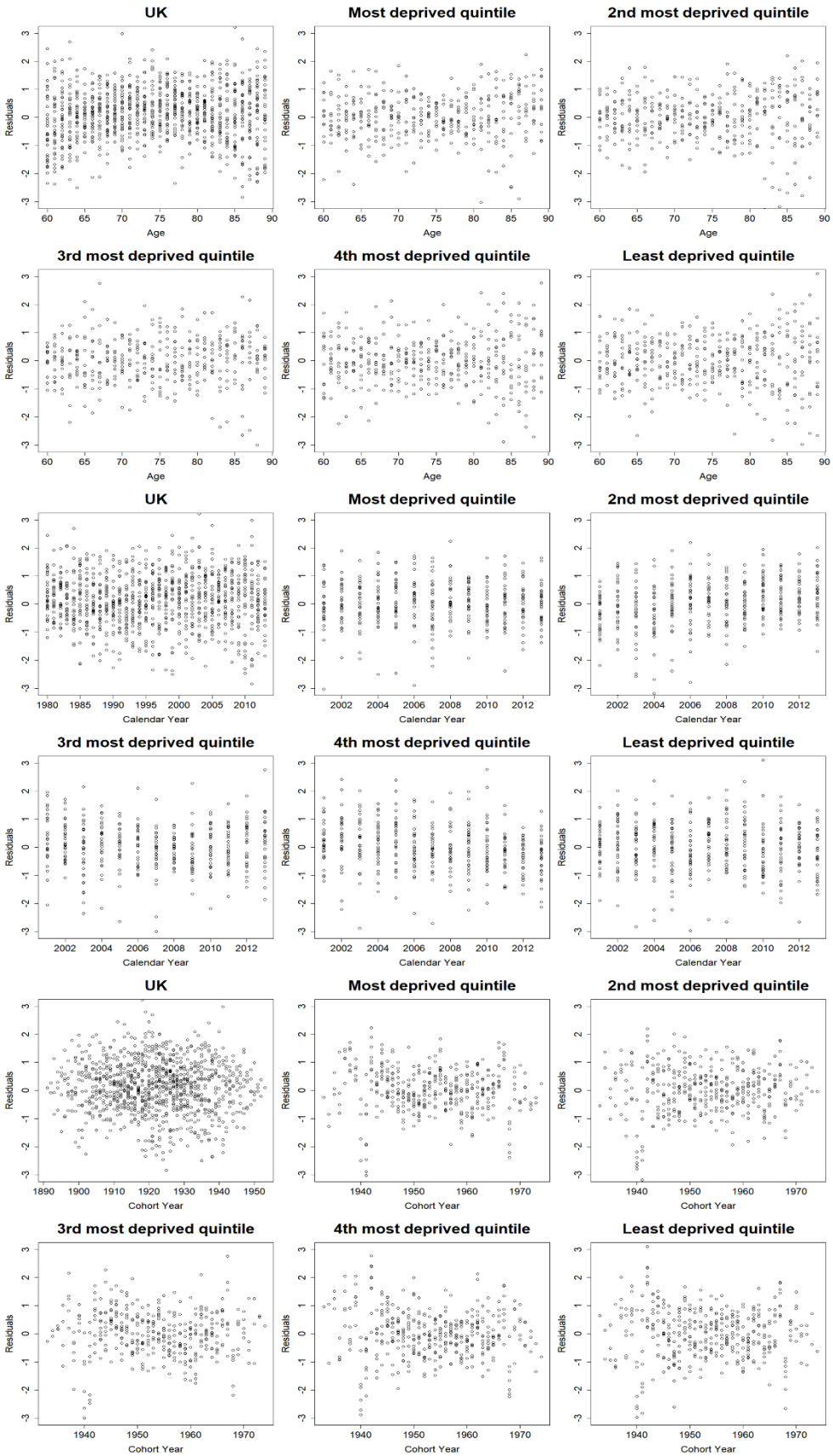


Figure 4.2.6: Standardised deviance residuals for the Lee-Carter model with cohort (reference) age and period factor model (book). The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

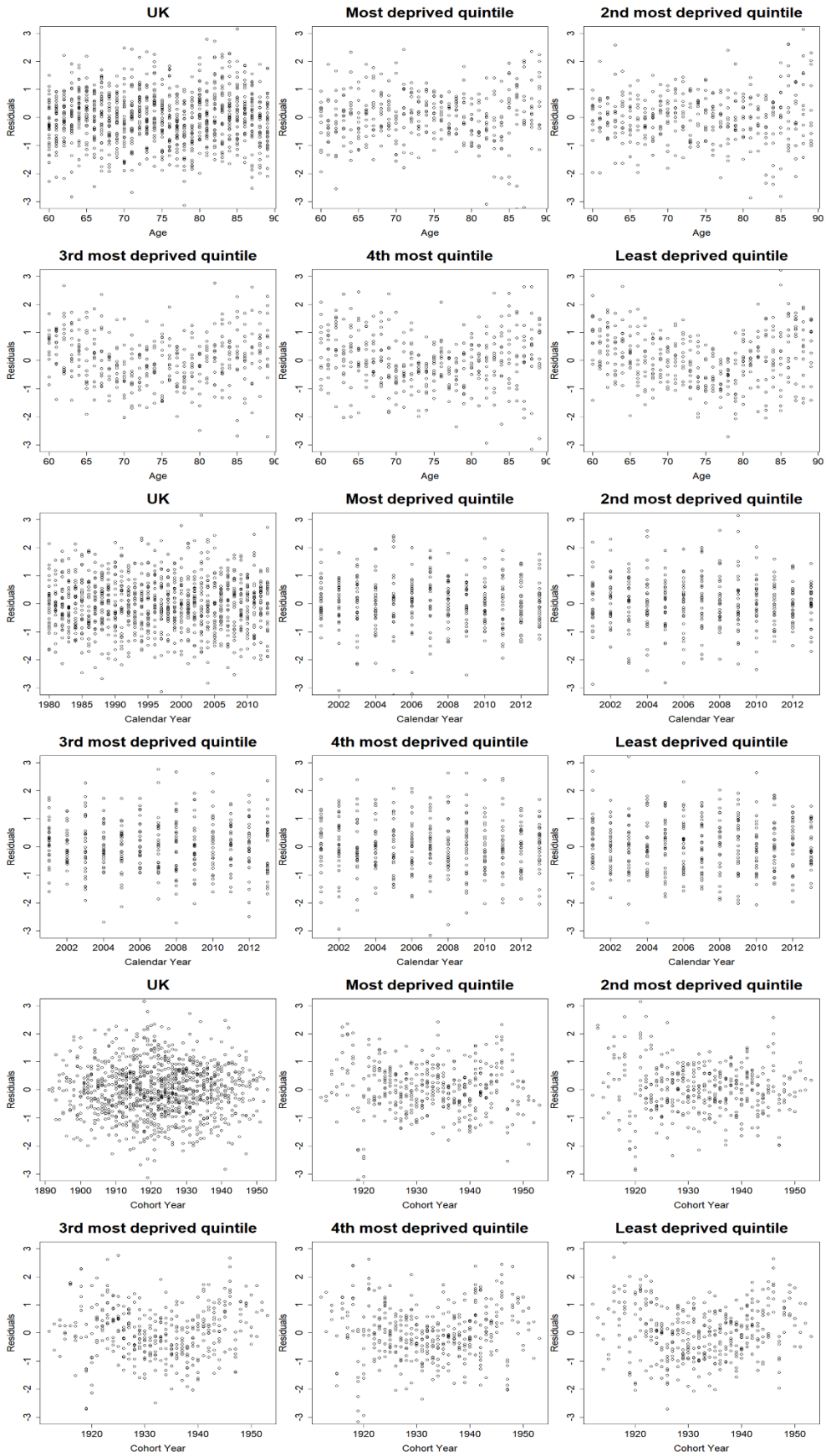


Figure 4.2.7: Standardised deviance residuals for the CBD M7 model (reference) and CBD M5 model (book). The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

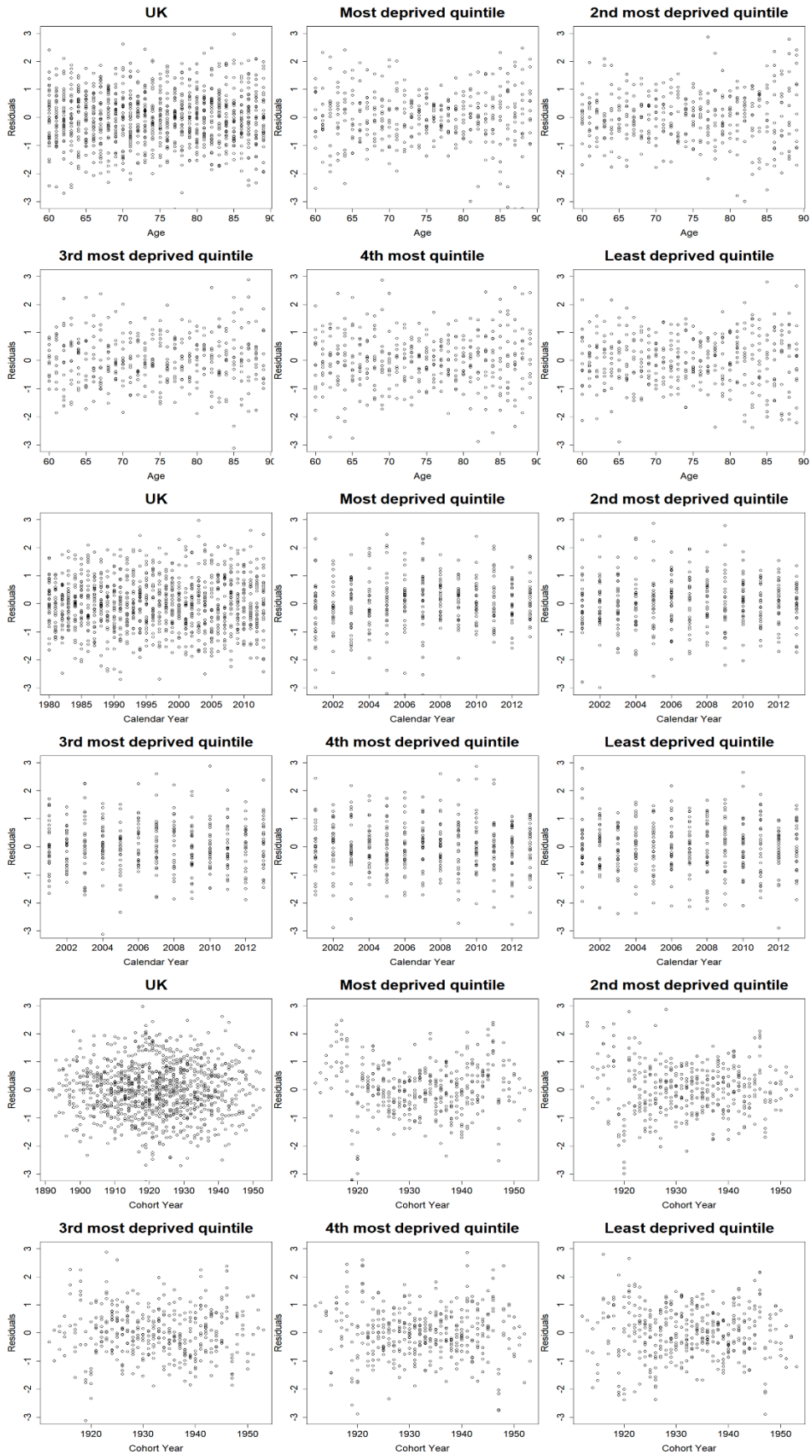


Figure 4.2.8: Standardised deviance residuals for the CAE + cohort model (reference) and CAE model (book). The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

Chapter 5

Model Projection

5.1 Forecasting future mortality rates

It is important to determine the projection period prior to conducting any prediction of future mortality rates. There is a usual convention that the maximum length of the projection period should be aligned to the fitting period (Booth et al., 2002). The fitting period of the UK population is 34 years from 1980 to 2013, and it is 13 years for the deprivation subgroups from 2001 to 2013. According to this convention, it would be plausible to use 13 years as the projection period for the deprivation subgroups. However, such a short projection period would limit further analysis from the mortality projections such as predicting long term survival probabilities for insurance applications and pension valuations. Since the UK population dataset has a larger spectrum of data spanning over 34 years, as well as having larger exposures compared to the deprivation quintiles, the projection period is chosen to be 27 years from 2014 to 2040.

The fitting performance in Chapter 4 indicates that the mortality improvements of the most deprived quintile differ from the rest, while the other quintiles exhibit similar trends among themselves. Chapter 3 also notes that the mortality rates of the UK population hover around those of the three middle deprived quintiles. Accordingly, the remainder of this study will focus on the UK population, as well as the most and least deprived quintile in England[3].

[3] The results for the second, third, and fourth most deprived quintiles are available upon request.

The time varying factors of the 11 models are projected from 2014 to 2040 under the forecasting method specified in Chapter 3. As the fitting processes differ for each model[4], the time varying factors and the cohort factors cannot be used directly to compare the forecasting performances of the models. The future mortality death rates are then used for the comparison as they can be projected through the specifications in Chapter 3.

Figures 5.1.1 to 5.1.11 show the 95% prediction intervals of the simulated 1-year death rate $q_{65,t}$ and $q_{75,t}$ over time under each model. The prediction intervals of $q_{85,t}$ are only displayed in the Appendix since they have similar patterns to q_{75}. The width of the prediction intervals of those models with the cohort effect (i.e models (5), (6), (9), (10), (11)) can be broadly separated into two periods. The period during which the future death rates require projected cohort values will have wider prediction intervals than otherwise. For example, from year 2019 onwards, the prediction intervals of $q_{65,t}$ are much wider than in previous years (2014-2018). Thus, the prediction intervals of these models are generally wider when compared to those without the cohort effect.

The prediction intervals of the deprivation subgroups under models (8) to (11) are generally wider than those under models (1) to (7), since under the former the projection of the death rates for the deprivation subgroups follows only after the projection of the death rates for the UK population. The prediction intervals under model (11) are much wider compared to the other models. However, it does not imply that the model itself is inadequate, as the prediction intervals are heavily dependent on the choice of time series model. A further discussion of this issue can be found in Chapter 6.

It is difficult to determine which model is better simply from judging the width of the prediction intervals, as they incorporate different levels of uncertainty, which can be considered as a part of model risk. It is also difficult to compare the accuracy of the central estimates directly between the models since there is no available data. In the next section, an out-of-sample analysis is conducted as a quantitative measure to compare the forecasting performances of the models.

[4] Models (1) to (3) have the same fitting procedure for the death rates but not for the time varying factors. The situation is the same for models (5) and (6).

Model (1)

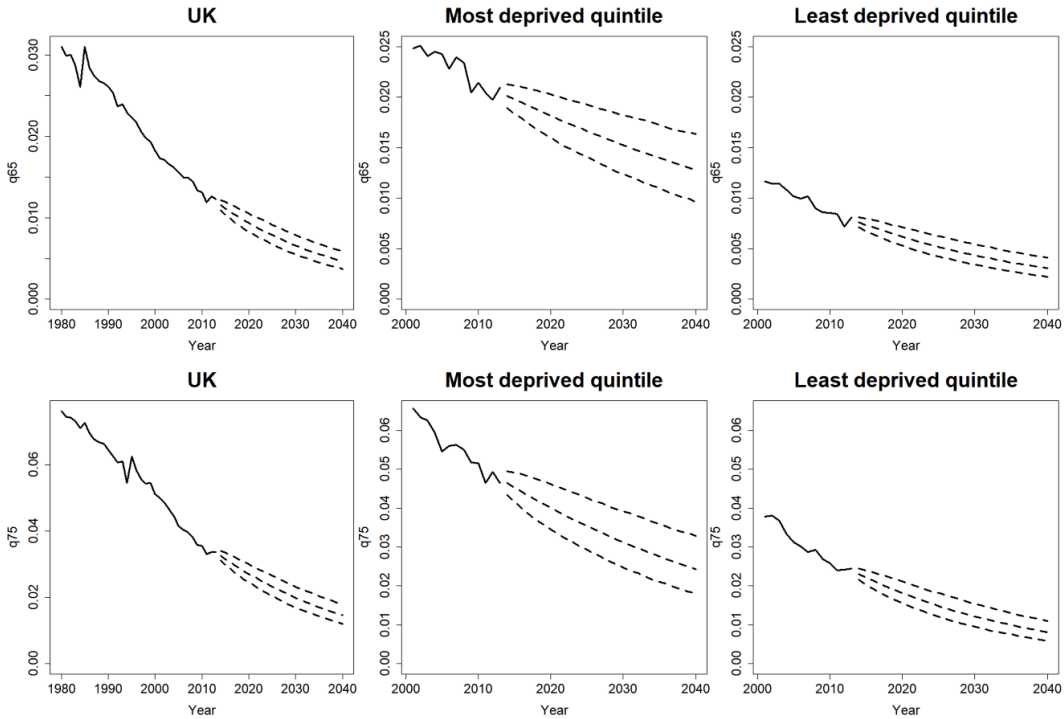


Figure 5.1.1: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (1).

Model (2)

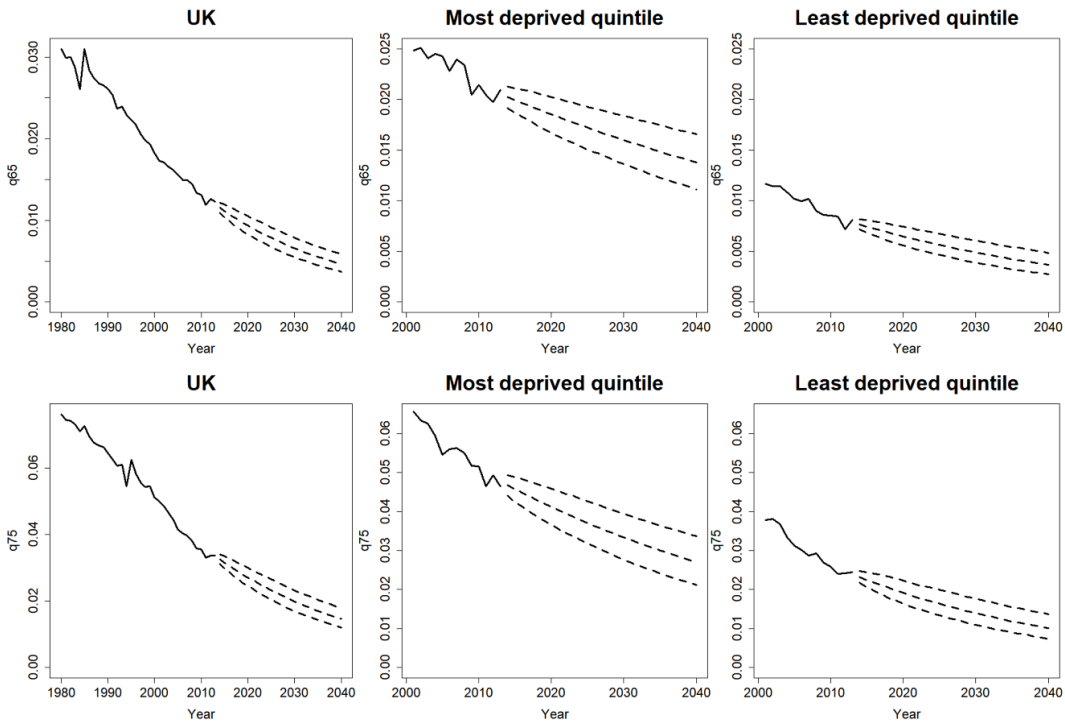


Figure 5.1.2: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (2).

Model (3)

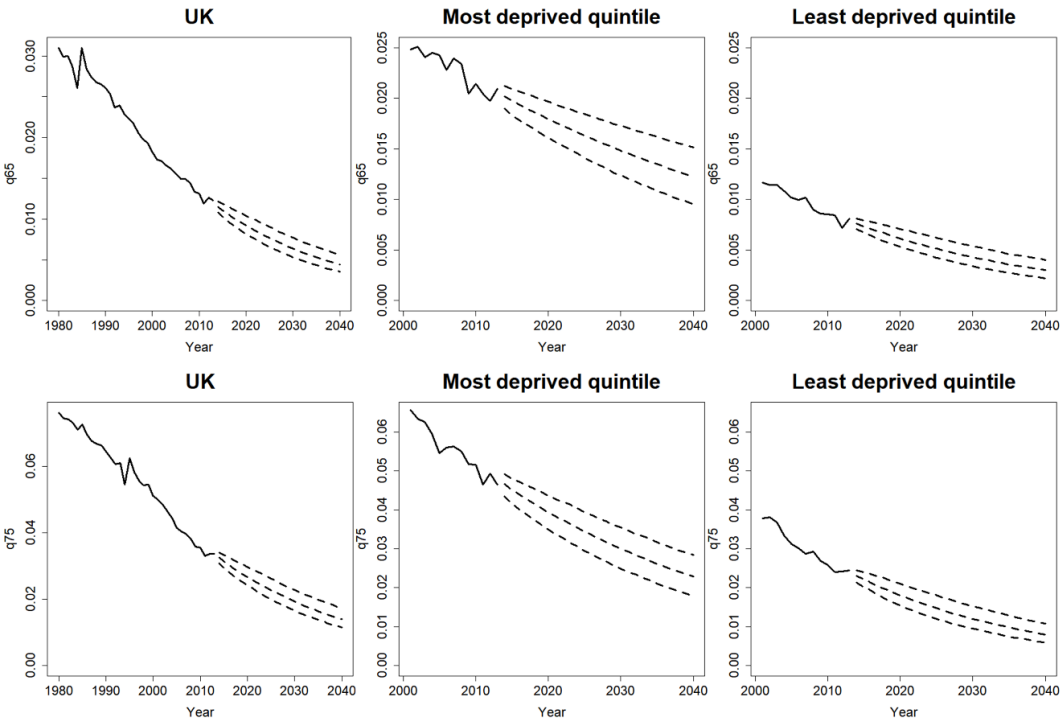


Figure 5.1.3: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (3).

Model (4)

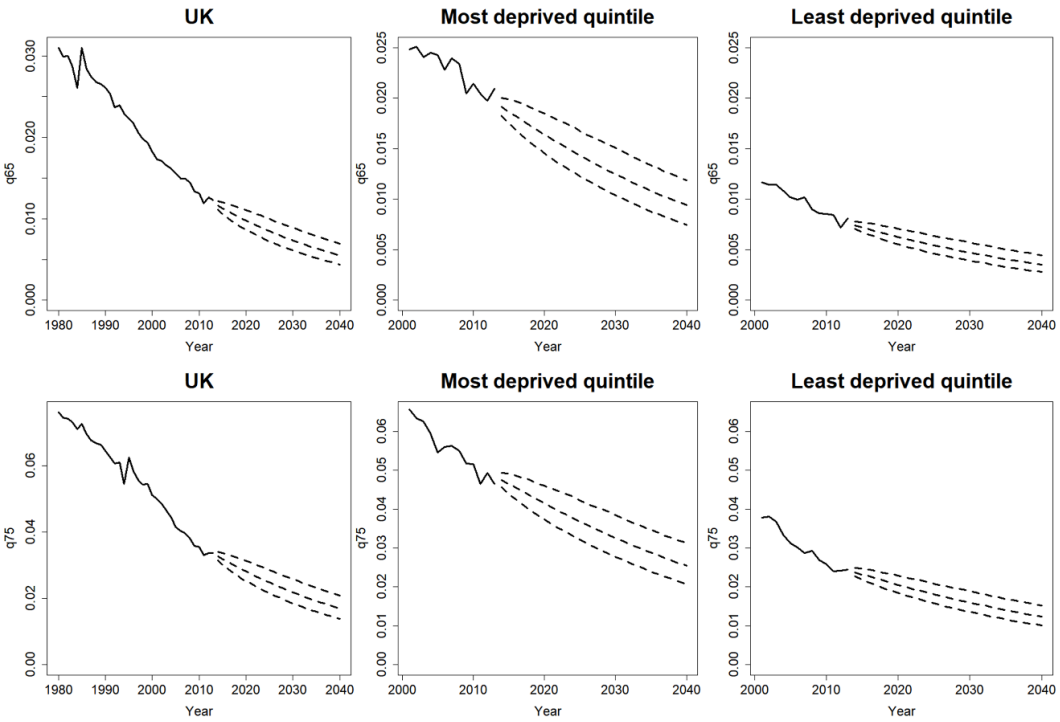


Figure 5.1.4: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (4).

Model (5)

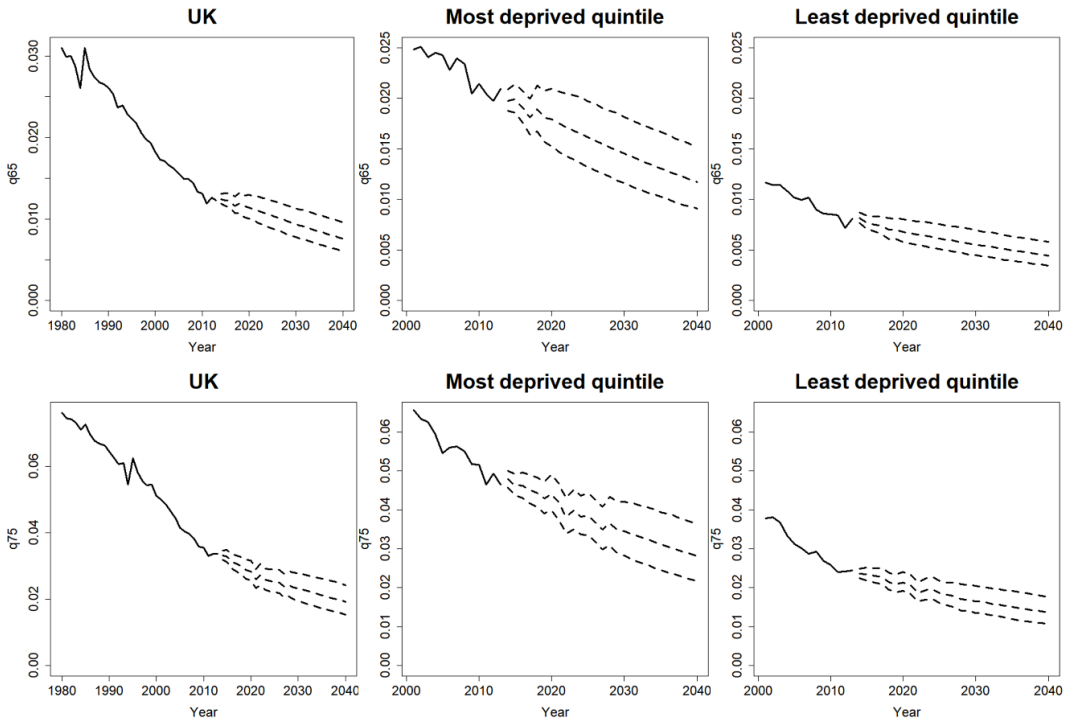


Figure 5.1.5: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (5).

Model (6)

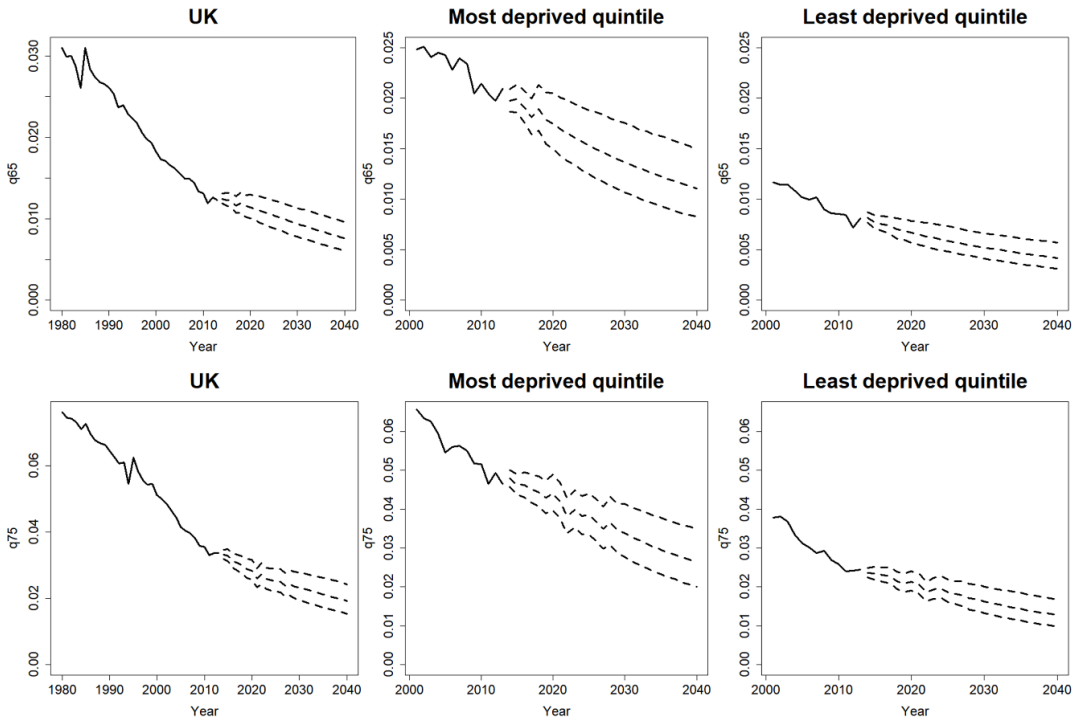


Figure 5.1.6: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (6).

Model (7)

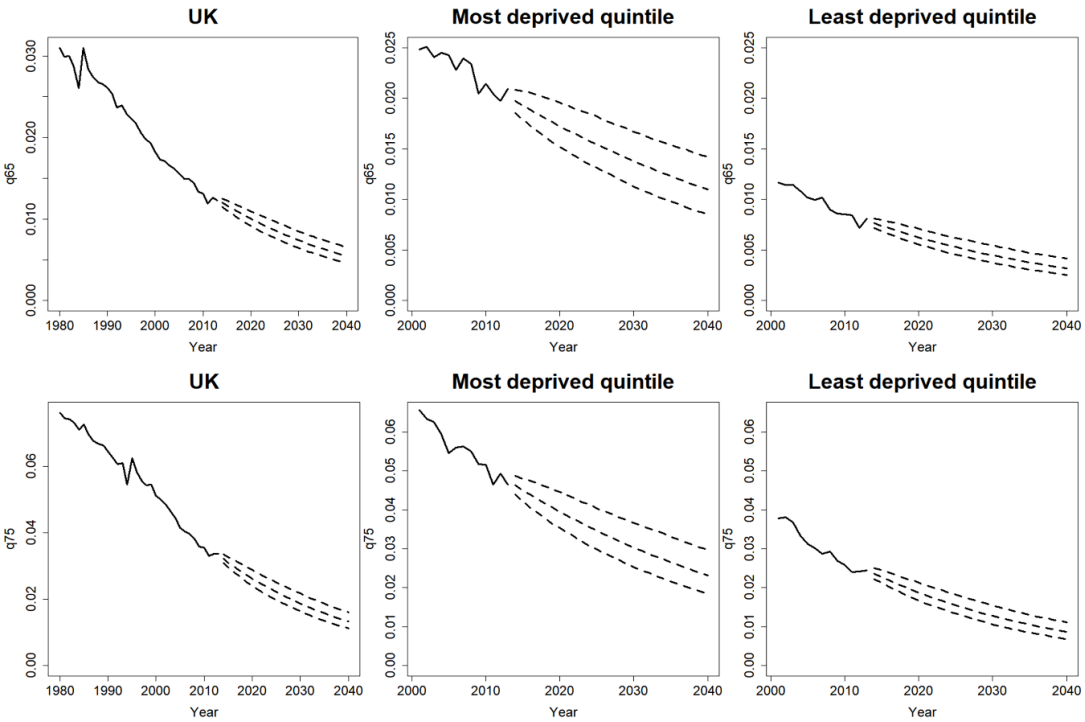


Figure 5.1.7: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (7).

Model (8)

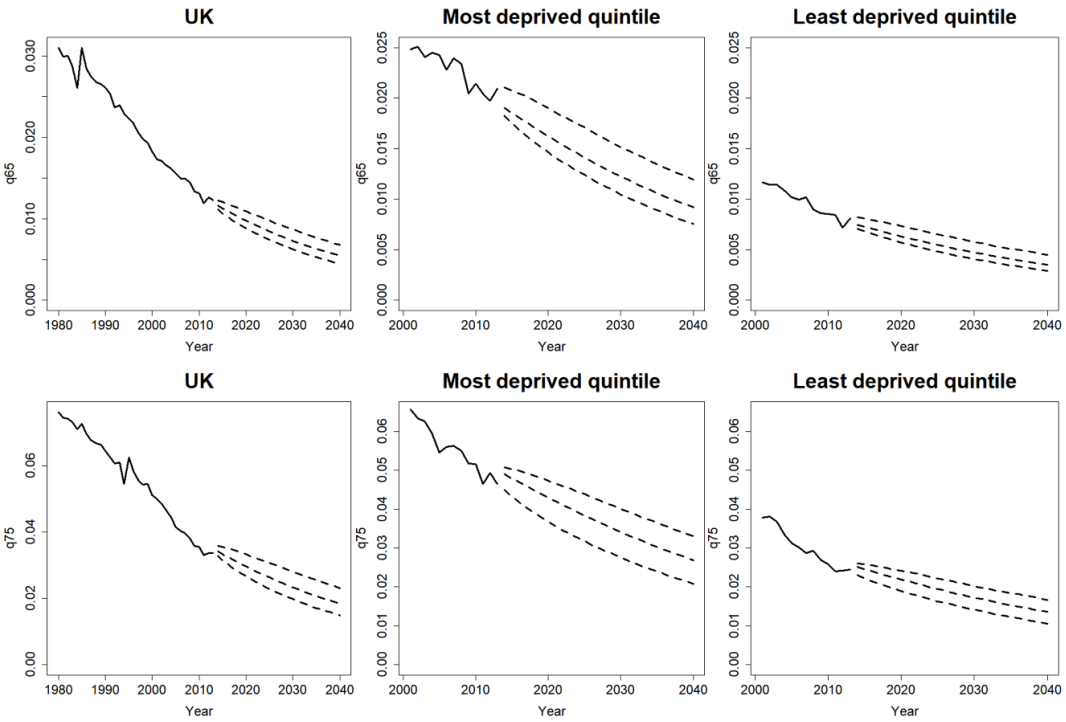


Figure 5.1.8: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (8).

Model (9)

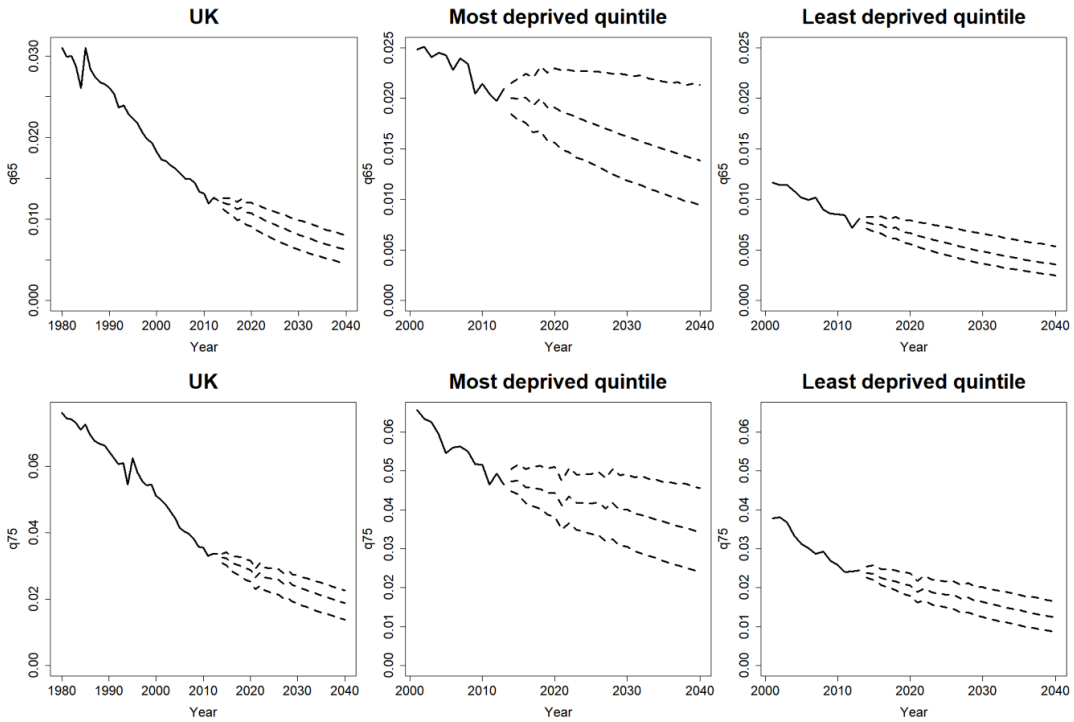


Figure 5.1.9: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (9).

Model (10)

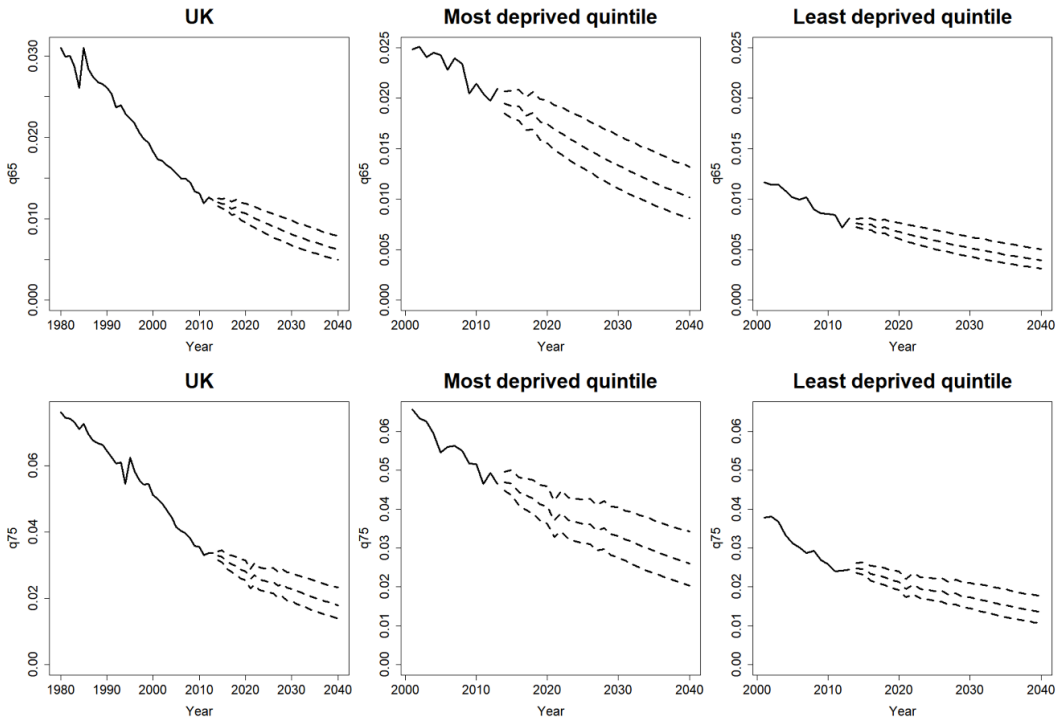


Figure 5.1.10: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (10).

Model (11)

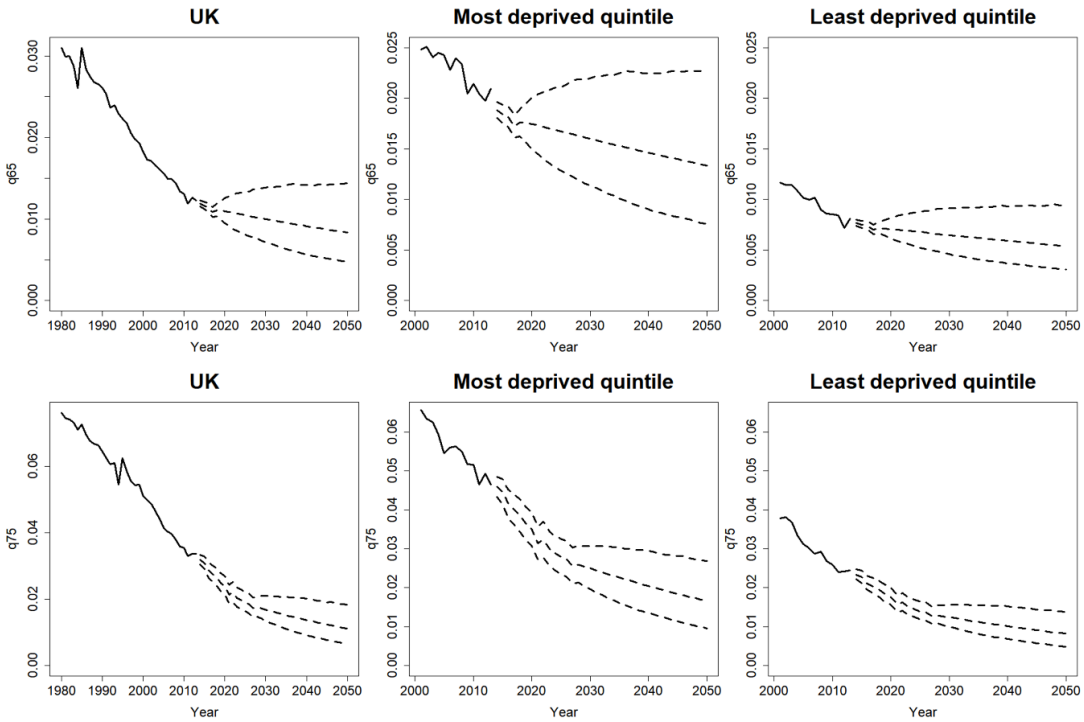


Figure 5.1.11: Projected 95% prediction intervals of $q_{65,t}$ and $q_{75,t}$ (dotted line) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England (right) under model (11).

5.2 Out-of-sample analysis

The out-of-sample test, otherwise known as backtesting, is conducted for the 11 models in the study for the comparison of forecast performances. The sample period is divided into two parts: in-sample and out-of-sample. The models are fitted to the in-sample period and then projected for the out-of-sample period. Afterwards, the MAPE (mean absolute percentage error) is calculated again similarly as in Chapter 4, where the fitted log central death rate is replaced by the projected log central death rate. To be consistent with Chapter 4, the comparison of the central death rates is also given in brackets in Tables 5.2.1 to 5.2.4.

For the UK population, 1980 to 2008 is taken as the in-sample period, and for the deprivation subgroups in England, 2001 to 2008 is set as the in-sample period. Then the out-of-sample period is 2009 to 2013. Tables 5.2.1 to 5.2.2 provide the MAPE values for the most deprived quintile and the least deprived quintile[5]. The age-period-cohort models (models (5) and (6)) produce the lowest MAPE values. The absence of model (7) is explained in the next paragraph.

The prediction intervals are displayed in Figures 5.2.1 to 5.2.10 to examine whether the actual death rates of $q_{75,t}$ lie within the corresponding prediction intervals under different models. For most of the models, the actual death rates are within the 95% prediction intervals, which indicate adequate forecasting performances from all models. Note that the results from model (7) are not provided in this section as the estimated covariance matrix for the time varying factors is possibly unstable[6] due to the limited in-sample period for the residuals bootstrapping, so the MAPE value for model (7) is not shown.

In addition, an interval forecast accuracy test has been conducted as a quantitative measure to the proportion of the actual death rates for all ages that lies within the corresponding prediction intervals under different models. The figures are shown in Table 5.2.3 to 5.3.4. The tables show all models provide sufficient forecasting performances.

Another out-of-sample analysis is conducted using another split for the sample period, in which the UK population has 1980 to 2010 as the in-sample period and the deprivation sub groups in England

[5] The other quintiles are shown in the Appendix.

[6] The covariance matrix is unstable when it is not positive definite, therefore it is unable to compute the Cholesky decomposition which is needed to simulate the vector error terms.

have 2001 to 2010 as the in-sample period. The out-of-sample period is then 2011 to 2013. Tables 5.2.5 to 5.2.6 provide the MAPE values for the most and least deprived quintiles. Most MAPE values drop for the shorter projection period, except for model (11), as a longer in-sample period generally increases the accuracy in projection. Interestingly, the M7M5 model in model (10) produces the lowest MAPE value, which differs from the previous split. The 95% prediction intervals are plotted in Figures 5.2.11 to 5.2.21 and the interval forecast accuracy test have been conducted for this split and the results are shown in Table 5.2.7 to 5.2.8. Again, it can be seen that most of the prediction intervals capture the actual death rates.

There are some inconsistent results in the ranking of the MAPE values and the interval forecast accuracy between the two different splits. One might argue that the in-sample period is too short for the first split. Only 8 years are used as the in-sample period for the deprivation subgroup population to project the death rates for the next 5 years. Another possible reason is that due to the projected log death rates being fairly close between different models, actually all 11 models deliver a satisfactory forecasting performance based on the MAPE analysis.

Table 5.2.1 MAPE values for the projected death rates (2009-2013) for UK & most deprived quintile.

Model	UK	Most deprived quintile	Total	Rank
(1)	1.27% (4.08%)	2.09% (6.22%)	1.68% (5.15%)	[6]
(2)	1.48% (4.48%)	2.07% (6.17%)	1.77% (5.33%)	[7]
(3)	1.23% (4.14%)	2.02% (5.80%)	1.63% (4.97%)	[5]
(4)	1.47% (4.46%)	2.42% (7.04%)	1.94% (5.75%)	[10]
(5)	1.01% (3.22%)	1.62% (4.74%)	1.32% (3.98%)	[1]
(6)	1.00% (3.17%)	1.65% (4.84%)	1.32% (4.01%)	[2]
(7)	-	-	-	-
(8)	1.43% (4.85%)	2.19% (6.41%)	1.81% (5.63%)	[8]
(9)	1.33% (3.90%)	2.38% (7.05%)	1.86% (5.48%)	[9]
(10)	1.12% (3.24%)	2.12% (6.11%)	1.62% (4.67%)	[4]
(11)	1.32% (3.78%)	1.46% (3.95%)	1.39% (3.86%)	[3]

Table 5.2.2 MAPE values for the projected death rates (2009-2013) for UK & least deprived quintile.

Model	UK	Least deprived quintile	Total	
(1)	1.27% (4.08%)	1.63% (4.97%)	1.45% (4.52%)	[4]
(2)	1.48% (4.48%)	1.47% (4.56%)	1.47% (4.52%)	[5]
(3)	1.24% (4.18%)	1.80% (5.46%)	1.52% (4.82%)	[6]
(4)	1.46% (4.43%)	1.60% (5.29%)	1.53% (4.86%)	[7]
(5)	1.01% (3.22%)	1.24% (4.38%)	1.13% (3.80%)	[1]
(6)	1.00% (3.17%)	1.31% (4.70%)	1.15% (3.94%)	[2]
(7)	-	-	-	-
(8)	1.43% (4.85%)	1.89% (6.69%)	1.66% (5.77%)	[9]
(9)	1.33% (3.90%)	2.02% (6.22%)	1.68% (5.06%)	[10]
(10)	1.12% (3.24%)	1.63% (5.58%)	1.37% (4.41%)	[3]
(11)	1.32% (3.78%)	1.61% (5.13%)	1.61% (4.46%)	[8]

Table 5.2.3 Interval forecast accuracy for the projected death rates (2009-2013) for UK & most deprived quintile.

Model	UK	Least deprived quintile	Total	Rank
(1)	72%	83%	78%	[8]
(2)	74%	83%	78%	[7]
(3)	88%	93%	90%	[2]
(4)	74%	77%	75%	[10]
(5)	93%	81%	87%	[4]
(6)	93%	81%	87%	[4]
(7)	-	-	-	-
(8)	78%	91%	84%	[6]
(9)	70%	82%	76%	[9]
(10)	95%	95%	95%	[1]
(11)	89%	89%	89%	[3]

Table 5.2.4 Interval forecast accuracy for the projected death rates (2009-2013) for UK & least deprived quintile.

Model	UK	Least deprived quintile	Total	Rank
(1)	73%	85%	79%	[8]
(2)	74%	89%	82%	[6]
(3)	82%	89%	86%	[5]
(4)	73%	71%	72%	[10]
(5)	93%	83%	88%	[2]
(6)	93%	83%	88%	[2]
(7)	-	-	-	-
(8)	78%	78%	78%	[9]
(9)	70%	91%	81%	[7]
(10)	95%	85%	90%	[1]
(11)	89%	83%	86%	[4]

Projection period 2009-2013

Model (1)

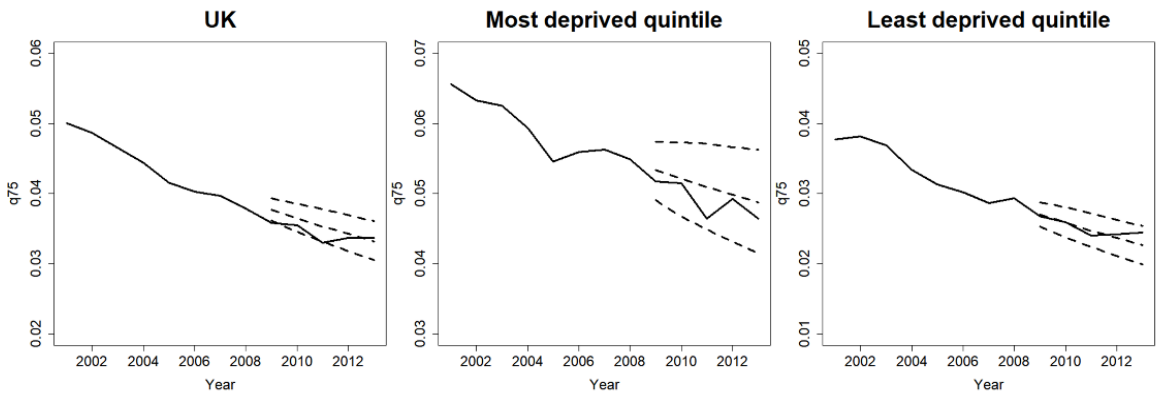


Figure 5.2.1: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (1).

Model (2)

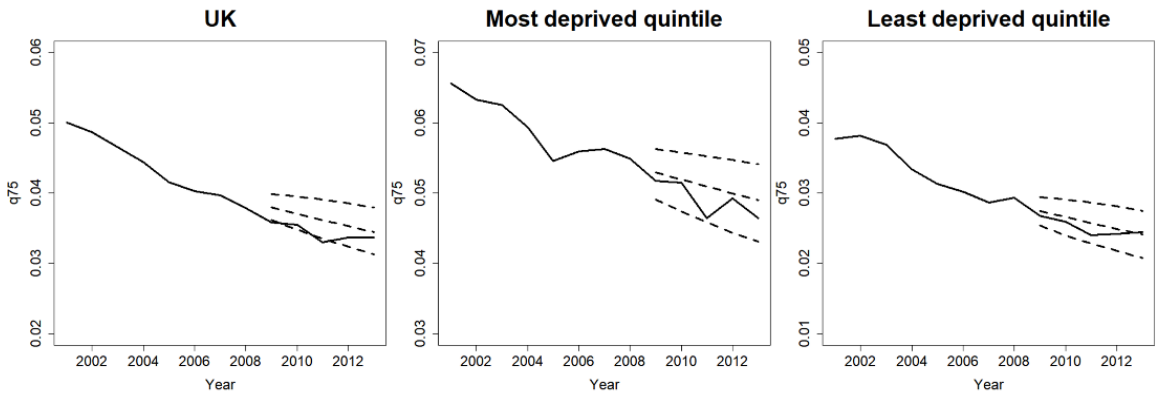


Figure 5.2.2: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (2).

Model (3)

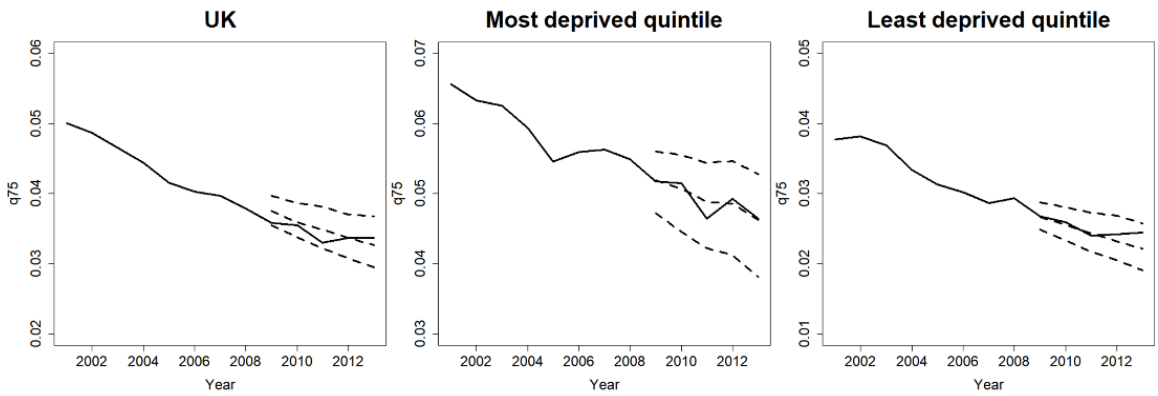


Figure 5.2.3 Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (3).

Model (4)

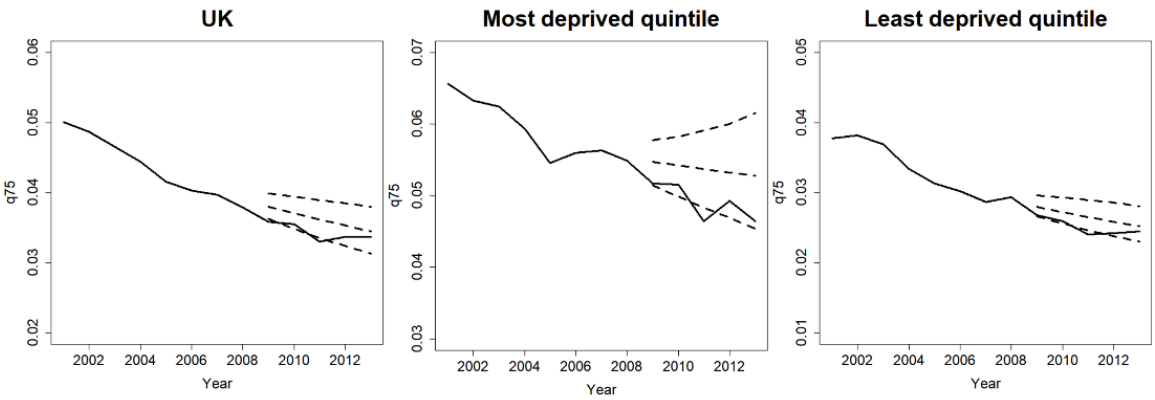


Figure 5.2.4 Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (4).

Model (5)

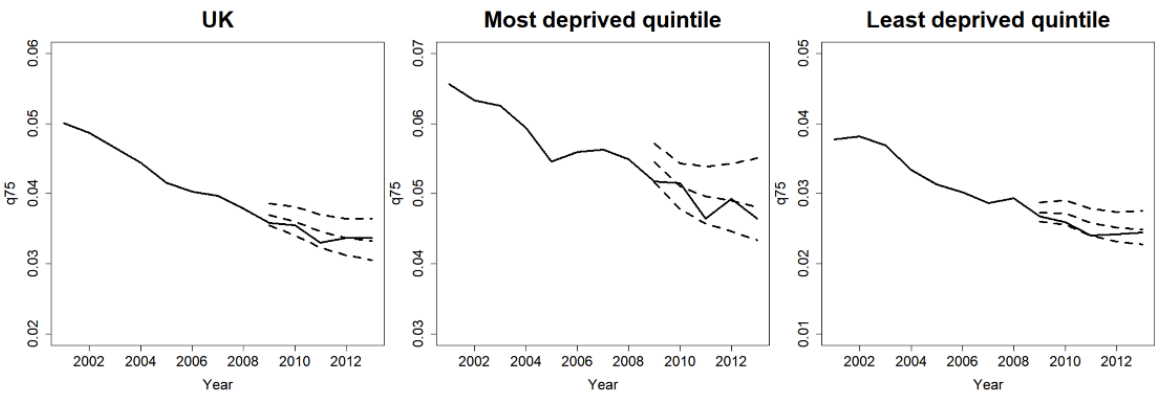


Figure 5.2.5: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (5).

Model (6)

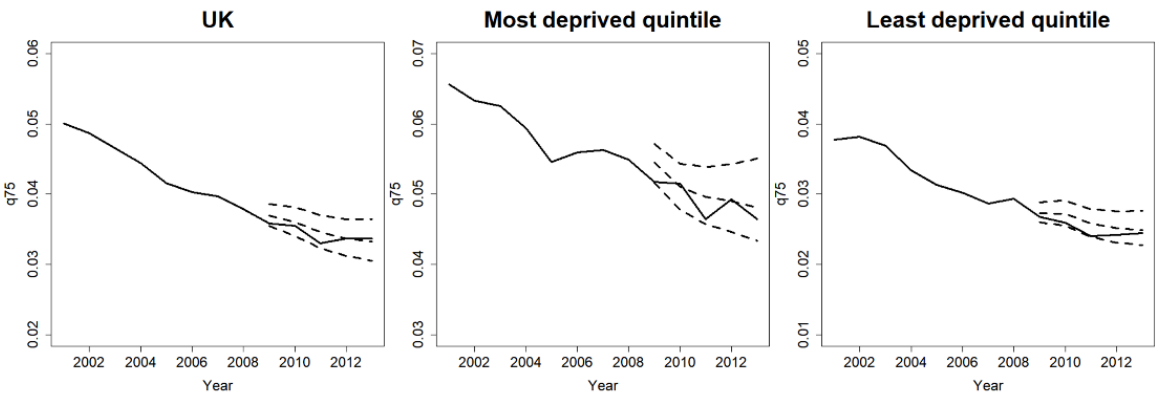


Figure 5.2.6: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (6).

Model (8)

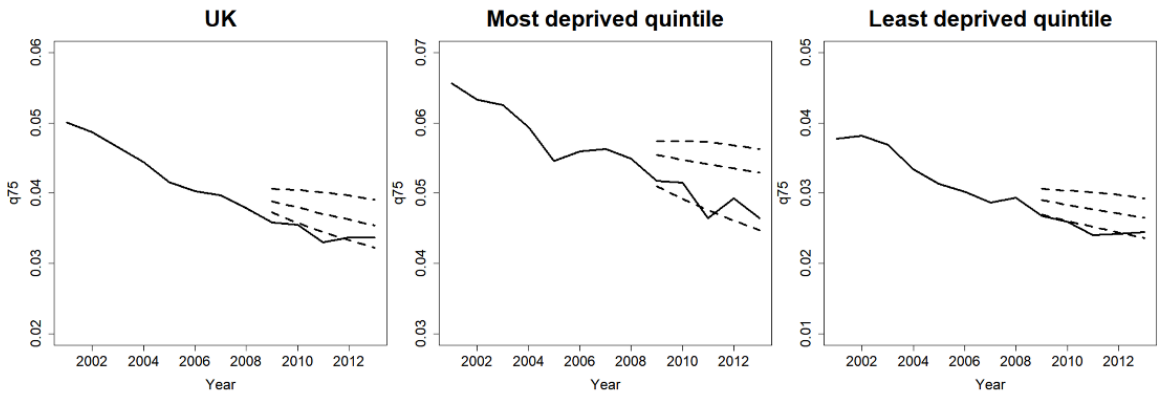


Figure 5.2.7: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (8).

Model (9)

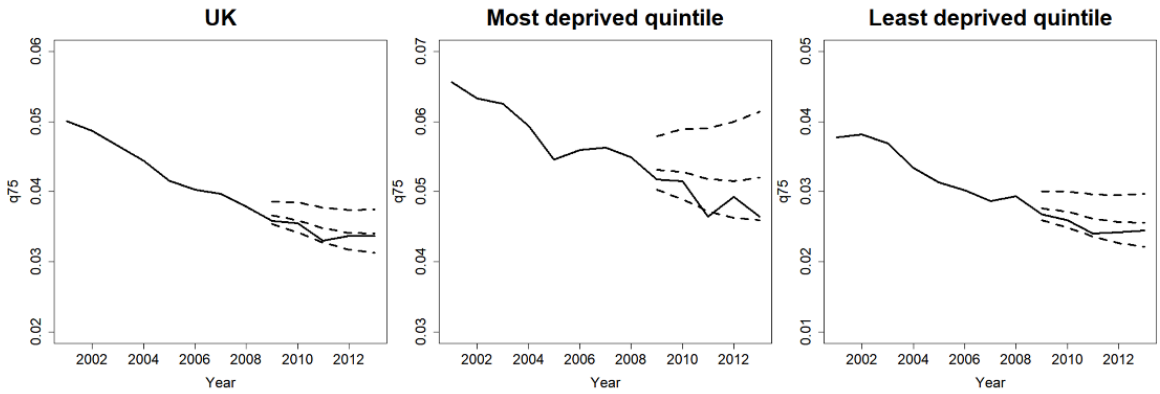


Figure 5.2.8: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (9).

Model (10)

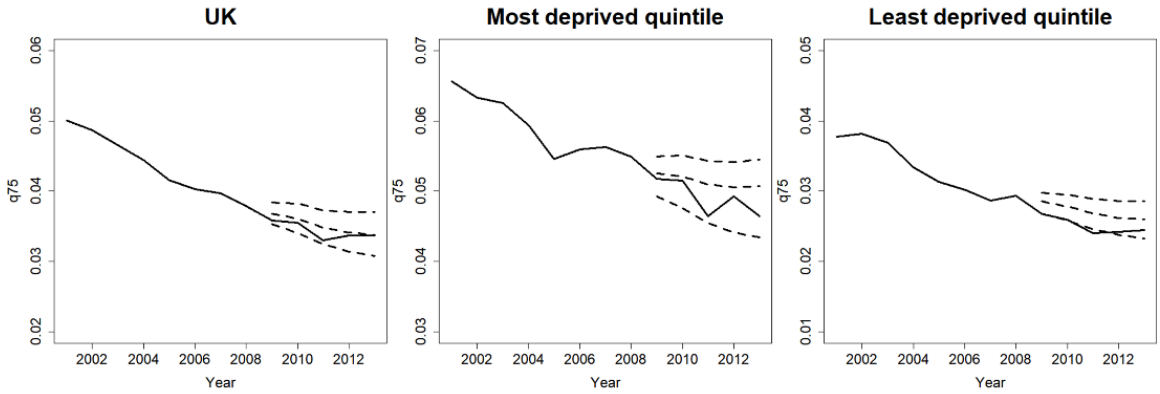


Figure 5.2.9: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (10).

Model (11)

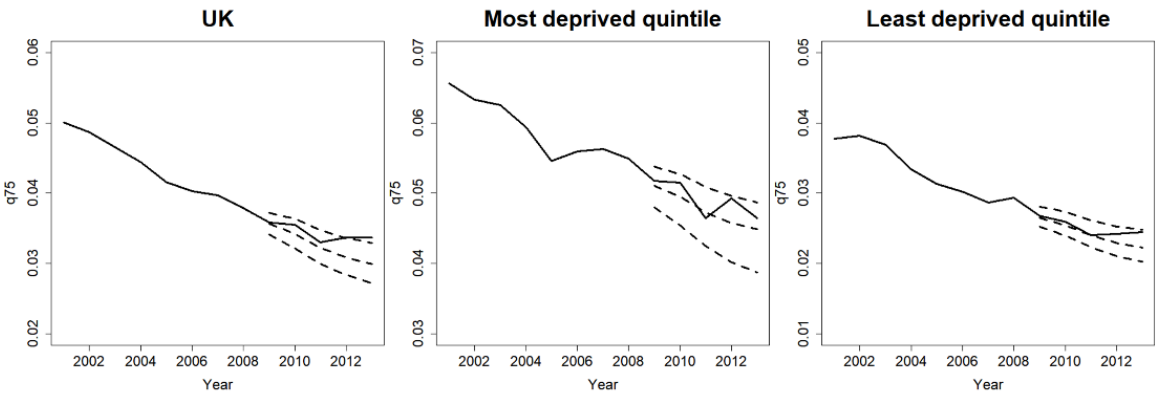


Figure 5.2.10: Simulated prediction intervals and actual $q_{75,t}$ (2009-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (11).

Table 5.2.5 MAPE values for the projected death rates (2011-2013) for UK & most deprived quintile.

Model	UK	Most deprived quintile	Total	RANK
(1)	1.13% (3.92%)	1.51% (4.09%)	1.32% (4.01%)	[8]
(2)	1.02% (3.46%)	1.38% (3.95%)	1.20% (3.70%)	[5]
(3)	1.15% (4.00%)	1.48% (4.12%)	1.32% (4.06%)	[9]
(4)	1.01% (3.43%)	1.38% (3.88%)	1.19% (3.65%)	[4]
(5)	1.14% (3.21%)	1.39% (3.73%)	1.27% (3.47%)	[7]
(6)	1.14% (3.20%)	1.38% (3.68%)	1.26% (3.44%)	[6]
(7)	1.03% (3.09%)	1.32% (3.50%)	1.17% (3.30%)	[3]
(8)	1.41% (4.42%)	1.94% (4.87%)	1.68% (4.64%)	[10]
(9)	0.73% (2.22%)	1.40% (4.05%)	1.07% (3.13%)	[2]
(10)	0.63% (1.96%)	1.33% (3.57%)	0.98% (2.77%)	[1]
(11)	1.65% (4.61%)	2.33% (6.07%)	1.99% (5.34%)	[11]

Table 5.2.6 MAPE values for the projected death rates (2011-2013) for UK & most deprived quintile.

Model	UK	Least deprived quintile	Total	RANK
(1)	1.13% (3.92%)	1.24% (4.33%)	1.18% (4.13%)	[8]
(2)	1.02% (3.46%)	1.14% (4.00%)	1.08% (3.73%)	[3]
(3)	1.09% (3.62%)	1.18% (4.13%)	1.13% (3.88%)	[7]
(4)	1.01% (3.41%)	1.45% (4.91%)	1.23% (4.16%)	[9]
(5)	1.14% (3.21%)	1.11% (4.15%)	1.13% (3.68%)	[6]
(6)	1.14% (3.20%)	1.09% (4.02%)	1.11% (3.61%)	[5]
(7)	1.03% (3.09%)	1.14% (3.90%)	1.09% (3.50%)	[4]
(8)	1.41% (4.42%)	1.56% (5.57%)	1.49% (4.99%)	[10]
(9)	0.73% (2.22%)	1.42% (4.69%)	1.07% (3.46%)	[2]
(10)	0.63% (1.96%)	1.29% (4.57%)	0.96% (3.27%)	[1]
(11)	1.65% (4.61%)	1.65% (5.46%)	1.65% (5.04%)	[11]

Table 5.2.7 Interval forecast accuracy for the projected death rates (2011-2013) for UK & most deprived quintile.

Model	UK	Least deprived quintile	Total	Rank
(1)	69%	89%	79%	[10]
(2)	82%	81%	82%	[8]
(3)	71%	88%	79%	[9]
(4)	82%	83%	83%	[7]
(5)	83%	88%	86%	[2]
(6)	83%	88%	86%	[2]
(7)	78%	91%	84%	[4]
(8)	79%	90%	84%	[4]
(9)	88%	79%	83%	[6]
(10)	98%	93%	96%	[1]
(11)	69%	64%	67%	[11]

Table 5.2.8 Interval forecast accuracy for the projected death rates (2011-2013) for UK & least deprived quintile.

Model	UK	Least deprived quintile	Total	Rank
(1)	69%	78%	73%	[10]
(2)	82%	84%	83%	[3]
(3)	74%	79%	77%	[8]
(4)	82%	70%	76%	[9]
(5)	83%	82%	83%	[4]
(6)	83%	81%	82%	[5]
(7)	79%	80%	79%	[7]
(8)	78%	83%	81%	[6]
(9)	88%	83%	86%	[2]
(10)	98%	79%	88%	[1]
(11)	71%	70%	71%	[11]

Projection period 2011-2013

Model (1)

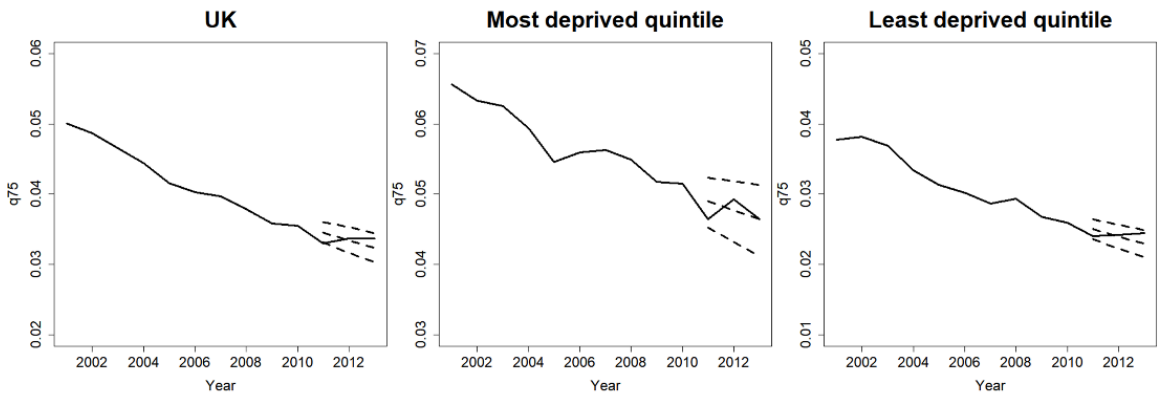


Figure 5.2.11: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (1).

Model (2)

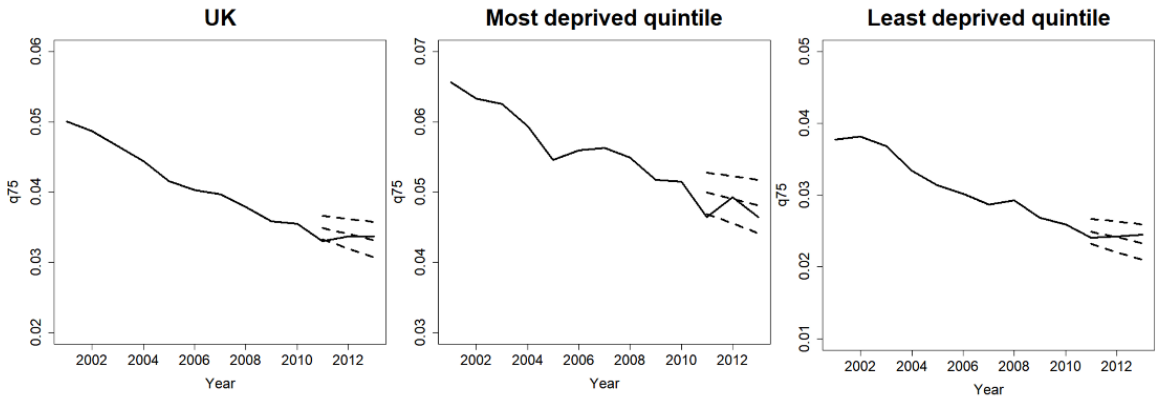


Figure 5.2.12: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (2).

Model (3)

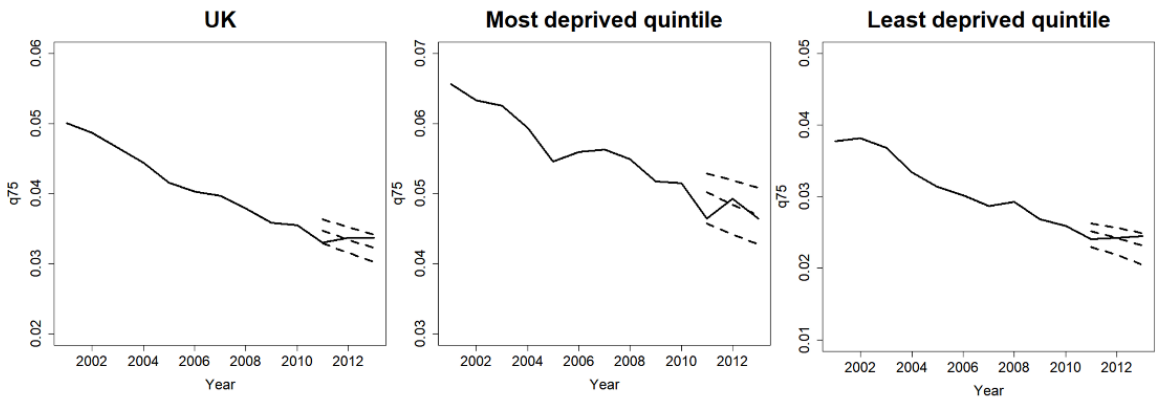


Figure 5.2.13: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (3).

Model (4)

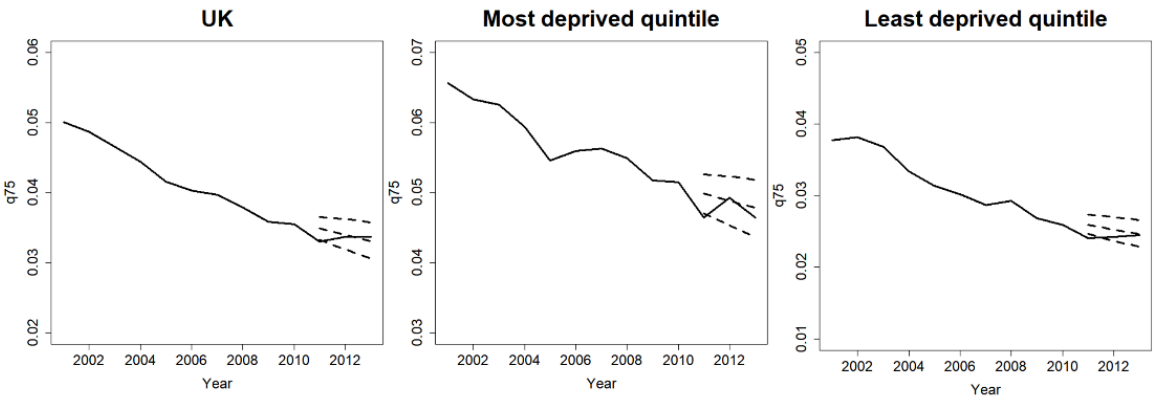


Figure 5.2.14: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (4).

Model (5)

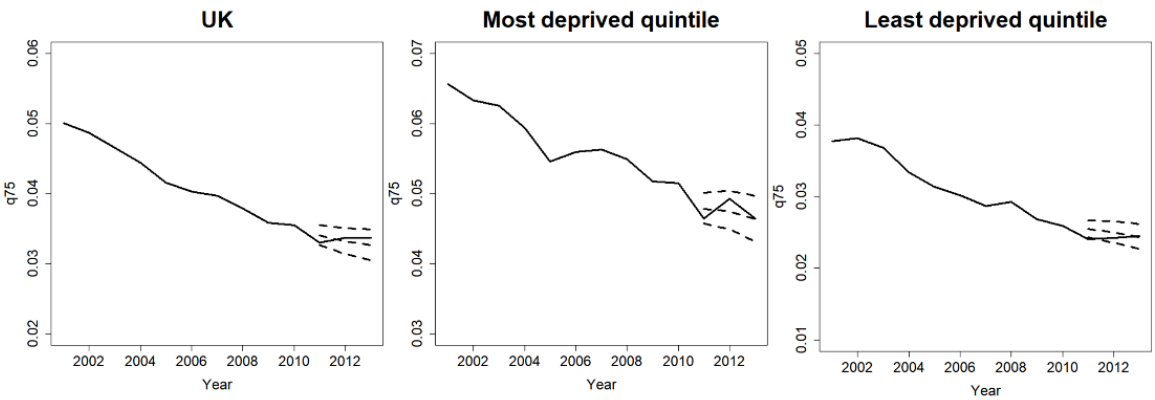


Figure 5.2.15: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (5).

Model (6)

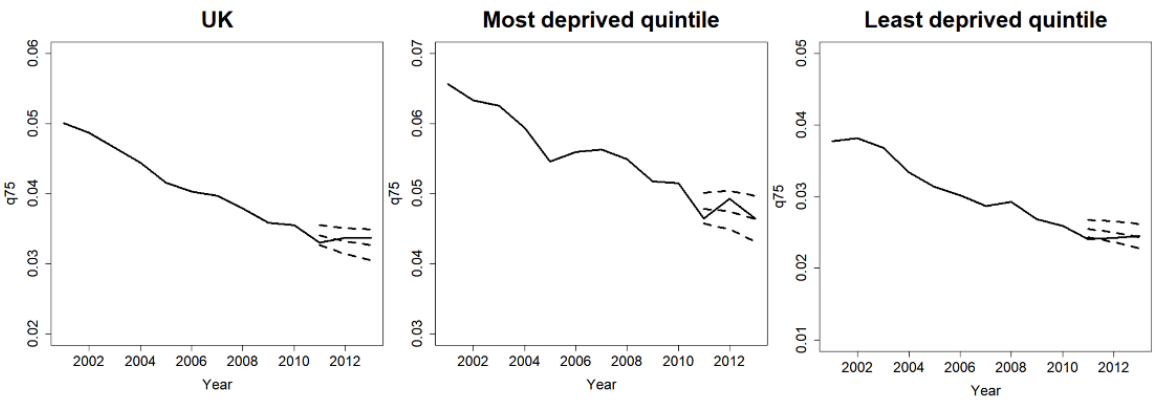


Figure 5.2.16: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (6).

Model (7)

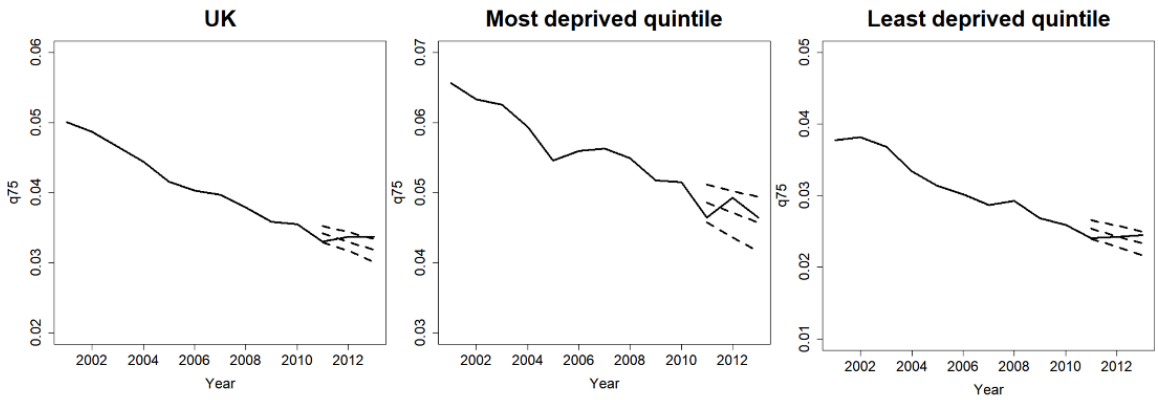


Figure 5.2.17: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (7).

Model (8)

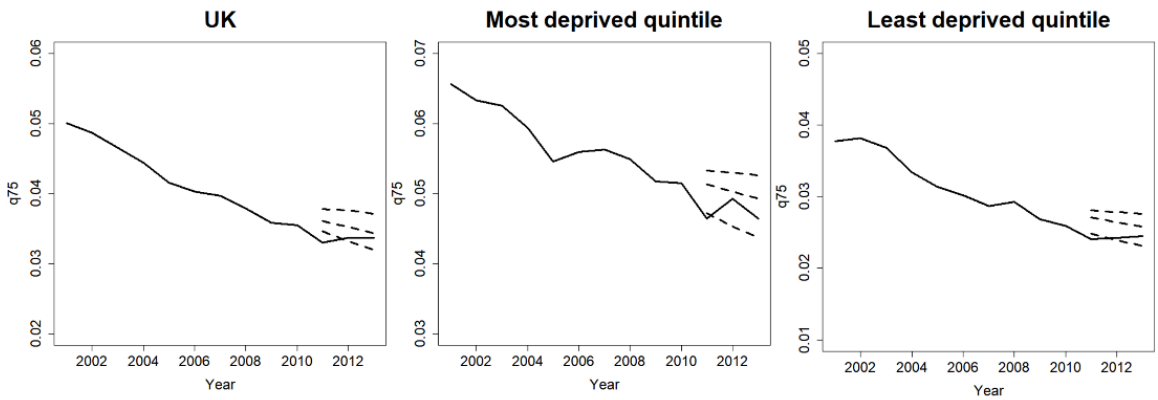


Figure 5.2.18: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (8).

Model (9)

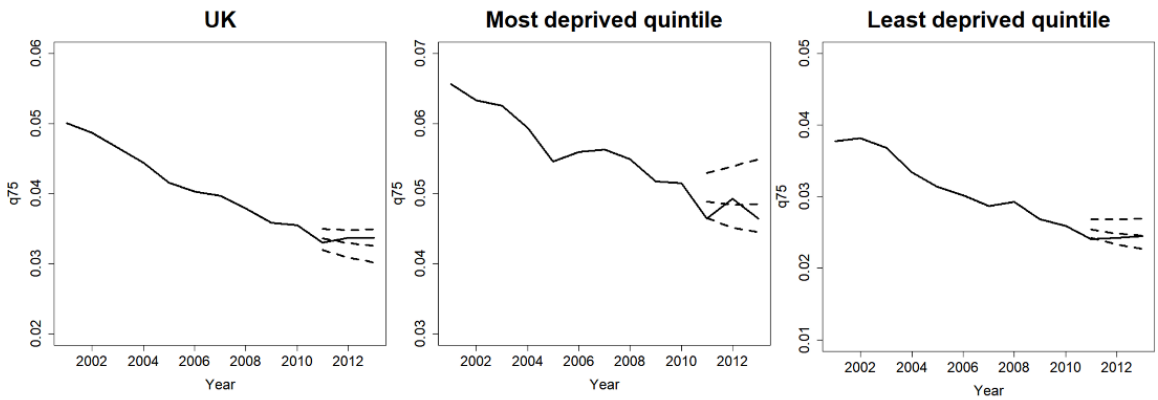


Figure 5.2.19: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (9).

Model (10)

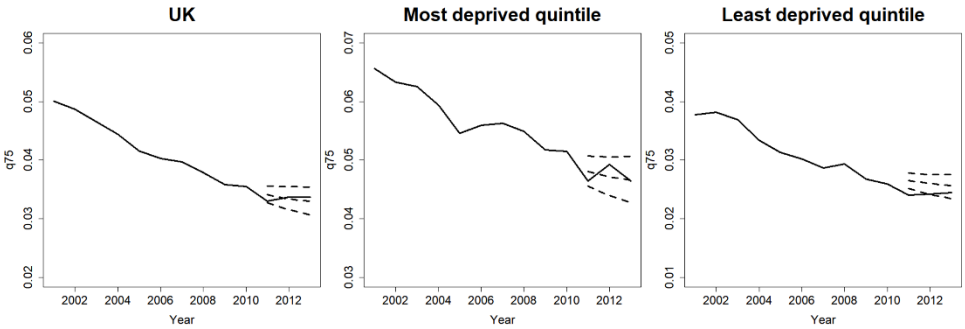


Figure 5.2.20: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (10).

Model (11)

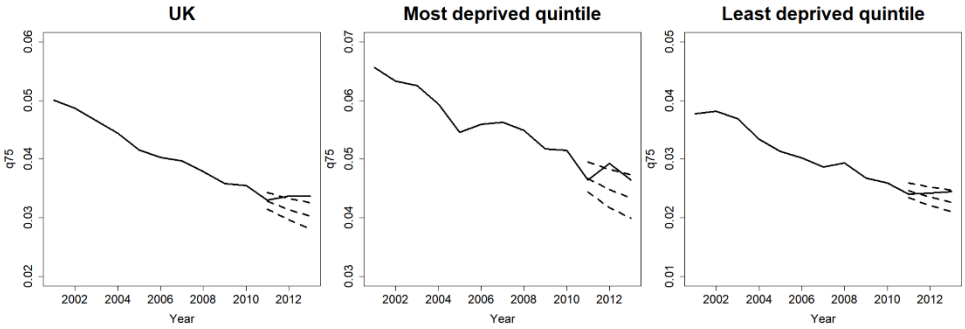


Figure 5.2.21: Simulated prediction intervals and actual $q_{75,t}$ (2011-2013) for the UK population (left), most deprived quintile in England (middle), and least deprived quintile in England under models (21).

5.3 A proxy hedge for the survival probabilities

In this section, a simple survival index hedge is used as a proxy to examine the hedging effectives under the multi-population models. As the annuity values are dependent mainly on the survival probabilities, further assumptions such as the interest rate dynamics are not considered here. The hedging effectiveness can be calculated in terms of variance reduction[7] and it is defined as:

$$1 - \frac{\mathrm{Var}\left({}_{s}p_{x,2014}^{(B)} - {}_{s}p_{x,2014}^{(R)}\right)}{\mathrm{Var}\left({}_{s}p_{x,2014}^{(B)}\right)}$$

where ${}_{s}p_{x,2014}^{(i)}$ refers to the probability that a person aged exactly x in year 2014 will survive to age $(x+s)$, (B) in the superscript refers to the book population (the deprivation subgroups), and (R) in the superscript refers to the reference population (UK population). Year 2014 is chosen to be the starting year as it is the first year in the projection period. The survival probability ${}_{s}p_{x,2014}^{(i)}$ is calculated on a cohort basis:

$${}_{s}p_{x,2014} = \left(1 - q_{x,2014}\right)\left(1 - q_{x+1,2015}\right) \dots \left(1 - q_{x+s-1,2014+s-1}\right)$$

where $q_{x,t}$ represents the probability that an individual aged x years old at time t will die within a year. For convenience in notation, ${}_{s}p_{x}^{(i)} = {}_{s}p_{x,2014}^{(i)}$ will be used from now on.

Seven different survival probabilities are examined in the study to cover a range of ages and duration, namely ${}_{10}p_{60}$, ${}_{10}p_{70}$, ${}_{10}p_{80}$, ${}_{15}p_{60}$, ${}_{15}p_{70}$, ${}_{20}p_{60}$ and ${}_{20}p_{60}$. Models (1) and (10) are chosen for a visual comparison regarding the variance reduction. Figures 5.3.1 to 5.3.4 show the density plots of the simulated survival probabilities, where the y-axis refers to the density and x-axis refers to the simulated survival probabilities standardised by its mean. The black line refers to the unhedged book which is the simulated distribution of ${}_{s}p_{x}^{(B)}$, while the red line refers to the hedge book which is the simulated distribution of ${}_{s}p_{x}^{(B)} - {}_{s}p_{x}^{(R)}$. As shown in the density plots, the distributions of the hedged book are usually narrower than the distributions of the unhedged book. The narrower the hedged distribution compared to the unhedged distribution, the more effective the hedge. The variance reduction calculated is given in Tables 5.3.1 to 5.3.2[8] for all models.

Tables 5.3.1 and 5.3.2 show some interesting patterns regarding the hedging effectiveness on the survival probabilities for a longer duration and an older starting age. A longer duration usually results in better hedging since the majority of the models ensure that the projected mortality rates of the two

[7] The detail of the proxy hedge can be found in the IFoA/LLMA Phase 1 Report (section 8.4)

[8] Tables 5.3.1 to 5.3.2 include the most and least deprived quintile. The other quintiles are shown in the Appendix.

populations converge at each age in the long term. Starting at an older age also results in better hedging performance, probably because the gap between the two populations reduces at older ages, as discussed in the previous sections.

Figure 5.3.1 and Table 5.3.1 indicate that the hedging performance are similar across all ages and durations under the Lee Carter model with bivariate random walk with drift in model (1) for the most deprived quintile. Surprisingly, as shown in Figure 5.3.2 and in Table 5.3.2, only 6% is hedged for ${}_{10}p_{60}$. The hedging effectiveness is not high because the time series process allows the two populations' mortality levels to go into different directions, which would result in a low hedging performance.

This argument is further supported by examining the hedging values under the Lee-Carter model with different time series structures in models (2) and (3). From the tables, the hedging effectiveness is similar between these two models since there is a connection between the time varying factors of the two populations, and so their mortality rates are more likely to move in the same direction, resulting in the smaller variance of the hedged book.

Under model (4), the hedging values are even higher across all ages and durations compared to the previous three models, because the two populations have the common sensitivity measure alongside with the connection in the time varying factors.

The age-period-cohort models (5) and (6) have a considerably low hedging performance for starting age 60, as the projected cohort values are required for the future mortality rates, while most or all the cohort values are known for the older ages. In Table 5.3.2, negative values indicate that the variance of the hedged book is even larger than the unhedged book.

Model (7) has a poor performance in hedging due to the use of the approximation method stated in Chapter 3, in which the two populations are not linked.

Models (10) and (11) have the similar issues with models (5) and (6). Survival probabilities of younger ages require the future projected cohort values but the drop-in performance is not as severe. Alongside with model (8), they follow the usual trend in hedging effectiveness according to age and duration.

Finally, model (9) shows a trend opposite to the general trend, since the log ratios of the death rates at older ages are closer to zero and the time varying factor for the ratio is chosen to follow a random walk with drift, so the mortality levels are less likely to move in the same direction.

From the results, models (4), (8), (10), (11) provide a good hedging performance for the survival probabilities. Note that the choice of the time series model has a great impact on the hedge effectiveness. The results in this section will be combined with the results from the other sections for the conclusion about the model comparison in the last chapter.

Table 5.3.1 Variance reduction of the survival probabilities for the most deprived quintile in England.

Model	${}_{10}p_{60}$	${}_{10}p_{70}$	${}_{10}p_{80}$	${}_{15}p_{60}$	${}_{15}p_{70}$	${}_{20}p_{60}$	${}_{20}p_{70}$
(1)	46%	47%	43%	47%	49%	48%	45%
(2)	87%	93%	69%	93%	92%	96%	74%
(3)	79%	89%	58%	45%	75%	93%	70%
(4)	84%	91%	94%	89%	96%	94%	95%
(5)	39%	87%	94%	49%	93%	57%	96%
(6)	45%	88%	94%	54%	94%	62%	96%
(7)	28%	36%	57%	29%	39%	9%	55%
(8)	55%	79%	93%	68%	90%	81%	94%
(9)	56%	71%	69%	64%	63%	71%	70%
(10)	71%	78%	90%	78%	88%	86%	92%
(11)	85%	82%	86%	87%	88%	90%	92%

Table 5.3.2 Variance reduction of the survival probabilities for the least deprived quintile in England

Model	${}_{10}p_{60}$	${}_{10}p_{70}$	${}_{10}p_{80}$	${}_{15}p_{60}$	${}_{15}p_{70}$	${}_{20}p_{60}$	${}_{20}p_{70}$
(1)	6%	45%	48%	26%	49%	39%	49%
(2)	64%	96%	88%	84%	97%	95%	91%
(3)	66%	96%	86%	84%	96%	94%	90%
(4)	66%	87%	94%	80%	95%	90%	98%
(5)	-14%	61%	72%	12%	74%	34%	81%
(6)	-12%	63%	73%	16%	75%	37%	81%
(7)	-19%	41%	28%	3%	47%	9%	28%
(8)	63%	88%	99%	79%	98%	91%	99%
(9)	64%	48%	14%	70%	41%	62%	21%
(10)	70%	90%	97%	84%	97%	93%	98%
(11)	76%	91%	96%	85%	97%	92%	99%

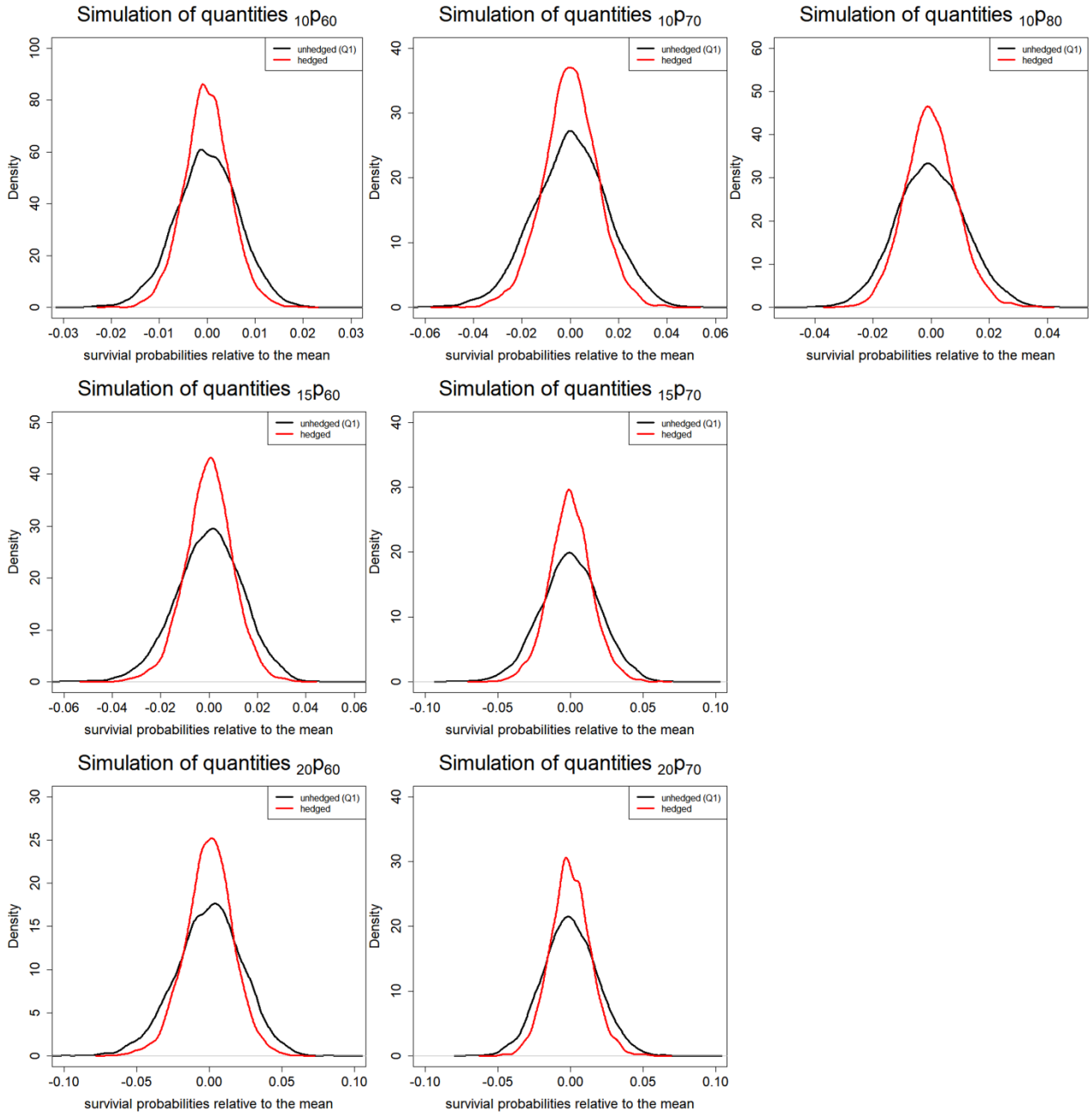


Figure 5.3.1 Simulated densities of ${}_tp_x$ for the most deprived quintile in model (1) where the red line represents the hedged book and the black line represents the unhedged book.

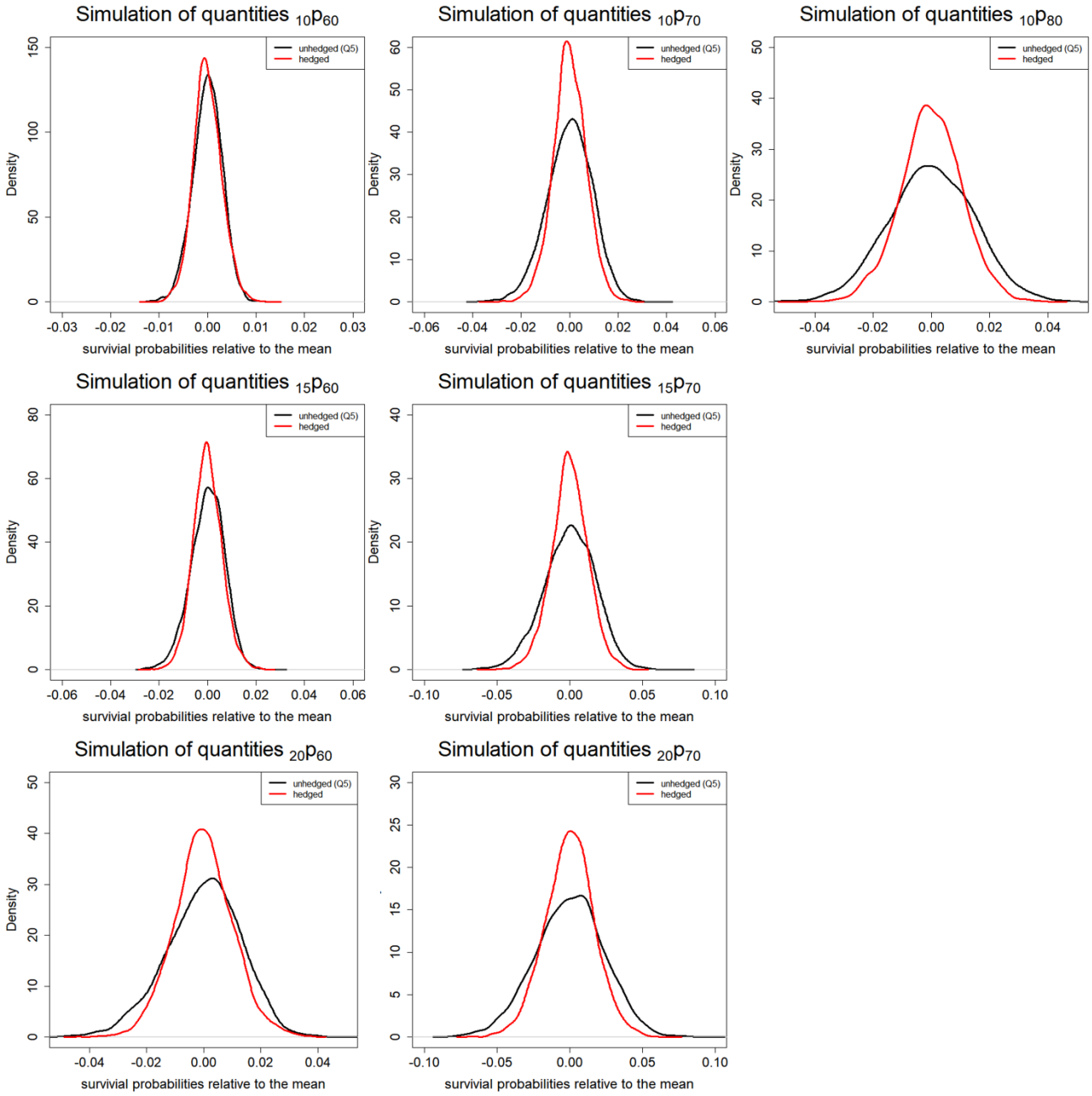


Figure 5.3.2 Simulated densities of ${}_tp_x$ for the least deprived quintile in model (1) where the red line represents the hedged book and the black line represents the unhedged book.

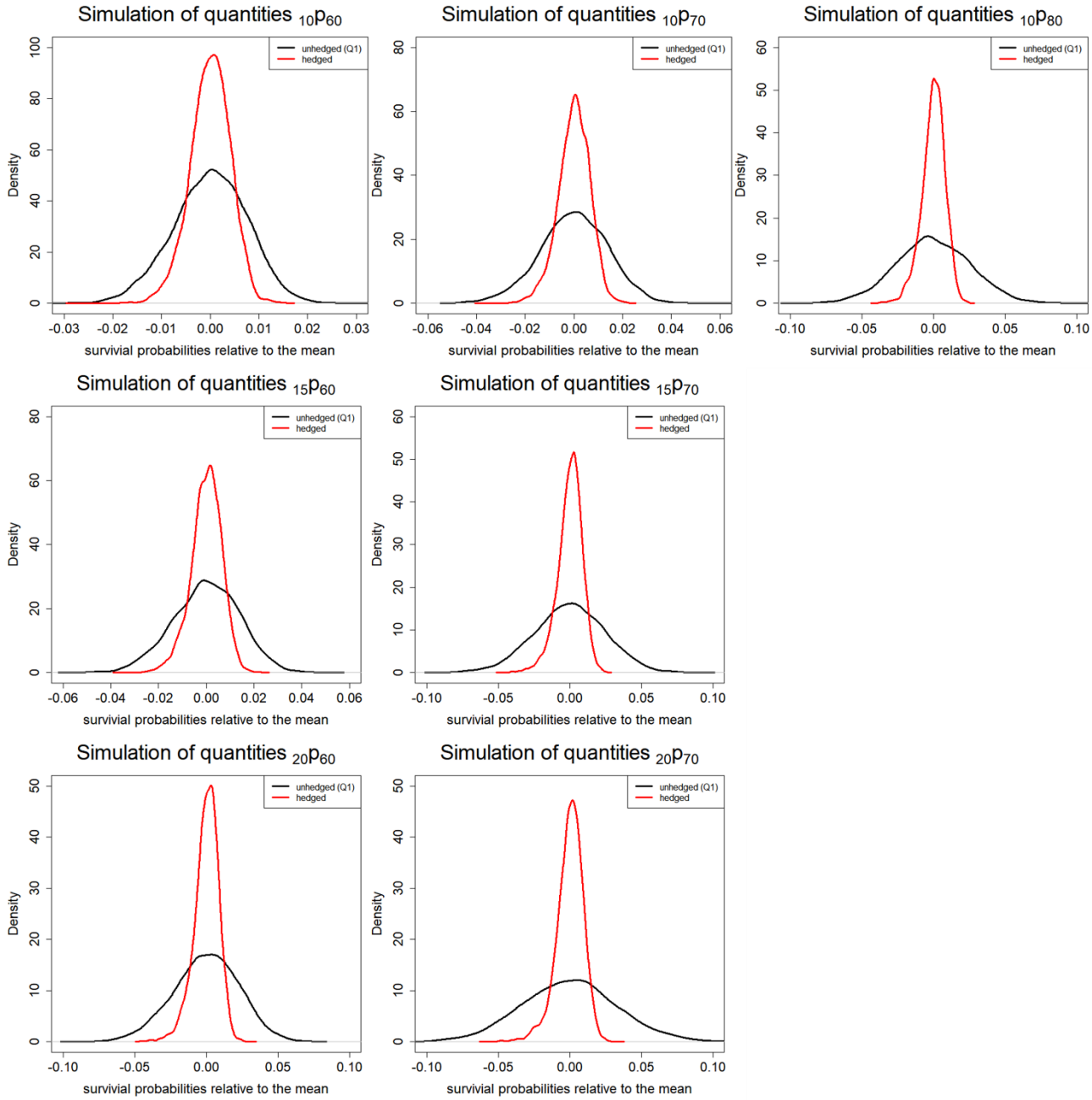


Figure 5.3.3 Simulated densities of ${}_tp_x$ for the most deprived quintile in model (10) where the red line represents the hedged book and the black line represents the unhedged book.

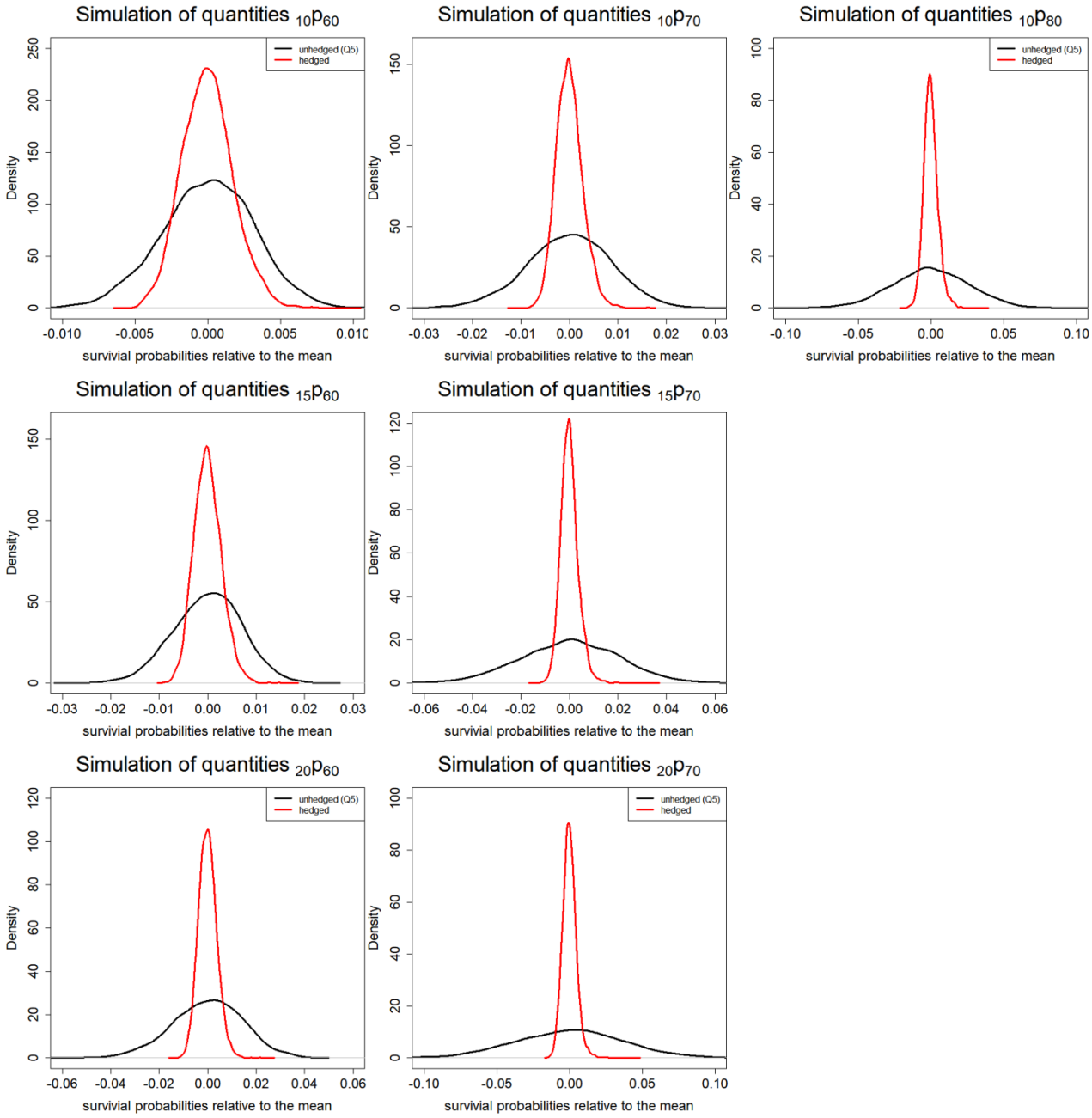


Figure 5.3.4 Simulated densities of ${}_tp_x$ for the least deprived quintile in model (10) where the red line represents the hedged book and the black line represents the unhedged book.

Chapter 6

Concluding remarks

This study has studied 11 multi-population models in the literature and compared their fitting, forecasting performances, and the implied hedging effectiveness. It can be seen from the fitting results that the models with a cohort effect are more desirable for the UK population with a lower BIC, lower MAPE values for the fitted death rates, and no systematic patterns in the residuals plotted against the cohort. The forecasting results show that all models are adequate based on the MAPE for the central estimates. For the hedging effectiveness, models (4), (8), (10), and (11) provide a consistent hedging effectiveness across the age and duration. Looking at these results, as a whole, the M7-M5 Model and CAE + Cohort model outperform the others. This conclusion is also consistent with the recent study done by Villegas et al (2017).

There is one potential issue regarding the MAPE values for the central estimates. The measure does not clearly distinguish the prediction accuracy between the models because of the limited data period available for the deprivation subgroups. One possible way to address this is to combine the recent deprivation subgroups data with an older dataset and conduct the analysis again. Another issue is that the choice of time series model has an important impact on the prediction intervals and hedging effectiveness. Using different time series models for the same fitting model produces a variety of results as shown in Chapter 5 regarding the Lee-Carter model. The same procedure can be extended to all fitting models by testing different time series models. For further research, a more extensive analysis on the hedging performance can be carried out by considering different mortality instruments and interest rate dynamics in addition to the models discussed in this study.

References

Bray, I. (2002). Application of Markov chain Monte Carlo methods to projecting cancer incidence and mortality. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *51*(2), 151-164.

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

Booth, H., Maindonald, J., & Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population studies*, *56*(3), 325-336.

Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of actuarial science*, *3*(1-2), 3-43.

Brillinger, D. R. (1986). A biometrics invited paper with discussion: the natural variability of vital rates and associated statistics. *Biometrics*, 693-734.

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

Cairns, A. J., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, *73*(4), 687-718.

Cairns, A. J., Blake, D., & Dowd, K. (2008). Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, *2008*(2-3), 79-113.

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

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., & Khalaf-Allah, M. (2011). Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin: The* Journal *of the IAA*, *41*(1), 29-59.

Chen, H., MacMinn, R., & Sun, T. (2015). Multi-population mortality models: A factor copula approach. *Insurance: Mathematics and Economics*, *63*, 135-146.

Currie, I. (2006). Smoothing and forecasting mortality rates with P-Splines, Heriot Watt University.

Danesi, I. L., Haberman, S., & Millossovich, P. (2015). Forecasting mortality in subpopulations using Lee–Carter type models: A comparison. *Insurance: Mathematics and Economics*, *62*, 151-161.

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

Haberman, S., Kaishev, V., Millossovich, P., Villegas, A., Baxter, S., Gaches, A., Gunnlaugsson, & Sison, M. (2014). Longevity Basis Risk: A Methodology for Assessing Basis Risk. Sessional research paper, Institute and Faculty of Actuaries, 8 December 2014. Tech. rep., Cass Business School, City University London and Hymans Robertson LLP.

Haberman, S., & Renshaw, A. (2012). Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, *50*(3), 309-333.

Hatzopoulos, P., & Haberman, S. (2013). Common mortality modeling and coherent forecasts. An empirical analysis of worldwide mortality data. *Insurance: Mathematics and Economics*, *52*(2), 320-337.

Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography*, *50*(1), 261-283.

Jarner, S. F., & Kryger, E. M. (2011). Modelling adult mortality in small populations: The SAINT model. *ASTIN Bulletin: The Journal of the IAA*, *41*(2), 377-418.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting US mortality. *Journal of the American statistical association*, *87*(419), 659-671.

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

Li, J. (2012). Mortality projection methods: Extensions and modifications, Macquarie University, Australia.

Li, J. (2013). A Poisson common factor model for projecting mortality and life expectancy jointly for females and males. *Population studies*, *67*(1), 111-126.

Li, J., Dacorogna, M., & Tan, C. I. (2014). The impact of joint mortality modelling on hedging effectiveness of mortality derivatives. In *Tenth International Longevity Risk and Capital Markets Solutions Conference, Santiago, Chile*.

Li, J. S. H., Zhou, R., & Hardy, M. (2015). A step-by-step guide to building two-population stochastic mortality models. *Insurance: Mathematics and Economics*, *63*, 121-134.

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

Li J., Yeo K. L., Pakshong C., Chan W. S., Kogure A., & Li J. S. H. (2012). Mortality experience in Asia-Pacific and modeling and management of longevity risk, Insurance Risk and Finance Research Centre Report.

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

Luy, M. (2003). Causes of male excess mortality: insights from cloistered populations. *Population and Development Review*, *29*(4), 647-676.

Kleinow, T. (2015). A common age effect model for the mortality of multiple populations. *Insurance: Mathematics and Economics*, *63*, 147-152.

Ngai, A., & Sherris, M. (2011). Longevity risk management for life and variable annuities: The effectiveness of static hedging using longevity bonds and derivatives. *Insurance: Mathematics and Economics*, *49*(1), 100-114.

Plat, R. (2009). Stochastic portfolio specific mortality and the quantification of mortality basis risk. *Insurance: Mathematics and Economics*, *45*(1), 123-132.

R Core Team (2017). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

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

Stevenson, M., & Wilson, A. (2008, May). Mortality of Public Sector Scheme Pensioners 2005-2007 Update. In *Institute of Actuaries of Australia 4th Financial Services Forum*.

Tan, C. I., Li, J., Li, J. S. H., & Balasooriya, U. (2014). Parametric mortality indexes: From index construction to hedging strategies. *Insurance: Mathematics and Economics*, *59*, 285-299.

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

Villegas, A. M., Haberman, S., Kaishev, V. K., & Millossovich, P. (2017). A comparative study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin: The Journal of the IAA*, *47*(3), 631-679.

Wang, C. W., Yang, S. S., & Huang, H. C. (2015). Modeling multi-country mortality dependence and its application in pricing survivor index swaps—A dynamic copula approach. *Insurance: Mathematics and Economics*, *63*, 30-39.

White, K. M. (2002). Longevity advances in high-income countries, 1955–96. *Population and Development Review*, *28*(1), 59-76.

Willets, R. C. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, *10*(4), 833-877.

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.

Wilson, C. (2001). On the scale of global demographic convergence 1950–2000. *Population and Development Review*, *27*(1), 155-171.

Wilson, C. (2011). Understanding global demographic convergence since 1950. *Population and Development Review*, *37*(2), 375-388.

Wong, K. (2016). Alternative parameterisations of the Poisson common factor model for modelling mortality jointly for both sexes, Macquarie University, Australia.

Xu, J. (2014). Joint mortality modelling and forecasting: a new joint model based on the Wang Transform, Macquarie University, Australia.

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

Yang, B., Li, J., & Balasooriya, U. (2016). Cohort extensions of the Poisson common factor model for modelling both genders jointly. *Scandinavian Actuarial Journal*, *2016*(2), 93-112.

Zhou, R., Li, J. S. H., & Tan, K. S. (2013). Pricing Standardized Mortality Securitizations: A Two-Population Model With Transitory Jump Effects. *Journal of Risk and Insurance*, *80*(3), 733-774.

Appendix

Standardized deviance residuals

The standardized deviance residuals are defined similarly as Li (2013) and Villegas and Haberman (2014)

$$r_{x,t,i} = sign(d_{x,t,i} - \hat{d}_{x,t,i})\sqrt{\frac{2(d_{x,t,i}\,/\hat{d}_{x,t,i}) - d_{x,t,i} + \hat{d}_{x,t,i}}{\widehat{\emptyset}}}$$

where n_p is the number of parameters being estimated and n_d is the number of observations and the over dispersion parameter $\widehat{\emptyset} = \frac{\text{deviance}}{n_d - n_p}$, the deviance for the Poisson death count is expressed as

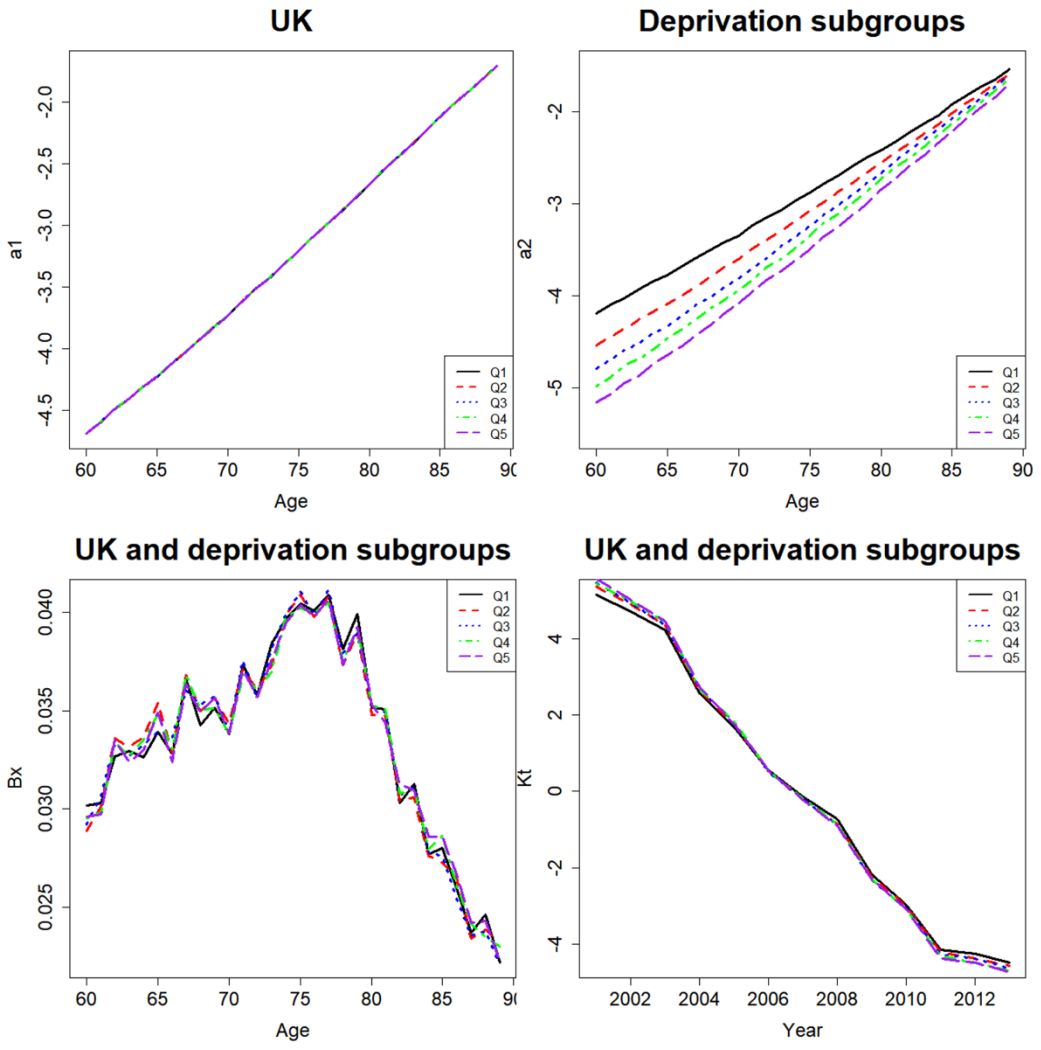
deviance $= 2\left[d_{x,t,i}\ln\left(\frac{d_{x,t,i}}{\hat{d}_{x,t,i}}\right) - d_{x,t,i} + \hat{d}_{x,t,i}\right]$

Fitting result for (model (A1) and model (A2))

Table A1 BIC values for using different numbers of additional factors for model (A1) and model (A2).

Additional factors	Most deprived quintile	2nd most deprived quintile	3rd most deprived quintile	4th most deprived quintile	Least deprived quintile
0	10663	**10174**	**10052**	**10068**	**10001**
1	**10578**	10331	10195	10226	10165
2	10759	10535	10427	10463	10388
3	10990	10826	10743	10773	10685
4	11301	11162	11103	11124	11041
5	11660	11536	11502	11500	11429
6	12048	11944	11920	11907	11847

Table A1 shows that 1 additional factor (model (A1)) generates the lowest BIC value for the most deprived quintile and 0 additional factors needed for the lowest BIC values for the other four quintiles. This indicates that the base mortality level $\alpha_{x,i}$ is enough to distinguish the mortality experience for the four quintiles. $\beta_{x,i}$ and $\kappa_{t,i}$. Model (A1) is needed for most deprived quintile as the additional factor with the lower BIC values. Model (A2) is not needed for the deprivation subgroups when the data only span for 13 years for both populations. The fitting result and residual for model (A1) is provided for all quintiles for comparison. Note that the large fluctuation of $\beta_{x,1}$ and $\beta_{x,2}$ is due to the small corresponding $\kappa_{t,1}$ and $\kappa_{t,2}$ values. The residual plot is also provided and the systematic pattern for residual plot against cohort is much weaker with less data available for the UK population. Forecasting is not performed since most quintiles do not require any additional factors and the period of data is not consistent with the majority of models used in the main study.


UK
Deprivation subgroups
UK and deprivation subgroups
UK and deprivation subgroups
a1
a2
Bx
Kt
Age
Year
Q1
Q2
Q3
Q4
Q5

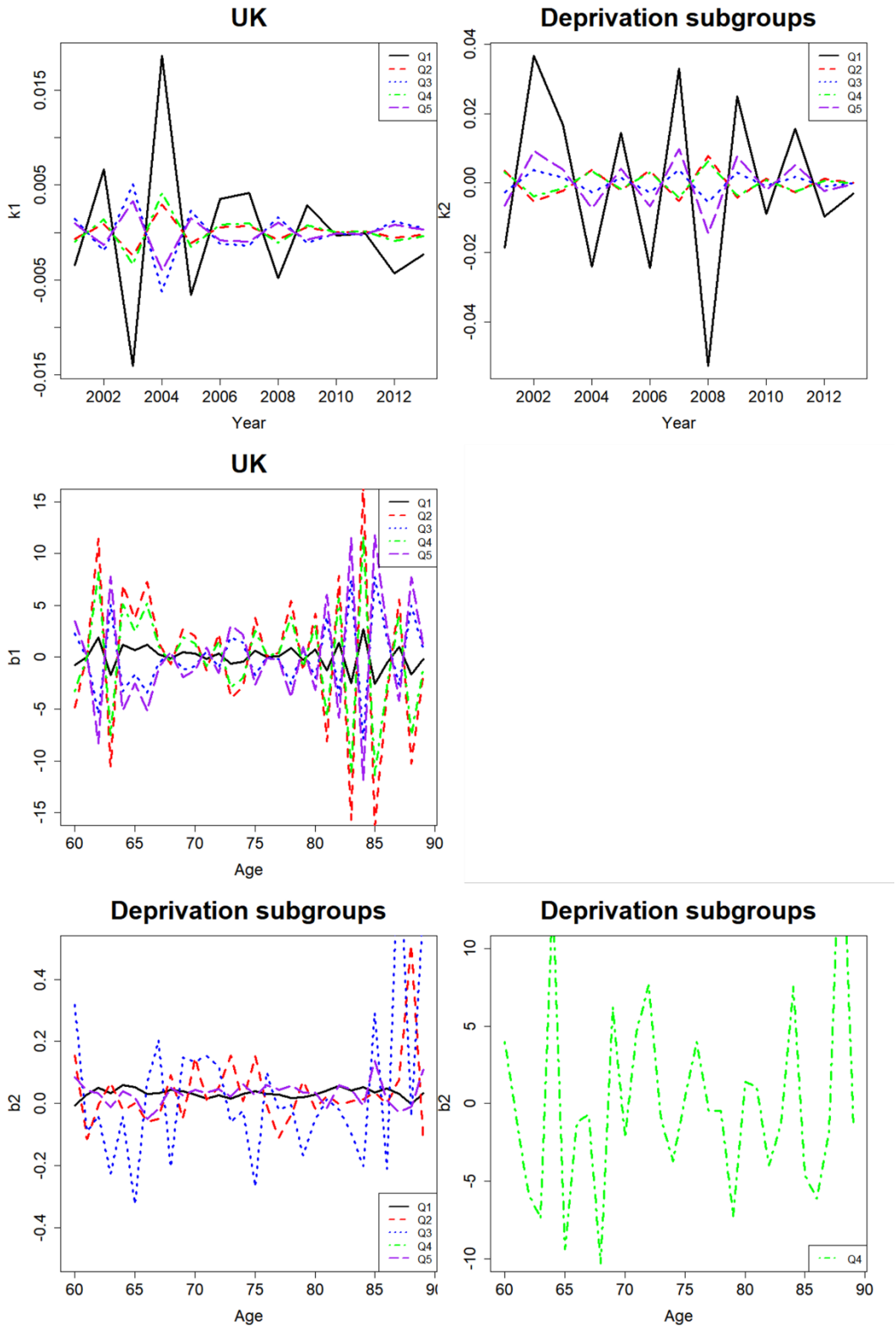


Figure A1: Parameter estimates of the augmented common factor model for UK population and the deprivation subgroups in England.

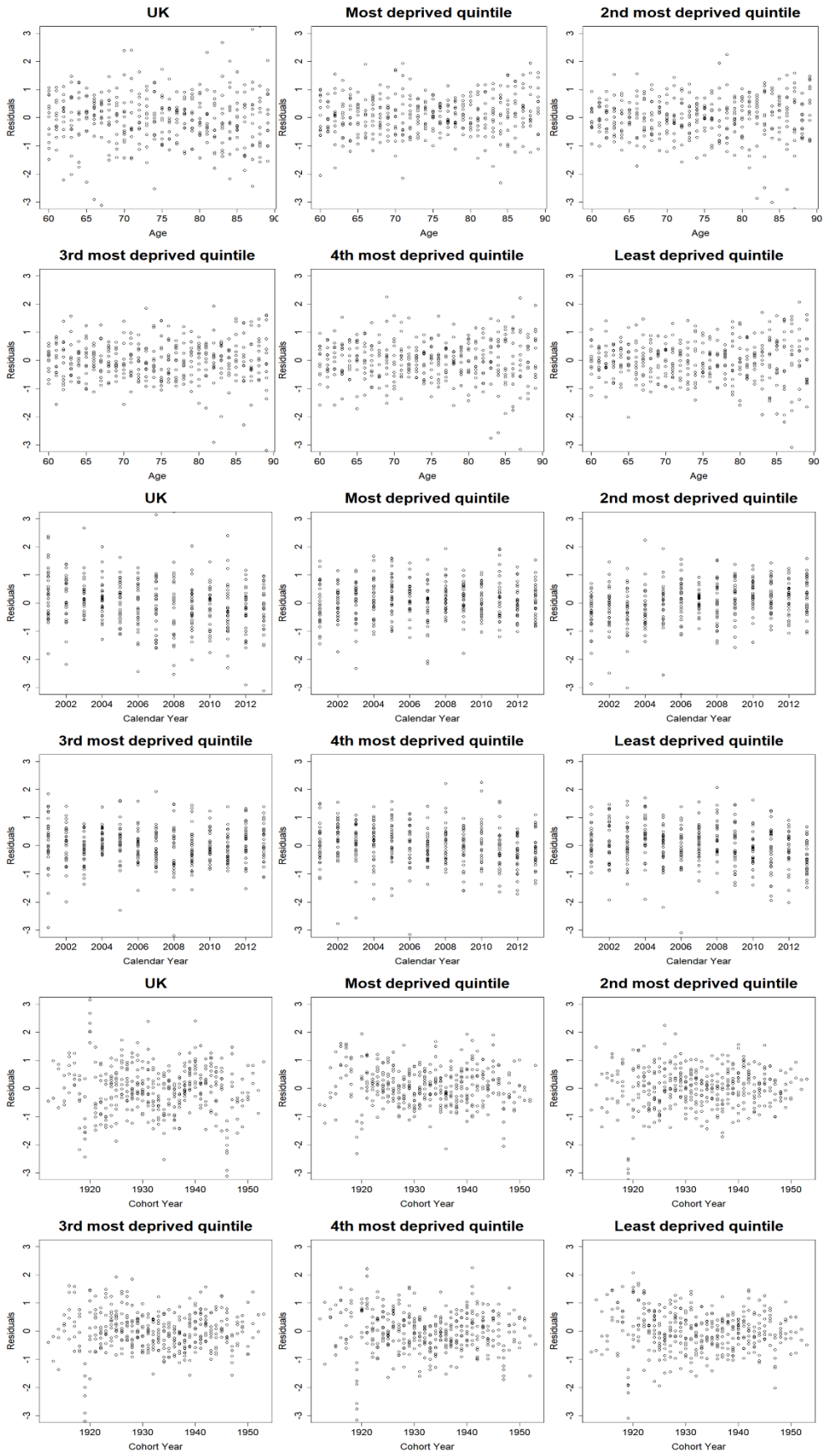
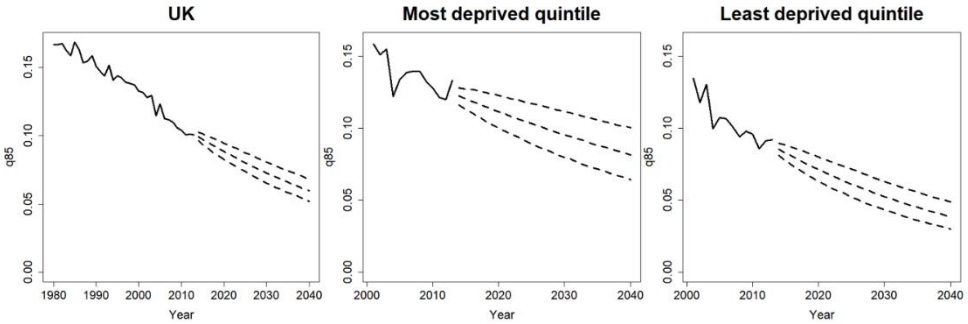


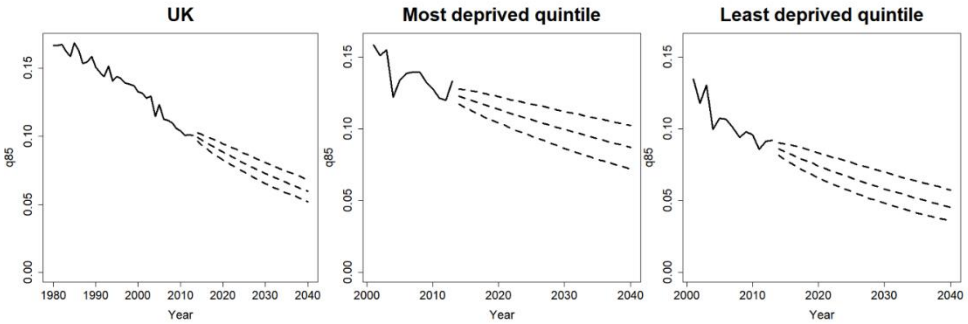
Figure A2: Standardised deviance residuals for the augmented common factor model, UK and the deprivation subgroups. The top two rows are plotted against age, middle two rows are plotted against calendar year, and bottom two rows are plotted against cohort year.

The prediction interval for q_{85} in (model (1) – (11))

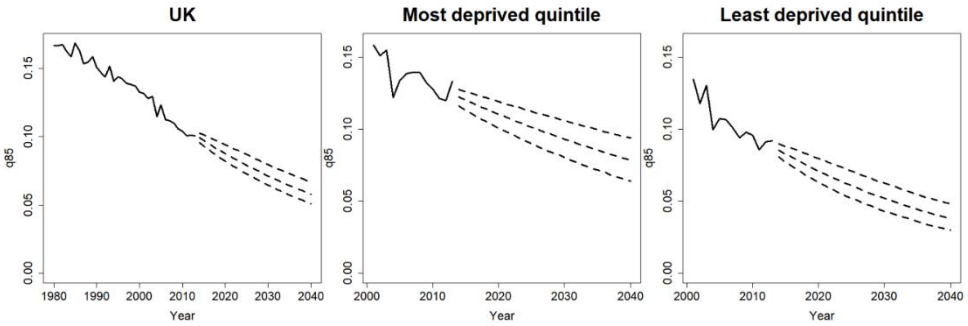
Model (1)



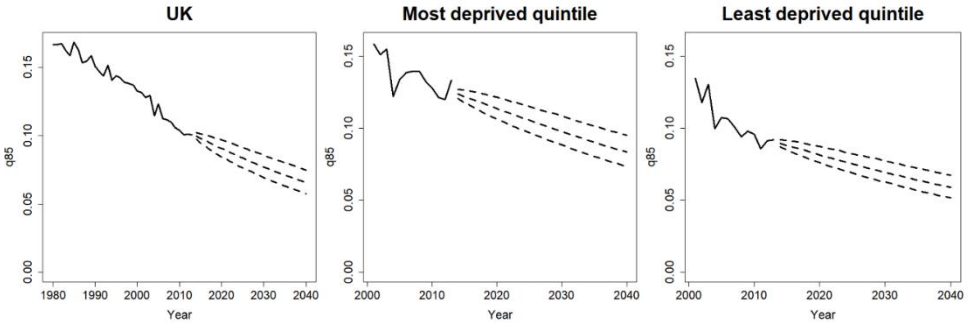
Model (2)



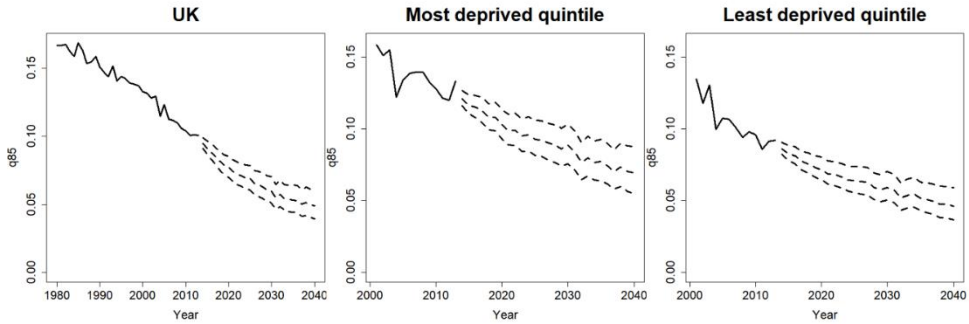
Model (3)



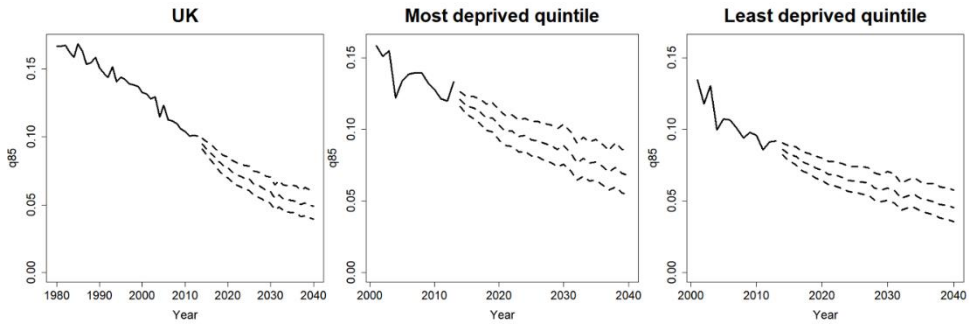
Model (4)



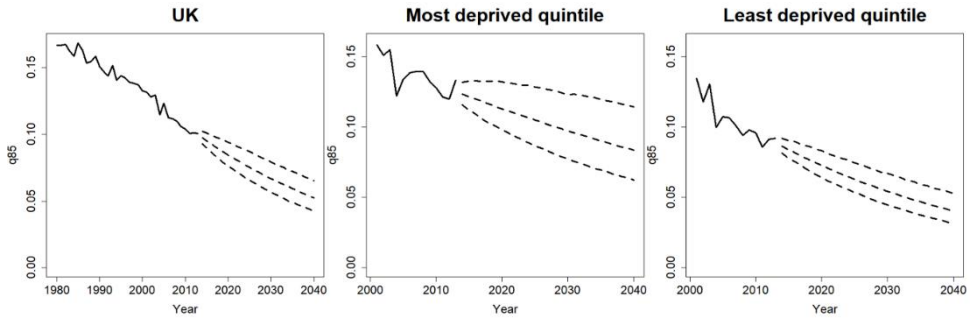
Model (5)



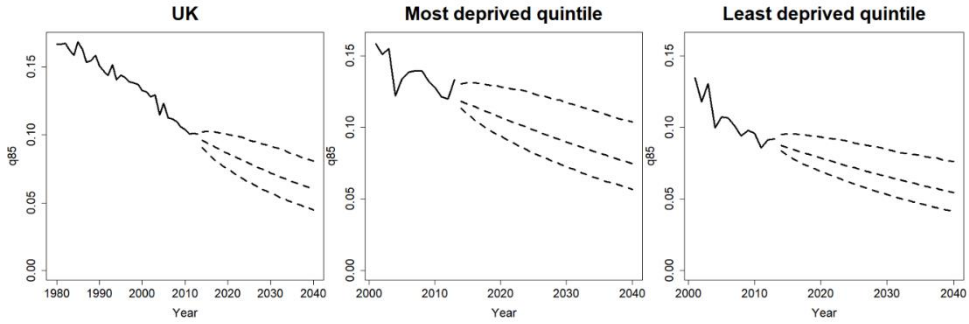
Model (6)



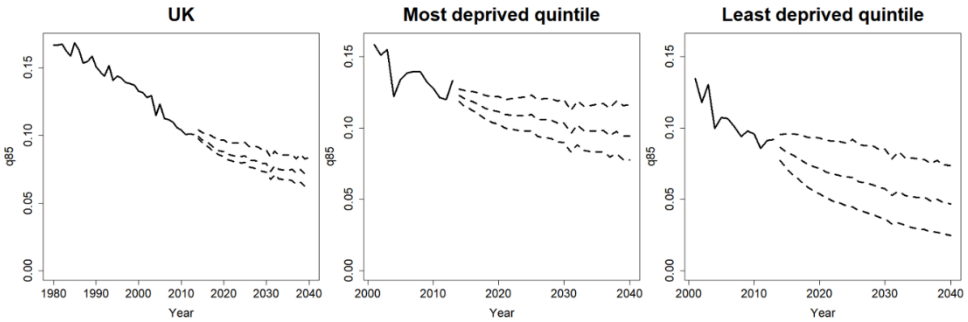
Model (7)



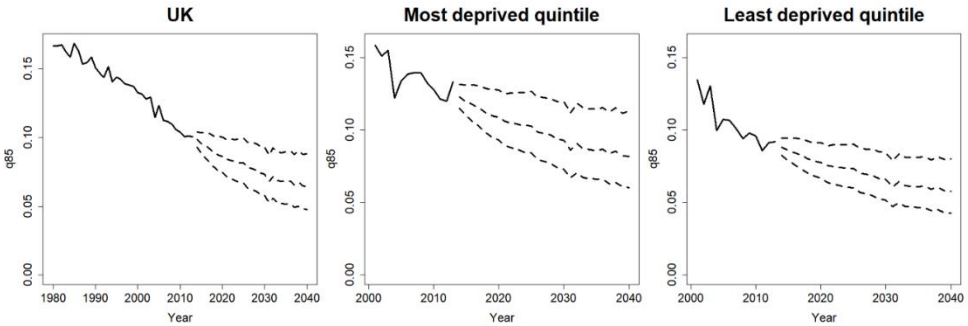
Model (8)



Model (9)



Model (10)



Model (11)

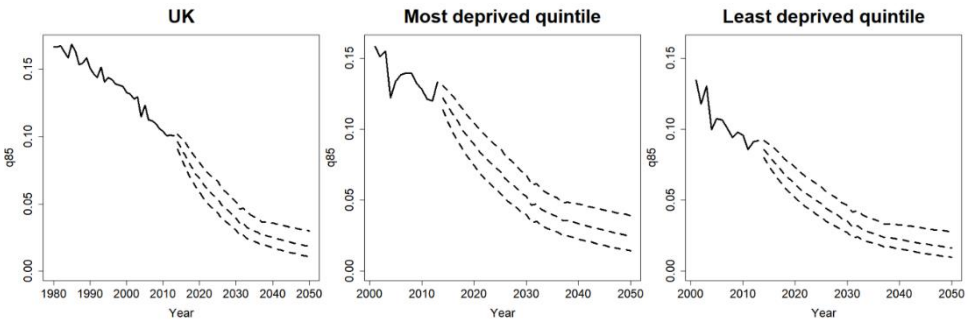


Figure A3-A13: Projected 95% prediction intervals for q_{85} (dotted line) for the UK population (left), most deprived quintile in England (mid) and least deprived quintile in England (right) for models (1) – (11)

Extra table for Chapter 5.2 (second, third and fourth most deprived quintile)

Table A2 MAPE values for the projected death rates (2009-2013) for UK & second most deprived quintile.

Model	UK	2nd most deprived quintile	Total	RANK
(1)	1.27% (4.08%)	2.02% (5.47%)	1.65% (4.78%)	[7]
(2)	1.48% (4.48%)	1.94% (5.45%)	1.71% (4.97%)	[9]
(3)	1.26% (4.09%)	2.11% (5.71%)	1.69% (4.90%)	[8]
(4)	1.46% (4.45%)	1.29% (3.87%)	1.38% (4.16%)	[5]
(5)	1.01% (3.22%)	1.30% (3.93%)	1.16% (3.58%)	[2]
(6)	1.00% (3.17%)	1.37% (4.19%)	1.18% (3.68%)	[3]
(7)	-	-	-	-
(8)	1.43% (4.85%)	1.56% (4.70%)	1.50% (4.77%)	[6]
(9)	1.33% (3.90%)	1.33% (3.97%)	1.33% (3.94%)	[4]
(10)	1.12% (3.24%)	1.06% (3.19%)	1.09% (3.21%)	[1]
(11)	1.32% (3.78%)	2.17% (5.72%)	1.74% (4.75%)	[10]

Table A3 MAPE values for the projected death rates (2009-2013) for UK & third most deprived quintile.

Model	UK	3rd most deprived quintile	Total	RANK
(1)	1.27% (4.08%)	1.94% (5.56%)	1.60% (4.82%)	[7]
(2)	1.48% (4.48%)	1.78% (5.35%)	1.63% (4.91%)	[9]
(3)	1.24% (4.10%)	2.00% (5.69%)	1.62% (4.89%)	[8]
(4)	1.46% (4.44%)	1.46% (4.69%)	1.46% (4.56%)	[4]
(5)	1.01% (3.22%)	1.21% (3.85%)	1.11% (3.54%)	[1]
(6)	1.00% (3.17%)	1.26% (4.08%)	1.13% (3.63%)	[2]
(7)	-	-	-	-
(8)	1.43% (4.85%)	1.86% (5.97%)	1.65% (5.41%)	[10]
(9)	1.33% (3.90%)	1.78% (5.41%)	1.56% (4.66%)	[6]
(10)	1.12% (3.24%)	1.15% (3.86%)	1.13% (3.55%)	[3]
(11)	1.32% (3.78%)	1.74% (4.96%)	1.53% (4.37%)	[5]

Table A4 MAPE values for the projected death rates (2009-2013) for UK & fourth most deprived quintile.

Model	UK	4th most deprived quintile	Total	RANK
(1)	1.27% (4.08%)	1.83% (5.57%)	1.55% (4.83%)	[5]
(2)	1.48% (4.48%)	1.94% (6.07%)	1.71% (5.28%)	[10]
(3)	1.28% (4.11%)	1.84% (5.61%)	1.56% (4.86%)	[6]
(4)	1.46% (4.44%)	1.84% (5.82%)	1.65% (5.13%)	[7]
(5)	1.01% (3.22%)	1.19% (3.98%)	1.10% (3.60%)	[1]
(6)	1.00% (3.17%)	1.22% (4.14%)	1.11% (3.66%)	[2]
(7)	-	-	-	-
(8)	1.43% (4.85%)	1.92% (6.49%)	1.68% (5.67%)	[8]
(9)	1.33% (3.90%)	2.06% (6.53%)	1.70% (5.22%)	[9]
(10)	1.12% (3.24%)	1.59% (5.17%)	1.35% (4.21%)	[3]
(11)	1.32% (3.78%)	1.47% (4.71%)	1.40% (4.24%)	[4]

Table A5 MAPE values for the projected death rates (2011-2013) for UK & second most deprived quintile.

Model	UK	2nd most deprived quintile	Total	RANK
(1)	1.13% (3.92%)	1.36% (4.12%)	1.24% (4.02%)	[8]
(2)	1.02% (3.46%)	1.29% (3.97%)	1.15% (3.71%)	[5]
(3)	1.20% (4.19%)	1.56% (4.60%)	1.38% (4.39%)	[9]
(4)	1.01% (3.42%)	1.26% (3.81%)	1.13% (3.62%)	[4]
(5)	1.14% (3.21%)	1.27% (3.80%)	1.21% (3.51%)	[6]
(6)	1.14% (3.20%)	1.29% (3.87%)	1.21% (3.53%)	[7]
(7)	1.03% (3.09%)	1.15% (3.40%)	1.09% (3.24%)	[3]
(8)	1.41% (4.42%)	1.70% (4.87%)	1.56% (4.65%)	[10]
(9)	0.73% (2.22%)	1.28% (3.64%)	1.00% (2.93%)	[2]
(10)	0.63% (1.96%)	1.02% (2.98%)	0.83% (2.47%)	[1]
(11)	1.65% (4.61%)	2.21% (5.79%)	1.93% (5.20%)	[11]

Table A6 MAPE values for the projected death rates (2011-2013) for UK & third most deprived quintile.

Model	UK	3rd most deprived quintile	Total	RANK
(1)	1.13% (3.92%)	1.32% (4.27%)	1.22% (4.09%)	[8]
(2)	1.02% (3.46%)	1.17% (3.84%)	1.10% (3.65%)	[3]
(3)	1.19% (3.52%)	1.44% (4.62%)	1.31% (4.35%)	[9]
(4)	1.01% (3.41%)	1.28% (4.38%)	1.14% (3.90%)	[5]
(5)	1.14% (3.21%)	1.13% (3.57%)	1.14% (3.39%)	[4]
(6)	1.14% (3.20%)	1.15% (3.67%)	1.15% (3.43%)	[6]
(7)	1.03% (3.09%)	1.31% (4.20%)	1.17% (3.65%)	[7]
(8)	1.41% (4.42%)	1.82% (5.81%)	1.62% (5.12%)	[10]
(9)	0.73% (2.22%)	1.03% (3.49%)	0.88% (2.85%)	[2]
(10)	0.63% (1.96%)	1.12% (3.75%)	0.87% (2.86%)	[1]
(11)	1.65% (4.61%)	2.05% (5.83%)	1.85% (5.22%)	[11]

Table A7 MAPE values for the projected death rates (2011-2013) for UK & third most deprived quintile.

Model	UK	4th most deprived quintile	Total	RANK
(1)	1.13% (3.92%)	1.25% (4.20%)	1.19% (4.06%)	[7]
(2)	1.02% (3.46%)	1.19% (4.09%)	1.11% (3.78%)	[3]
(3)	1.19% (4.18%)	1.33% (4.42%)	1.26% (4.30%)	[9]
(4)	1.00% (3.41%)	1.38% (4.52%)	1.19% (3.96%)	[8]
(5)	1.14% (3.21%)	1.18% (3.90%)	1.16% (3.56%)	[6]
(6)	1.14% (3.20%)	1.16% (3.78%)	1.15% (3.49%)	[5]
(7)	1.03% (3.09%)	1.19% (3.85%)	1.11% (3.47%)	[4]
(8)	1.41% (4.42%)	1.48% (4.94%)	1.45% (4.68%)	[10]
(9)	0.73% (2.22%)	1.31% (4.23%)	1.02% (3.23%)	[2]
(10)	0.63% (1.96%)	1.04% (3.59%)	0.83% (2.78%)	[1]
(11)	1.65% (4.61%)	1.68% (5.36%)	1.67% (4.99%)	[11]

Extra table for Chapter 5.3 (second, third and fourth most deprived quintile)

Table A8 Variance reduction of the survival probabilities for the second most deprived quintile in England.

Model	${}_{10}p_{60}$	${}_{10}p_{70}$	${}_{10}p_{80}$	${}_{15}p_{60}$	${}_{15}p_{70}$	${}_{20}p_{60}$	${}_{20}p_{70}$
(1)	44%	46%	44%	46%	46%	47%	45%
(2)	91%	95%	89%	95%	95%	97%	94%
(3)	91%	93%	87%	96%	95%	85%	89%
(4)	93%	94%	94%	96%	96%	97%	97%
(5)	35%	83%	86%	47%	89%	56%	92%
(6)	37%	82%	86%	49%	88%	57%	92%
(7)	17%	51%	9%	38%	48%	44%	12%
(8)	75%	91%	92%	87%	95%	94%	96%
(9)	76%	78%	27%	76%	76%	81%	39%
(10)	93%	93%	95%	95%	96%	97%	98%
(11)	97%	94%	95%	97%	97%	98%	98%

Table A9 Variance reduction of the survival probabilities for the second most deprived quintile in England.

Model	${}_{10}p_{60}$	${}_{10}p_{70}$	${}_{10}p_{80}$	${}_{15}p_{60}$	${}_{15}p_{70}$	${}_{20}p_{60}$	${}_{20}p_{70}$
(1)	45%	55%	50%	51%	55%	55%	53%
(2)	86%	95%	90%	94%	96%	97%	94%
(3)	90%	95%	91%	98%	97%	99%	95%
(4)	99%	99%	99%	99%	99.7%	99.8%	99.8%
(5)	38%	82%	86%	52%	88%	62%	91%
(6)	40%	81%	85%	53%	87%	63%	91%
(7)	34%	42%	35%	-141%	49%	-15%	38%
(8)	93%	97%	98%	96%	99%	99%	99%
(9)	72%	89%	21%	80%	88%	90%	37%
(10)	95%	99%	99%	98%	99%	99%	99.5%
(11)	99%	99%	99%	99%	99.6%	99.7%	99.7%

Table A10 Variance reduction of the survival probabilities for the second most deprived quintile in England.

Model	${}_{10}p_{60}$	${}_{10}p_{70}$	${}_{10}p_{80}$	${}_{15}p_{60}$	${}_{15}p_{70}$	${}_{20}p_{60}$	${}_{20}p_{70}$
(1)	54%	63%	59%	58%	63%	62%	61%
(2)	82%	97%	90%	91%	97%	97%	93%
(3)	79%	96%	90%	89%	96%	52%	88%
(4)	92%	96%	96%	94%	97%	97%	98%
(5)	14%	65%	68%	29%	62%	37%	55%
(6)	15%	64%	68%	28%	61%	36%	53%
(7)	9%	53%	41%	30%	63%	48%	48%
(8)	87%	96%	99%	93%	99%	97%	99.5%
(9)	68%	82%	37%	79%	86%	87%	57%
(10)	90%	97%	98%	95%	99%	98%	99%
(11)	93%	97%	98%	95%	99%	98%	99%