Quantifying biometric life insurance risks with non-parametric smoothing methods

Tomas, J.

Citation for published version (APA):
Chapter 5

Entity specific prospective mortality tables

5.1 Introduction

It is now well documented that the human mortality has globally declined over the 20th century. Life expectancy is greater than ever before and increasing rapidly, see Pitacco et al. (2009, Ch.3). These mortality improvements pose a challenge for the pricing and reserving in life insurance and for the management of public pension regimes. In a pension plan, the longevity risk is transferred from the policyholder to the insurer. The latter has to evaluate his liability with appropriate mortality tables. It is in this context that since 1993 the French regulatory tables for annuities have been prospective mortality tables taking in account the increase of the life expectancy.

Prospective mortality tables allow to determine the remaining lifetime for a group, not according to the conditions of the moment, but given the future developments of living conditions. However, applying exogenous tables to the group considered may result in under-provisioning the annuities, when the mortality of the group is lower than of the reference population. With the international regulations Solvency II and IFRS insurers are required to evaluate their liabilities from realistic assumptions leading to an evaluation of the best estimate. In consequence, for pensions regimes and more generally due to the longevity risk, insurers have to build specific mortality tables, taking into account the expected evolution of the mortality of their insured population, see Planchet and Kamega (2011).

Probably because it was not understood initially in which respects demographic sciences differed from natural sciences, it was believed that mortality
laws similar to those discovered in astronomy and physics could be found. However, none were found and it is far from certain that there are any. As a consequence it is impossible to predict the evolution of mortality as the expected movements of the stars, as notes Henry (1987).

Not being able to predict from laws, but being forced to forecast, Henry (1987) suggests that it is in the experience that we should seek the best means to do it. His view is not radically different from the one expressed by de Laplace (1829) where ignorance is temporary and research will increase our understanding and help formulating accurate forecasts. « Imagine [...] an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings which compose it [...] to it, nothing would be uncertain and the future, as the past, would be present in its eyes ».

Laws can be replaced by assumptions about the future characteristics of a population to deduce future perspectives, in numbers and structure, of this population. In the absence of laws, we observe some regularities and patterns in mortality, either permanent or specific to a certain period, from which we can produce forecasts that most of the time are sufficient for our needs. It is generally accepted that the demographic phenomenon of inertia is sufficient (apart from periods of war) for extrapolation of past trends, see Booth (2006).

In this chapter, we present a two steps approach to build entity specific mortality tables. From portfolios of several insurance companies, the first step consists in constructing global prospective mortality tables by gender. For clarity, only results about the male population are presented. By reasoning globally, this table summarizes the male mortality experience of these portfolios. The heterogeneity present between the portfolios is taken into account in a second step. The male prospective table is then used to adjust the mortality specifically to each male insured portfolio. The computations are carried out with the help of R, R Development Core Team (2012).

When the size of the group is sufficiently large, we can construct a prospective mortality table with the intention of identifying the behavior of the insured population that would differ from the regulatory tables or more generally from the national standard. However, in practice the size of the group may be limited and the past experience is observed over a short period. As mentioned in Planchet and Lelieur (2007), two approaches can be proposed to smooth the crude data and project the future mortality using past observations. We distinguish

i. Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality to obtain a smooth surface representing the data correctly, and yield a realistic projection. In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
ii. Models using an external reference mortality table (exogenous approaches) that present a solution to overcome the difficulties associated with having a small volume of data. The idea is to adjust a reference table to the experience of a given set of data.

Considering the limited volume of data available, our attention, in the first step of our methodology, is focused on the second class of models even though comparisons with the first approach are presented.

As a part of the construction of such tables, it is necessary to describe the risks we are facing according to their nature: poolable (hazard on different strata of the population) or systematic (the financial impact can be potentially more important for the insurer or the pension regime). More precisely Planchet and Kamega (2011), similar to Alho (1990), classified the risks into four different but related sources:

i. The risk that can be pooled, originating from random variations of the empirical expectancy around the mathematical expectation due to the sampling variations.

ii. The systematic risk of parameters estimation, originating from a wrong estimation of the model parameters given the sampling variations.

iii. The systematic risk of errors in expert judgment when taking into account external information.

iv. The systematic risk of model due to model misspecification or a change in the trend over time.

The poolable risk, referring to the random character of individual deaths, is not treated here. Extensive studies have discussed the issue of systematic risk of parameter estimation due to the sampling fluctuations. The variance and covariance of parameter estimates are derived either by standard estimation or by bootstrapping, resampling from the original data to create replicate datasets from which the variance and covariance can be estimated. See Planchet and Kamega (2011) for an application of parametric bootstrap.

In this chapter, we focus on the model risk and, to a lesser extent, on the risk of expert judgment related to the choice of the external references used. There is a need for awareness of model risk when assessing longevity-related liabilities, especially for annuities and pensions regimes. The problem is that one can quantify uncertainty within a given model, but one cannot quantify the uncertainty about the model itself. If recent studies, Sibberstein et al. (2008) or Richards and Currie (2009), suggest a rather general analytical framework for the pricing of financial derivatives, the establishment of a standard framework for mortality and longevity models remains to be done. The model risk is particularly difficult to measure, because we cannot measure the uncertainty on a number, as with a price. We have to measure the
uncertainty on a much more complex object, which is the survival distribution. A pragmatic approach is to define a set of possible models and measure the differences on variables of interest.

In our first step, we do not take into account the heterogeneity between the different portfolios. The mortality of the entire male population is not specific to any male subpopulation. The second step of our approach is then to build entity specific male prospective mortality tables by adjusting the reference table validated in the first step to the mortality of each male portfolio. For this purpose, we use a Poisson generalized linear model including interactions with age and calendar year.

The chapter is organized as follows. Section 5.2 specifies the notation and assumptions used in this chapter. It also briefly describes the data and presents our approach to construct specific prospective mortality tables. Section 5.3 covers the extrapolation method for the surfaces obtained by local likelihood smoothing. The extrapolation is performed by identifying the mortality components and their importance over time using functional data analysis. Time series methods are used to extrapolate the time-varying coefficients. The construction of a global prospective reference table for the male population is illustrated in Section 5.4 with the assessment of model risk. Section 5.5 illustrates the construction of entity specific prospective tables. A Poisson GLM including interactions with age and calendar year gives a solution to this problem. Finally, Section 5.6 summarizes the conclusions drawn in the chapter.

5.2 Notation, assumption, data and approach

5.2.1 Notation

We analyze the changes in mortality as a function of both the attained age $x$ and the calendar year $t$. The force of mortality at attained age $x$ for the calendar year $t$, is denoted by $\varphi_x(t)$. We denote $D_{x,t}$ the number of deaths recorded at attained age $x$ during calendar year $t$ from an exposure-to-risk $E_{x,t}$ that measures the time during individuals are exposed to the risk of dying. It is the total time lived by these individuals during the period of observation.

5.2.2 Piecewise constant forces of mortality

We assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next, $\varphi_{x+\tau}(t+\xi) = \varphi_x(t)$ for $0 \leq \tau < 1$ and $0 \leq \xi < 1$. 
We denote by \( p_x(t) \) the probability that an individual aged \( x \) in calendar year \( t \) reaches age \( x+1 \), and by \( q_x(t) = 1 - p_x(t) \) the corresponding probability of death. The expected remaining lifetime of an individual reaching age \( x \) during calendar year \( t \) is denoted by \( e_x(t) \). Under the assumption of piecewise constant forces of mortality, we have for integer age \( x \) and calendar year \( t \),

\[
p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)),
\]

\[
e_x(t) = \frac{1 - \exp(-\varphi_x(t))}{\varphi_x(t)} + \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \exp(-\varphi_{x+j}(t+j)) \right\} \frac{1 - \exp(-\varphi_{x+k}(t+k))}{\varphi_{x+k}(t+k)}.
\]

### 5.2.3 The data

Data are originating from 8 portfolios of various French insurance companies, denoted P1, P2, ..., P8. Table 5.1 displays the observed statistics of the male data.

<table>
<thead>
<tr>
<th>Portfolios</th>
<th>Mean Age In</th>
<th>Mean Age Out</th>
<th>Mean Expo</th>
<th>Mean Age at death</th>
<th>Beginning</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>38.36</td>
<td>43.42</td>
<td>5.06</td>
<td>53.47</td>
<td>01/01/1996</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P2</td>
<td>44.28</td>
<td>45.76</td>
<td>1.48</td>
<td>51.68</td>
<td>01/01/2005</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P3</td>
<td>43.18</td>
<td>45.44</td>
<td>2.26</td>
<td>76.98</td>
<td>01/07/2004</td>
<td>30/06/2007</td>
</tr>
<tr>
<td>P4</td>
<td>51.43</td>
<td>61.74</td>
<td>10.31</td>
<td>77.92</td>
<td>01/01/1996</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P5</td>
<td>42.48</td>
<td>44.60</td>
<td>2.12</td>
<td>54.42</td>
<td>01/01/2003</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P6</td>
<td>47.42</td>
<td>51.15</td>
<td>3.73</td>
<td>71.84</td>
<td>01/01/1996</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P7</td>
<td>55.77</td>
<td>56.78</td>
<td>1.01</td>
<td>72.44</td>
<td>01/01/2006</td>
<td>31/12/2007</td>
</tr>
<tr>
<td>P8</td>
<td>53.65</td>
<td>55.06</td>
<td>1.41</td>
<td>62.28</td>
<td>01/01/2005</td>
<td>31/12/2007</td>
</tr>
</tbody>
</table>

Table 5.1: Observed statistics by portfolios.

The second column of Table 5.1 displays the mean age at entrance of the period of observation for the male population while the third column is the mean age at exit. The fourth column is the average exposure to the risk, and the fifth presents the average age at death. The period of observation of each portfolio is displayed in the sixth and seventh columns. It illustrates that we are facing two difficulties. On one hand the period of observation is small, spreading over 12 years. On the other hand, the structure of the heterogeneity is changing over time as the portfolios are not observed during the same period. Figure 5.1 displays the difference between the portfolios.

From Figure 5.1, the differences in structure by age between the portfolios are apparent. Portfolios P3, P6 and P7, top left corner, are marked by a
relative low average exposure and high average age at death, while portfolios P1, P2 and P5 have a lower average age at death. But a low average age at death does not necessarily mean a higher mortality because observations are censored and truncated.

### 5.2.4 The approach

With the notation of Section 5.2.2 and under the assumption of a piecewise constant force of mortality, the likelihood becomes

\[
L(\varphi_x(t)) = \exp(-E_{x,t} \varphi_x(t))(\varphi_x(t))^{D_{x,t}}.
\]

The associated log-likelihood is

\[
\ell(\varphi_x(t)) = \log L(\varphi_x(t)) = -E_{x,t} \varphi_x(t) + D_{x,t} \log \varphi_x(t).
\]

Similarly to Section 4.4.1, maximizing the log-likelihood \(\ell(\varphi_x(t))\) gives \(\hat{\varphi}_x(t) = D_{x,t}/E_{x,t}\) which coincides with the central death rates \(\hat{m}_x(t)\). Then it is apparent that the likelihood \(\ell(\varphi_x(t))\) is proportional to the Poisson likelihood based on

\[
D_{x,t} \sim \text{Poisson}(E_{x,t}\varphi_x(t)), \quad (5.1)
\]

and it is equivalent to work on the basis of the true likelihood or on the basis of the Poisson likelihood, as recalled in Gschlössl et al. (2011). Thus, under the assumption of constant forces of mortality between non-integer values of \(x\) and \(t\), we consider (5.1) to take advantage of the Generalized Linear Models (GLMs) framework.
Our approach to construct entity specific mortality tables is in two steps. From our collection of portfolios originating from several insurance companies, the first step consists in constructing global prospective mortality tables by gender. For clarity, only results about the male population are presented. By reasoning globally, the male mortality table summarizes the mortality experience of the male portfolios. For this purpose, following Hyndman and Ullah (2007) and Hyndman and Booth (2008), we use principal component analysis (PCA) of functional data combined with a preliminary smoothing, and fit time series models to each component coefficient to obtain forecasts of the forces of mortality. The preliminary smoothing reduces some of the inherent randomness in the observed data. For this purpose we compare the following models described in Table 5.2.

<table>
<thead>
<tr>
<th>Model</th>
<th>Formula</th>
<th>Ref. table</th>
<th>Local lik.</th>
<th>Max. lik.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>[ D_{x,t} \sim \text{Poisson} \left( E_{x,t} \exp(f_1(x,t)) \right) ]</td>
<td></td>
<td>M1</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>[ D_{x,t} \sim \text{Poisson} \left( E_{x,t} \exp(f_2(\log(\phi_{x,t}^\text{ref}(t)))) \right) ]</td>
<td>INSEE</td>
<td>M2.A</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>[ D_{x,t} \sim \text{Poisson} \left( E_{x,t} \phi_{x,t}^\text{ref}(t) \exp(f_1(x,t)) \right) ]</td>
<td>INSEE</td>
<td>M3.A</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>[ D_{x,t} \sim \text{Poisson} \left( E_{x,t} \exp(f_1(x,t) + f_2(\log(\phi_{x,t}^\text{ref}(t)))) \right) ]</td>
<td>INSEE</td>
<td>M4.A</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>[ \text{logit } \phi_x(t) = \alpha + \beta \text{logit } \phi_x^\text{ref}(t) + \epsilon_{x,t} ]</td>
<td>INSEE</td>
<td>M5.A</td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.2**: Description of the models and estimation method used in the first step.

The functions \( f_1 \) and \( f_2 \) are unspecified smooth functions of attained age \( x \) and calendar year \( t \), and forces of mortality according to a reference table \( \phi_x^\text{ref}(t) \), respectively. Model M1 is an endogenous non-parametric approach. Model M2 is an exogenous non-parametric relational model. Models M3 and M4 are mixtures of endogenous and exogenous approaches. Model M3 includes the expected number of deaths \( E_{x,t} \phi_x^\text{ref}(t) \) according to an external reference table. In model M4, \( f_1 \) targets cells for which enough exposure is available (and \( f_2 \approx 0 \)) whereas \( f_2 \) allows to borrow strength from the reference prospective table when the exposure is too limited (\( f_1 \approx 0 \)). Finally model M5 is a semi-parametric Brass-type relational model.

The models M1, M2, M3 and M4 are estimated by non-parametric methods. We considered local kernel-weighted log-likelihood methods presented in Chapter 3 to estimate the smooth functions \( f_1(x,t) \) and \( f_2(\log(\phi_x^\text{ref}(t))) \) for \( x \in [x_1, x_n] \) and \( t = 1, \ldots, m \). The extrapolation, for \( t = m+1, \ldots, m+h \),
relies only on the information contained in the smoothed surface. It is performed by identifying the mortality components and their importance over time using functional PCA. Then time series methods are used to extrapolate the time-varying coefficients. In model M5, the logits of the crude forces of mortality are regressed on the logits of the forces of mortality according to a reference table. The estimation is done by minimizing a weighted distance between the estimated and observed forces of mortality. We refer to Planchet and Théron (2011, Ch.7) for details. Moreover, M5 has the advantage of integrated estimation and forecasting, as the parameters $\alpha$ and $\beta$ are constant.

Finally, we consider Model (6) in Thatcher (1999, p9) to complete the tables until age 120: logit $\varphi_x(t) \approx \log(\alpha_t) + \beta_t x$. It is a robust model that has been found to give good results when fitted to data below age 100 and then extrapolated to higher ages.

We consider two external prospective tables for the first step of our approach as references for the relational models. One is the national demographic projections for the French population over the period 2007-2060, provided by the French National Office for Statistics, INSEE, Blanpain and Chardon (2010). These projections are based on assumptions concerning fertility, mortality and migrations. We choose the baseline scenario among a total of 27 scenarios. The baseline scenario is based on the assumption that until 2060, the total fertility rate is remaining at a very high level (1.95). The decrease in gender-specific and age-specific mortality rates is greater for men over 85 years old. The baseline assumption on migration consists in projecting a constant annual net-migration balance of 100,000 inhabitants. We complete this table by adding the years 1996-2006 from a previous INSEE table. The tables being relatively wiggly, we smoothed the forces of mortality of the completed table using local kernel weighted log-likelihood. The second external reference table, denoted TG05, is a market table built for the entire French market provided by the French Institute of Actuaries, Planchet (2006). Originally the table is generational and covers the period 1900-2005. We adapted it to our needs and to cover the period 1996-2035. It is worth to mention that this table was constructing using mortality trends originating from the INSEE table where a prudence has been added. As a consequence, this table is not fully faithful to the data but incorporates prudence in an arbitrary manner.

In our first step, we do not take into account the heterogeneity between the different portfolios. The mortality of the entire male population is not specific to any male subpopulation. The second step of our approach is then to build entity specific male mortality tables by adjusting the reference table, validated in the first step, to the mortality of each male portfolio. A Poisson GLM including interactions with age and calendar year gives a solution to this problem.
5.3 Extrapolative method

Stochastic methods of mortality forecasting have received considerable attention, see Booth (2006) and Booth and Tickle (2008) for recent reviews. The most widely used are those involving some forms of extrapolation often using time series methods. Extrapolative methods assume that future trends will essentially be a continuation of the past. In mortality forecasting, this is usually a reasonable assumption because of historical regularities. Functional data methods fall into this category, but they have only recently been adopted in mortality forecasting, see Hyndman and Ullah (2007) and Hyndman and Booth (2008).

Lee-Carter or its variants are now the dominant methods of mortality forecasting in actuarial sciences. The Lee-Carter method, Lee and Carter (1992), has a number of advantages, among them simplicity. The Lee-Carter method involves using the first principal component of the log-mortality matrix. In contrast to parametric approaches which specify the functional form of the age pattern of mortality in advance, principal components approaches estimate the age pattern from the data. Improvements to the Lee-Carter estimation basis have been proposed. A Poisson log-likelihood approach has been developed in Brouhns et al. (2002b), Brouhns et al. (2002a) and Renshaw and Haberman (2003) to remedy to some of the drawbacks of the Lee-Carter approach, such as for instance the assumed homoskedasticity of the errors. Cosette et al. (2007) use a binomial maximum likelihood, and a negative binomial version of the Lee-Carter model has been developed by Delwarde et al. (2007) to take into account the over-dispersion phenomenon. The methodology proposed by Hyndman and Ullah (2007) and Hyndman and Booth (2008) can be considered as a successor to the Lee-Carter estimation in that it also involves a principal component decomposition of the mortality surface. However the approach differs in that it uses the functional data paradigm, see Ramsay and Silverman (2005).

Semi-parametric relational models such as M5 have the advantage of integrated estimation and forecasting. This section covers the extrapolation method for the smooth surfaces obtained by local likelihood for models M1, M2, M3 and M4. The extrapolation is performed by identifying the mortality components and their importance over time using functional data analysis, see Ramsay and Silverman (2005, CH.8) and Hyndman and Ullah (2007). Time series methods are used to extrapolate the time-varying coefficients. It can be summarized as follows:

i. Smooth the aggregated data using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion using
the following model:

\[ y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \varepsilon_t(x) \quad \text{with} \quad \varepsilon_t(x) \sim \text{Normal}(0, \nu(x)), \]

where \( y_t(x) = \log \hat{\varphi}_x(t) \), \( \mu_x \) is the mean of \( \log \hat{\varphi}_x(t) \) across years and \( \{\phi_{k,x}\} \) is a set of orthonormal basis functions.

iii. Fit ARIMA models to each of the coefficients \( \{\beta_{t,k}\}, k = 1, \ldots, K \).

iv. Extrapolate the coefficients \( \{\beta_{t,k}\}, k = 1, \ldots, K \), for \( t = m+1, \ldots, m+h \) using the fitted time series models.

v. Use the forecast coefficients with (5.2) to obtain forecasts of \( y_t(x) \), \( t = m+1, \ldots, m+h \), and hence of \( \varphi_x(t) \).

A smoothed version of principal component analysis for functional data is discussed in Silverman (1996). Following the approach of Ramsay and Dalzell (1991) and Hyndman and Ullah (2007), we prefer smoothing the observed data first rather than smoothing the principal component directly to place relevant constraints on the smoothing more easily.

### 5.3.1 Functional principal components analysis

The decomposition using an orthonormal basis (step ii.) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991). In the following, we proceed similarly to Hyndman and Ullah (2007). A more general presentation can be found in Ramsay and Silverman (2005, Ch.8).

We want to find a set of \( K \) orthonormal functions \( \phi_k(x) \) so that the expansion of each curve in terms of the basis functions approximates the curve as closely as possible. For a given \( K \), the optimal orthonormal basis functions \( \{\phi_k(x)\} \) minimize the mean integrated squared error

\[ \text{MISE} = \frac{1}{n} \sum_{t=1}^{m} \int \varepsilon_t^2(x) \, dx \]

This basis set provides informative interpretation and coefficients \( \{\beta_{t,k}\} \) that are uncorrelated, simplifying the forecasting method as multivariate time series models are not required.

The parameter \( \mu(x) \) is estimated as the mean of \( \log \hat{\varphi}_x(t) \) across years. Then we estimate \( \{\beta_{t,k}\} \) and \( \{\phi_k(x)\} \) using a principal components decomposition of \( \hat{y}_t^*(x) = \hat{y}_t(x) - \hat{\mu}(x) \). Our aim is to find the functions \( \phi_k(x) \) that maximize the variance of the scores

\[ z_{t,k} = \int \phi_k(x) \hat{y}_t^*(x) \, dx, \]
subject to the constraints
\[ \int \phi_k^2(x) \, dx = 1 \quad \text{and} \quad \int \phi_k(x) \phi_{k-1}(x) \, dx = 0 \text{ if } k \geq 2. \]

These are defined iteratively for \( k = 1, \ldots, K \) where \( k \leq m - 1 \). The number \( K \) of basis functions depends on many considerations, as explained by Ramsay and Silverman (2005). It depends on the number of discrete points \( m \) in the original data, whether some level of smoothing is imposed by using \( K < m \), on the efficiency of the basis functions in reproducing the behavior of the original functions, and so on. For our application, 12 sampling points are available per curve and actually for these data a value of \( K \) as small as 3 captures most of the interesting variation in the original data.

Assume that we can rewrite each smoothed function \( \hat{y}_t^*(x) \) in an alternative basis expansion
\[ \hat{y}_t^*(x) = \sum_{j=1}^{p} a_{t,j} \xi_j(x). \]

We denote \( A \) the \( m \times p \) matrix of the coefficients \( a_{t,j} \). Let \( J \) be a \( p \times p \) matrix with \( (i,k) \)th element \( J_{ik} = \int \xi_i(x)\xi_k(x) \, dx \). We find the Choleski decomposition \( J = U^T U \) and define
\[ \phi_k(x) = \left( U^{-1} g^{(k)} \right)^T \xi(x), \]

where \( g^{(k)} \) is the \( k \)th normalized eigenvector of \( (U^{-1})^T J S J^T U^{-1} \), \( S = (m - 1)^{-1} A^T A \) and \( \xi(x) = (\xi_1(x), \ldots, \xi_p(x))^T \). Now, if \( \Phi \) denotes an \( n \times (m - 1) \) matrix with \( (i,k) \)th element \( \phi_k(x_i) \), and \( Y \) is a \( m \times n \) matrix with \( (t,i) \)th element \( \hat{y}_t^*(x) \), then \( \hat{\beta}_{t,k} \) is the \( (t,k) \)th element of \( \beta = Y \Phi \).

This procedure is a simplified version of the approach presented in Hyndman and Ullah (2007). In addition, the authors propose a robust method to avoid difficulties with outlying years. For the presentation of their approach, we refer to the mentioned article.

5.3.2 Extrapolation of the time-varying coefficients

The estimated \( \hat{\beta}_{t,k} \)'s can be extrapolated using Box-Jenkins time series methods. The Box-Jenkins approach is one of the most powerful forecasting techniques available and it can be tailored to analyze almost any set of data.

We need to forecast \( \hat{\beta}_{t,k} \) for \( k = 1, \ldots, K \) and \( t = m + 1, \ldots, m + h \). For \( K > 1 \) this is a multivariate time series problem. However, as mentioned previously, because of the way the basis functions \( \phi_k(x) \) have been chosen, the coefficients \( \hat{\beta}_{t,k} \) and \( \hat{\beta}_{t,l} \) are uncorrelated for \( k \neq l \). As a consequence, univariate time series methods are adequate for forecasting each series \( \{ \hat{\beta}_{t,k} \} \).

It is expressed through the development of an ARIMA\((p,d,q)\) model where \( p, d, \) and \( q \) are integers, greater than or equal to zero and refer to the order of the autoregressive, integrated and moving average parts of the model.
Given the time series $\{\hat{\beta}_{t,k}\}$, where $t$ is an integer index, an ARIMA $(p,d,q)$ model is described by

$$(1 - B)^d \phi(B) \hat{\beta}_{t,k} = \theta(B) Z_t, \text{ and } \{Z_t\} \sim \text{White Noise} (\sigma^2), \quad (5.3)$$

where $B$ is the backshift operator, $B \hat{\beta}_{t,k} = \hat{\beta}_{t-1,k}$, expressing the length of the previous data that the model uses to provide the forecasts, and $\phi()$ and $\theta()$ are polynomials of degrees $p$ and $q$ respectively. The parameter $d$ controls the level of differencing. If $d = 0$, the ARIMA is equivalent to an ARMA model. If $d \geq 1$, we can add an arbitrary polynomial trend of degree $(d - 1)$ to $\{\hat{\beta}_{t,k}\}$, without violating the difference equation (5.3). Therefore, ARIMA models are useful for representing data with trend. The AR stands for autoregressive and describes a stochastic process that can be described by a weighted sum of its previous values and a white noise error, while MA stands for moving average and describes a stochastic process that can be described by a weighted sum of a white noise error and the white noise error from the previous periods.

We consider a full range of ARIMA$(p,d,q)$ models with $d = 0, 1, 2$ and $p, q = 0, 1, 2, 3, 4$ as candidates for the period effects. The Bayes information criterion (BIC) is calculated for each ARIMA model and, on the basis of this information, the parameters $p$, $d$ and $q$ are selected. We refer the reader to Brockwell and Davis (2002) for a useful theoretical introduction to time series methods and to Delwarde and Denuit (2003) for an exhaustive application to the Lee-Carter model.

Then, extrapolated forces of mortality are derived using estimated $\mu(x)$ and $\Phi$, the set of the basis functions, and extrapolated $\{\hat{\beta}_{t,k}\}$. Then conditioning on the observed data $\mathcal{J} = \{\varphi_t(x_i); t = 1, \ldots, m; i = 1, \ldots, n\}$ and on the set of the basis function $\Phi$, we deduce $h$-step ahead forecasts of $\varphi_{m+h}(x)$

$$\hat{\varphi}_{m,h}(x) = \mathbb{E} [\varphi_{m+h}(x)|\mathcal{J}, \Phi] = \hat{\mu}(x) + \sum_{k=1}^{K} \hat{\beta}_{m,h,k} \hat{\phi}_k(x),$$

where $\hat{\beta}_{m,h,k}$ denotes the $h$-step ahead forecasts of $\beta_{m+h,k}$ using the estimated series $\hat{\beta}_{1,k}, \ldots, \hat{\beta}_{m,k}$. Hyndman and Ullah (2007) and Hyndman and Booth (2008) provide a procedure to approximate the forecast variance. We refer to the mentioned articles for the presentation of their method.

### 5.4 Construction of a global prospective table

From our collection of portfolios originating from several insurance companies presented in Section 5.2.3, the first step consists in constructing global prospective mortality tables by gender.
By reasoning globally, these tables summarize the mortality experience of these portfolios. We focus on the measurements of the forces of mortality as a function of the attained age $x$ and the calendar year $t$.

### 5.4.1 The aggregated data

We aggregate the portfolios by attained age $x$ and calendar year $t$. The range of attained ages is 30-90 and the observations cover the period 01/01-1996-31/12/2007. Figures 5.2 displays the observed statistics of the aggregated datasets for the male population.

![Figure 5.2: Observed surfaces of the aggregated datasets, male population.](image)

For years 1996 to 2002 only portfolios P1, P4 and P6 are contributing to the surface. After the year 2002, we observe an increase of the number of deaths and exposures due to the aggregation of the other portfolios. As a consequence, the structure of the heterogeneity is changing over time. It may impact the estimation of the mortality trend over the years and ideally we should have stuck to the same structure of the heterogeneity. By aggregating the portfolios, we are therefore making a trade-off between the constitution of a relatively long history and a situation where the structure of the heterogeneity would be stable.

### 5.4.2 Comparisons of the fits

We fitted the models presented in Table 5.2. Figure 5.3 displays the fits in the log scale for the 9 models over the years for several ages. It gives us the opportunity to visualize the similarities and differences between the smoothed surfaces obtained by the models.

Figure 5.3a shows the smoothed fits at attained age 30. It is apparent that the relational models using the table TG05 as reference lead to higher forces of mortality at this age while the models using the national population table originating from INSEE produce smoothed fits in the neighborhood of the endogenous model M1.

In Figure 5.3b, the decreasing trend of the forces of mortality over the
years is sharper at age 40 for models using the national population table. Moreover, compared to the smoothed fits at attained age 30, we observe that the models using the national population table lead to higher forces of mortality than the models having the table TG05 as reference. This fact remains true for ages 50, Figure 5.3c, and 60, Figure 5.3d.

At attained age 50, Figure 5.3c, the fully exogenous non-parametric models M2 and semi-parametric models M5 lead to similar graduation when using the market national population table. Similar remarks can be made with respect to the reference table used, for the models M3 and M4, being mixtures between endogenous and exogenous modeling. Because the models having an exogenous component rely on the general shape of the reference table, the decreasing trend observed for models M2, M3, M4 and M5 is mostly linear. But for the fully endogenous model M1, we observe a non-linear trend.

In Figures 5.3d, 5.3e and 5.3f, the decreasing trend of the forces of mortality is steeper for the fully exogenous models M2 and M5 than models having an endogenous component.

We observe that the models have the following features in common. The overall level of mortality has been declining over time and these improvements have been greater at lower ages than at higher ages. However the models diverge in the speed of the improvement. The fully exogenous models M2 and M5 estimate a steeper decrease of the forces of mortality than models M1, M3 and M4 using an endogenous component. The models using a mixture of endogenous and exogenous modeling M3 and M4 behave similarly with respect to the reference table used. At the extreme ages, the models using the market table lead to higher estimated forces of mortality, while for ages in the center, the models using the national population table yield higher estimates. It gives us a first indication of the degree of model risk. In the following section, these visual comparisons are supplemented by a range of quantitative diagnostics which will increase our confidence in some models and question the suitability of others for our purposes.

5.4.3 Tests and quantities to compare graduations

We now carry out a number of tests to assess the impact of model choice. We apply the tests proposed by Forfar et al. (1988, p.56-58) and Debón et al. (2006, p.231). We have also obtained the values of the mean absolute percentage error MAPE and $R^2$ used in Felipe et al. (2002). In addition, we compute the relative difference between the observed number of deaths and the expected number of deaths obtained by the models $SMR - 1$, where the standardized mortality ratio (SMR) is defined by

$$SMR = \frac{\sum_{(x,t)} E_{x,t} \hat{\varphi}_x(t)}{\sum_{(x,t)} E_{x,t} \hat{\varphi}_{x\text{ref}}(t)} = \frac{\sum_{(x,t)} D_{x,t}}{\sum_{(x,t)} E_{x,t} \hat{\varphi}_{x\text{ref}}(t)}, \quad (5.4)$$
Figure 5.3: Comparisons of the Fits for several attained ages, log scale, male population.
for \((x, t)\) in the set of ages and calendar years of interest. We compare the crude forces of mortality rates to the graduated series to see whether the approaches lead to similar graduation. Table 5.3 presents the results.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitted DF</td>
<td>7.56</td>
<td>4.02</td>
<td>3.94</td>
<td>4.88</td>
<td>4.88</td>
<td>4</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Deviance</td>
<td>1302.53</td>
<td>1302.83</td>
<td>1328.04</td>
<td>1296.58</td>
<td>1351.66</td>
<td>1313.81</td>
<td>1366.89</td>
<td>1355.92</td>
<td>1417.08</td>
</tr>
<tr>
<td>Standardised residuals</td>
<td>&gt; 2</td>
<td>86</td>
<td>84</td>
<td>92</td>
<td>86</td>
<td>93</td>
<td>87</td>
<td>94</td>
<td>86</td>
</tr>
<tr>
<td>&gt; 3</td>
<td>24</td>
<td>22</td>
<td>24</td>
<td>23</td>
<td>24</td>
<td>23</td>
<td>26</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Signs</td>
<td>+(-)</td>
<td>327(405)</td>
<td>319(411)</td>
<td>336(396)</td>
<td>345(387)</td>
<td>343(389)</td>
<td>334(398)</td>
<td>337(395)</td>
<td>333(399)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0043</td>
<td>0.0005</td>
<td>0.0291</td>
<td>0.1296</td>
<td>0.0962</td>
<td>0.0198</td>
<td>0.0350</td>
<td>0.0162</td>
<td>0.9117</td>
</tr>
<tr>
<td>Runs</td>
<td>Nb of runs</td>
<td>346</td>
<td>326</td>
<td>322</td>
<td>328</td>
<td>302</td>
<td>332</td>
<td>306</td>
<td>334</td>
</tr>
<tr>
<td>test</td>
<td>Value</td>
<td>−1.55</td>
<td>−3.00</td>
<td>−3.46</td>
<td>−3.17</td>
<td>−5.02</td>
<td>−2.77</td>
<td>−4.82</td>
<td>−2.61</td>
</tr>
<tr>
<td>p-value</td>
<td>0.1188</td>
<td>0.0026</td>
<td>5.28e−4</td>
<td>0.0014</td>
<td>5.25e−7</td>
<td>0.0056</td>
<td>1.4e−6</td>
<td>0.0089</td>
<td>8.86e−5</td>
</tr>
<tr>
<td>Kolmogorov Smirnov test</td>
<td>Value</td>
<td>0.0327</td>
<td>0.0286</td>
<td>0.0601</td>
<td>0.0286</td>
<td>0.0286</td>
<td>0.0014</td>
<td>0.0300</td>
<td>0.0063</td>
</tr>
<tr>
<td>p-value</td>
<td>0.8262</td>
<td>0.9239</td>
<td>0.1419</td>
<td>0.9239</td>
<td>0.0978</td>
<td>0.9239</td>
<td>0.1257</td>
<td>0.8555</td>
<td>0.0657</td>
</tr>
<tr>
<td>(\chi^2)</td>
<td>1400.19</td>
<td>1402.53</td>
<td>1418.89</td>
<td>1405.81</td>
<td>1445.53</td>
<td>1421.85</td>
<td>1466.93</td>
<td>1473.63</td>
<td>1545.42</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.9326</td>
<td>0.9256</td>
<td>0.9325</td>
<td>0.9312</td>
<td>0.9325</td>
<td>0.9302</td>
<td>0.9006</td>
<td>0.9221</td>
<td>0.9006</td>
</tr>
<tr>
<td>SMR − 1 (%)</td>
<td>−0.79</td>
<td>0.29</td>
<td>0.21</td>
<td>0.11</td>
<td>0.30</td>
<td>1.22e−12</td>
<td>−7.77e−13</td>
<td>1.87</td>
<td>2.87</td>
</tr>
</tbody>
</table>

Table 5.3: Comparisons between the smoothing approaches.

The approaches display different results. Model M1, having the highest degrees of freedom and being fully endogenous, has the capacity to reveal many features in the data. Therefore, it has the highest number of runs, lowest \(\chi^2\) and MAPE and highest \(R^2\). We observe that M1 is the only model to lead to a higher number of expected deaths than observed. Conversely, the fully exogenous semi-parametric models M5, and to a lesser degree the non-parametric M2, lead to higher deviance, higher \(\chi^2\), lower \(R^2\), higher number of standardized residuals exceeding the thresholds 2 and 3 and higher relative difference between expected and observed number of deaths.

We observe that the fully exogenous models M2 and M5 do not behave similarly. The non-parametric models M2, being more flexible, perform better than the semi-parametric models M5. With respect to the reference table used, models M2 have a lower deviance, lower number of standardized residuals exceeding the thresholds 2 and 3, lower \(\chi^2\) and MAPE, and higher \(R^2\). Also, the expected and observed number of deaths are closer.

The mixtures of endogenous and exogenous modeling M3 and M4 have similar results with respect to the reference table used. Nevertheless, models M3, including the expected number of deaths according to the reference table INSEE or TG05, perform better than models M4. Models M3 have a better spread of the residuals between positive and negative signs, higher value for the sign test, lower deviance, \(\chi^2\) and MAPE. However, models M4
have the smallest relative difference between expected and observed number of deaths.

We observe that, in general, models incorporating the national population table originating from INSEE (models A) produce graduations that are closer to the data than models using the market table TG05 (models B) as reference. Using the market table leads to higher deviance, higher $\chi^2$, lower number of runs and higher number of standardized residuals exceeding the thresholds 2 and 3 compared to models incorporating the national population table. The market table TG05 is derived on mortality trends originating from the INSEE table where a prudence has been added. As a consequence, this table is not fully faithful to the data but incorporates prudence in an arbitrary manner.

The tests and quantities carried out in Table 5.3 show the strengths and weaknesses of each model to adjust the observed mortality. The choice between the models is only a matter of judgment and depends on the purpose for which the prospective mortality table would be used. It is up to potential users of the table to decide the weights they place on the different criteria. However, regarding the wide ranging set of model selection criteria, we can eliminate some models. We have seen that the non-parametric models, due to their flexibility, ensure a good fit. Hence models M2 would be preferred to M5. Within the mixture of endogenous and exogenous models, M3 would be preferred to M4. Compared to the fully endogenous model and to the fully exogenous models, relying partly on the national population table is beneficial according to the various tests and quantities used in assessing the adjustment of observed mortality, hence model M3.A would be preferred to models M1 and M2.

5.4.4 Extrapolation of the smoothed surfaces and completed tables

The extrapolation of the smoothed surface of models M1, M2, M3 and M4 is performed by identifying the mortality components and their importance over time using functional principal component analysis presented in Section 5.3.1. Then time series methods are used to extrapolate the time-varying coefficients. Model M5 has the advantage of integrated estimation and forecasting.

Figure 5.4 displays the basis functions and associated coefficients using (5.2) for the models M1, M2, M3 and M4. A decomposition of order $K = 3$ has been used.

The average log-mortality at attained ages is similar for the models over time except at the extreme ages, Figure 5.4a. Models using the market mortality table TG05 as reference lead to higher mortality around age 30 compared to models using the national population mortality table, as observed previously in Figure 5.3a.
Figure 5.4: Basis functions and associated coefficients with $K = 3$ for models $M_1, M_2, M_3$, and $M_4$, male population.
Figure 5.4b shows the first basis function for all models. The first term accounts for at least 99.6% of the variation in mortality. The coefficient, Figure 5.4e, indicates a fairly steady decline in mortality over time. The models lead to more or less the same results except for model M2.A that gives the steepest decrease. The models M3 and M4 using a mixture of endogenous and exogenous modeling produce similar results with respect to the reference table used. The basis function $\phi_1(x)$ indicates that the decline has been faster for the young adults and at ages 60–80 for the models using the national population table originating from INSEE (models A) as well as a fully endogenous model M1. But for models using the market table TG05 (models B) the decrease has been steady for ages 30–80. We observe that models M3 lead to the fastest and slowest improvement of the mortality for the young adults and individuals above 80, respectively.

The basis function $\phi_2(x)$, displayed in Figure 5.4c, models the differences between the young adults and those over 75. The coefficients in Figure 5.4f shows that this difference in mortality has falling from the beginning of the period of investigation to 2002 - starting date of observation of additional portfolios - and increasing since 2002 to the end of the period of investigation.

Similarly, Figure 5.4d displays difference between the young adults (up to 50) and those over 80. However, the shape of associated coefficient Figure 5.4g is more irregular than $\beta_{t,2}$. Again we observed that the choice of the reference table used leads to a different pattern of the basis functions and associated coefficients.

The time-varying coefficients are forecast using univariate time series methods. Table 5.4 summarizes the ARIMA models, introduced in Section 5.3.2. For each of the models M1, M2, M3 and M4, we considered a full range of ARIMA($p,d,q$) models with $d = 0, 1, 2$ and $p, q = 0, 1, 2, 3, 4$ as candidates for the period effects. The Bayes information criterion (BIC) was calculated for each ARIMA model and, on the basis of this information, the parameters $p, d$ and $q$ have been selected. Figure 5.5 displays the resulting projections for models M1, M2, M3 and M4 for $h = 28$, that is until year 2035. For clarity, the confidence intervals are omitted.

We notice that the coefficients $\tilde{\beta}_{m,h,2}$ and $\tilde{\beta}_{m,h,3}$ in Figures 5.5b and 5.5c are rapidly constant. As a consequence, we could have performed a decomposition using the first principal component as in the original Lee-Carter method. However, it may not be the case for other datasets, as illustrated in Hyndman and Ullah (2007) and Hyndman and Booth (2008). The use of several components is the main difference between this approach and the Lee-Carter method, which uses only the first component and also involves an adjustment. The extra principal components allow more accurate forecasting of age-specific forces of mortality, though in our application at least 99.6% of the variation is explained by the first component.
Table 5.4: Description of the models for the time-varying coefficients, male population.

<table>
<thead>
<tr>
<th>Model</th>
<th>Component</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>( k=1 )</td>
<td>ARIMA(1,2,1) with zero mean</td>
</tr>
<tr>
<td></td>
<td>( k=2 )</td>
<td>ARIMA(0,0,0) with non-zero mean</td>
</tr>
<tr>
<td></td>
<td>( k=3 )</td>
<td>ARIMA(2,0,1) with zero mean</td>
</tr>
<tr>
<td>M2.A &amp; M4.A</td>
<td>( k=1 )</td>
<td>ARIMA(0,1,0) with drift</td>
</tr>
<tr>
<td></td>
<td>( k=2 ), 3</td>
<td>ARIMA(1,0,0) with zero mean</td>
</tr>
<tr>
<td></td>
<td>( k=1 )</td>
<td>ARIMA(1,1,0) with drift</td>
</tr>
<tr>
<td></td>
<td>( k=2 )</td>
<td>ARIMA(0,0,0) with zero mean</td>
</tr>
<tr>
<td>M3.A</td>
<td>( k=1 )</td>
<td>ARIMA(0,2,0)</td>
</tr>
<tr>
<td></td>
<td>( k=2 )</td>
<td>ARIMA(2,0,0) with zero mean</td>
</tr>
<tr>
<td></td>
<td>( k=3 )</td>
<td>ARIMA(0,0,2) with zero mean</td>
</tr>
<tr>
<td>M3.B</td>
<td>( k=1 )</td>
<td>ARIMA(1,2,0)</td>
</tr>
<tr>
<td></td>
<td>( k=2 ), 3</td>
<td>ARIMA(0,0,0) with zero mean</td>
</tr>
</tbody>
</table>

Model \& component
Figure 5.5: Projections of the estimated coefficients $\beta_{t,k}$ for the models M1, M2, M3 and M4 obtained by ARIMA, and estimated regression parameters of model 6 of Thatcher (1999), male population.
The next step is to obtain completed tables until age 120. For this, we apply Model (6) in Thatcher (1999, p9) to the forces of mortality to extrapolate the data: \( \text{logit} \varphi_x(t) \approx \log(\alpha_t) + \beta_t \times x \). It is a robust model that has been found to give good results when fitted to data below age 100 and then extrapolated to higher ages. Figures 5.5d and 5.5e show the estimated regression parameters \( \alpha_t \) and \( \beta_t \), respectively. All models estimate a linear effect of time on the forces of mortality at high ages, Figure 5.5e. We observe that models using the national population table (models A) lead to a steeper increase of the linear component \( \beta_t \) over time, Figure 5.5e, than models using the market table as reference (models B). As a consequence, those models lead to a more rapid increase of the forces of mortality at the highest ages (70-90), which in turn results in a more rapid decrease of the forces of mortality at lower ages. The mixture of endogenous and exogenous modeling models M4 and fully exogenous semi-parametric model M5 produce very similar results, while the fully endogenous model M1 and fully exogenous model M2.A differ largely from the other models.

Figure 5.6 displays the fits in the log scale for the 9 models over the years for several ages. For clarity, the confidence intervals are omitted. The forecasts produced here are based on the first three principal components. The additional components may serve to incorporate relatively recent changes in pattern. The use of smoothing prior to modeling results in forecast age patterns that are relatively smooth.

As visualized in Figure 5.3, the overall level of mortality is declining over time and these improvements are greater at lower ages than at higher ages. However the models diverge in the level and speed of the improvement. At the extreme ages, the models using the market table (models B) lead to higher estimated forces of mortality, while for ages in the center, the models using the national population table (models A) yield higher estimates. The fully exogenous models M2 and M5 produce a steeper decrease of the forces of mortality than models M3 and M4. Model M1 stands out, leading to a non-linear decline of the forces of mortality and inducing the sharpest decrease.

5.4.5 Model risk and validation of the final table

We have seen in Figure 5.6 that models diverge in the level and speed of the improvement of the level of mortality across the age. It gives us a first indication of the degree of model risk. Figure 5.7 shows the survival indexes at several ages computed from the completed tables obtained with the different models. It represents the survival indexes of cohorts aged 30, 40, 50, 60, 70 and 80 in 1996 over 40 years. This measures the proportion from a group of males aged 30, 40, 50, 60, 70 or 80 at the start of 1996 who remain alive for the next 40 years.
<table>
<thead>
<tr>
<th>(a) Attained age</th>
<th>(b) Attained age</th>
<th>(c) Attained age</th>
<th>(d) Attained age</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>(e) Attained age</td>
<td>(f) Attained age</td>
<td>(g) Attained age</td>
<td>(h) Attained age</td>
</tr>
<tr>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

**Figure 5.6:** Comparisons of the fits and forecasts for several attained ages, log scale, male population.
It can be seen that these survival indexes are affected by the choice of the modeling, endogenous, exogenous or mixture of the two and to a lesser degree by the choice of the reference table used. Endogenous model M1, and exogenous models M2 and M5 with respect to the reference table used lead to higher survival indexes for cohorts aged 30, 40 and 50 in 1996, Figures 5.7a, 5.7b and 5.7c. For a cohort aged 60, 70 and 80 in 1996, Figures 5.7d, 5.7e and 5.7f, the survival indexes are consistent within the models.

We observe substantial differences in using the market table TG05 or the national population table as reference in the exogenous or mixture models. For cohort aged 30 in 1996, Figure 5.7a, the incorporation of the market table (models B) leads to a higher survival index for the models using of the national population table (models A). Conversely, when incorporating the national population table for cohort aged 40 and 50 in 1996, in Figures 5.7b and 5.7c the survival indexes are higher with respect to the models.

As a second example, we calculate some single figures summarizing the lifetime probability distribution for cohorts at several ages in 1996. Table 5.5 displays the indices.

The mixtures of endogenous and exogenous modeling, models M3 and M4, lead to the smallest partial life expectancies $40e_{30}$, $40e_{40}$, $40e_{50}$ and $40e_{60}$ for cohorts aged 30, 40, 50 and 60 in 1996. The fully endogenous model M1 yields the highest partial life expectancies $40e_{30}$ and $40e_{40}$ but leads to the smallest for cohorts aged 80 in 1996.

The semi-parametric models M5 produce higher partial life expectancies than the non-parametric models M2 except for $40e_{50}$ and $40e_{60}$. Similarly, the mixture models M4 yield higher partial life expectancies than models M3 incorporating the expected number of deaths according to a reference table except for $40e_{80}$.

We observe, once more, that the choice of the reference table affects the quantities. Using the national population table leads to higher life expectancy than incorporating the market table.

These results can be seen in the median age at death, Med$(40T)$. The exogenous models M2 and M5 produce close estimates, and the mixture models M3 and M4 lead to more or less similar results.

The mixture M3 and M4 models stand out as having a much higher standard deviation of the random life time, $40\sigma$, than the exogenous model, which would be expected. However it suggests that model risk might be an issue. For example, the price of a financial option that has the survival index as its underlying quantity is strongly dependent on its standard deviation; everything else being equal, the higher the variance, the higher the value of the option, as recalled in Cairns et al. (2009).

The entropy $H(40T)$ obtained with the exogenous models M2 and M5 is similar, also there is not much difference in $H(40T)$ for the mixture M3 and M4 models. However, we notice that using the market table leads to
Figure 5.7: Survival indices for cohorts at several ages at the start of 1996 over 40 years, male population.
To have a clear picture of the contribution of model risk to forecast uncertainty, we can make use of the first two robust principal component scores of quantities of interests such as the partial life expectancies of cohorts at several ages in 1996 with the Highest Density Regions (HDR) boxplots of Hyndman (1996). Hyndman and Shang (2010) have proposed this method with identification of outliers in mind. Our idea is to use this graphical method on single figure indices summarizing the lifetime probability distributions, such as the partial life expectancies for cohorts at several ages in 1996, to visualize similarity between the models and outliers and thus model risk in forecast uncertainty.
The bivariate HDR boxplot displays the mode, the highest density point, along with the 50% inner and 99% outer highest density regions. All points excluded from the outer HDR are outliers. Figure 5.8 displays the bivariate HDR boxplot of the first two robust principal component scores of the partial life expectancies for cohorts at several ages in 1996.

![Figure 5.8: Bivariate HDR boxplot of the first two robust principal component scores of the partial life expectancies for cohorts at several ages in 1996, male population.](image)

The dark and light gray regions show the 50% HDR and the outer HDR, respectively. The points outside the outer regions are identified as outliers, as model M5.A. The asterisk in Figure 5.8 marks the mode of the bivariate robust principal component scores, corresponding to model M2.B. It shows clearly that the non-parametric models are grouped more by the reference table used and less by the kind of modeling (non-parametric, semi-parametric, endogenous, exogenous and so on).

We have concentrated here on the contribution of model risk in extrapolating the future mortality. However, it is appropriate to allow for parameter uncertainty to provide a more complete picture of the level of risk on the valuations of an insurer, such as provisioning and capital requirement.

The overall model risk associated with a prospective mortality table should ideally take into account two factors,

i. the adjustment according to the past mortality, and

ii. the extrapolation of the future mortality.
From Section 5.4.3, we can eliminate some models, regarding the wide ranging set of model selection criteria. We have seen that the non-parametric models, due to their flexibility, ensure a good fit. Hence models M2 would be preferred to M5. Within the mixture of endogenous and exogenous models, M3 would be preferred to M4. Compared to the fully endogenous model and to the fully exogenous models, relying partly on the national population table is beneficial according to the various tests and quantities used in assessing the adjustment of observed mortality in Table 5.3, hence model M3.A would be preferred to models M1 and M2.

This choice could be refined by analyzing the extrapolated future mortality. We can apply the concept of biological reasonableness which was first proposed in Cairns et al. (2006) as an aid in assessing the forecasts. This concept is not based on hard scientific, biological or medical facts. It is rather subjective and asks the question where the data are originating from and based on this knowledge, what mixture of biological factors, medical advances and environmental changes would have to happen to cause this particular set of forecasts. For instance, in Figure 5.6, the projections for model M1, look rather more optimistic than the set of projections of the other models. If we cannot think about any good reason why this might happen, then we must disqualify the model on the basis of biological reasonableness. The projections of Model M3.A seem reasonable, in accordance with the set of projections with the other models. Hence, in the following section we adjust the entity specific portfolio experience to the baseline mortality surface obtained by the mixture of endogenous and exogenous modeling M3.A.

5.5 Adjustment to entity specific mortality experience

5.5.1 Entity specific mortality experience

In our first step, we do not take into account the heterogeneity between the different portfolios. The mortality of the aggregated male population is not specific to any male portfolio. We compare the mortality experiences of the 8 portfolios presented in Table 5.1 to the validated table constructed in the first step. The standardized mortality ratio (SMR), as defined in expression (5.4), appears to be a useful index. The observed deaths in a particular portfolio are compared with those that would be expected if the mortality validated in the first step applied. Table 5.6 displays the SMR of the 9 portfolios with the national population reference table originating from INSEE, the market table TG05 and the validated table obtained by the mixture of endogenous and exogenous modeling M3.A.
<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>mean((\text{SMR} - 1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>INSEE</td>
<td>−36.93</td>
<td>−57.36</td>
<td>−2.03</td>
<td>−21.22</td>
<td>−29.80</td>
<td>−35.41</td>
<td>−47.11</td>
<td>−42.75</td>
<td>34.08</td>
</tr>
<tr>
<td>TC05</td>
<td>52.31</td>
<td>18.09</td>
<td>59.62</td>
<td>34.47</td>
<td>88.43</td>
<td>9.88</td>
<td>0.18</td>
<td>30.39</td>
<td>37.91</td>
</tr>
<tr>
<td>M3.A</td>
<td>10.85</td>
<td>−18.67</td>
<td>32.41</td>
<td>10.73</td>
<td>30.98</td>
<td>−5.25</td>
<td>−22.66</td>
<td>−7.50</td>
<td>17.38</td>
</tr>
</tbody>
</table>

Table 5.6: Relative difference between expected and observed number of deaths by portfolios, \((\text{SMR} - 1\%))\), male population.

Table 5.6 illustrates the heterogeneity between the portfolios. We observe that the table validated in the first step under-estimates the number of deaths for portfolios P1, P3, P4 and P5, while it over-estimates the number of deaths for the other portfolios. It should be noted that the national population table constantly over-estimates the number of deaths, but the market table leads to an under-estimation. The relative difference between the observed and expected number of deaths obtained with model M3.A is similar for portfolios P1 and P4, P3 and P5 and P6 and P8, respectively. Relative differences are smaller when using the national population table for P3, or using the market table for P7. However, on average the validated table originating from M3.A leads to the smallest difference in absolute value, illustrating the usefulness of the first step of our approach.

5.5.2 Poisson GLM with age and calendar year interactions

In a Poisson regression, we include the portfolio dummies as a covariate and allow interactions with age and calendar year. We assume that the number of deaths for a portfolio \(i\) at attained age \(x\) and calendar year \(t\) is determined by

\[
D_{x,t,i} \sim \text{Poisson}(E_{x,t,i} \varphi_x(t,i)),
\]

with

\[
\log \varphi_x(t,i) = \alpha + \beta \log \hat{\varphi}^{ref}_x(t) + \sum_{j=1}^{n} \gamma_j I_i + \sum_{j=1}^{n} \delta_j x I_i + \sum_{j=1}^{n} \kappa_j t I_i + \sum_{j=1}^{n} \lambda_j x t I_i
\]

(5.5)

where \(\hat{\varphi}^{ref}_x(t)\) is the baseline force of mortality derived in our first step, the \(I_i\)’s are binary variables coding the portfolios and \(n\) represents the number of portfolios.

If we do not allow for interactions, we will observe parallel shifts of the forces of mortality according to the baseline mortality for each dimension. This view is certainly unrealistic and interactions need to be incorporated. We take the first portfolio P1 as reference level. The relative mortality of the portfolios is expressed with respect to this reference level P1.

We start by incorporating all interactions. We remove the calendar year effect for portfolios P2, P7 and P8, having less than 4 years of observation.
With a parsimonious principle in mind, we progressively exclude the insignificant interactions by computing the drop in deviance test (or likelihood-ratio test) for models with and without the interaction considered.

The final model is the following

\[
\varphi_x(t, i) = \alpha + \beta \log \hat{\varphi}_x^\text{ref}(t) + \delta_1 x + \kappa_1 t + \lambda_1 x t \\
+ \delta_2 x I_{i=2} + \delta_3 x I_{i=3} + \kappa_2 t I_{i=3} \\
+ \kappa_3 t I_{i=4} \\
+ \kappa_4 t I_{i=5} \\
+ \gamma_1 I_{i=6} + \delta_4 x I_{i=6} + \kappa_5 t I_{i=6} \\
+ \gamma_2 I_{i=7} \\
+ \gamma_3 I_{i=8} + \delta_5 x I_{i=8}.
\]  

(5.6)

The main effects and interactions included in the final model (5.6) are presented in Table 5.7.

<table>
<thead>
<tr>
<th>Regression coef.</th>
<th>Parameter est.</th>
<th>Std. error</th>
<th>z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>136.2</td>
<td>21.54</td>
<td>6.322</td>
<td>2.59e-10</td>
</tr>
<tr>
<td>( \beta )</td>
<td>1.648</td>
<td>9.610e-02</td>
<td>17.153</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>-36.11</td>
<td>9.510</td>
<td>-3.797</td>
<td>0.0001</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>-0.4028</td>
<td>3.910e-02</td>
<td>-10.301</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>0.6307</td>
<td>9.675e-02</td>
<td>6.519</td>
<td>7.08e-11</td>
</tr>
<tr>
<td>( \delta_1 )</td>
<td>-1.783</td>
<td>0.2929</td>
<td>-6.088</td>
<td>1.14e-09</td>
</tr>
<tr>
<td>( \delta_2 )</td>
<td>-2.168e-03</td>
<td>8.188e-04</td>
<td>-2.648</td>
<td>0.008</td>
</tr>
<tr>
<td>( \delta_3 )</td>
<td>-2.585e-02</td>
<td>1.673e-03</td>
<td>-15.452</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>( \delta_4 )</td>
<td>-5.658e-03</td>
<td>1.105e-03</td>
<td>-5.122</td>
<td>3.03e-07</td>
</tr>
<tr>
<td>( \delta_5 )</td>
<td>-1.178e-02</td>
<td>1.466e-03</td>
<td>-8.034</td>
<td>9.45e-16</td>
</tr>
<tr>
<td>( \kappa_1 )</td>
<td>-6.477e-02</td>
<td>1.089e-02</td>
<td>-5.951</td>
<td>2.67e-09</td>
</tr>
<tr>
<td>( \kappa_2 )</td>
<td>1.012e-03</td>
<td>6.345e-05</td>
<td>15.945</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>( \kappa_3 )</td>
<td>-8.897e-05</td>
<td>1.743e-05</td>
<td>-5.104</td>
<td>3.33e-07</td>
</tr>
<tr>
<td>( \kappa_4 )</td>
<td>1.585e-05</td>
<td>1.517e-05</td>
<td>10.445</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>( \kappa_5 )</td>
<td>1.810e-02</td>
<td>4.754e-03</td>
<td>3.807</td>
<td>0.0001</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>8.638e-04</td>
<td>1.467e-04</td>
<td>5.887</td>
<td>3.94e-09</td>
</tr>
</tbody>
</table>

Table 5.7: Results from the Poisson regression model (5.5), male population.
Model (5.5) is estimated over the observation period 1996-2007 and for the age range 30-90. The specific prospective mortality tables are now easily derived by incorporating the entire mortality table $\hat{\varphi}_{x}^{\text{ref}}(t)$ obtained in the first step of our approach. For instance, for portfolio P1, the forces of mortality are given by

$$\hat{\varphi}_{x}(t, 1) = \exp \left( \hat{\alpha} + \hat{\beta} \log \hat{\varphi}_{x}^{\text{ref}}(t) + \hat{\delta}_{1} x + \hat{\kappa}_{1} t + \hat{\lambda}_{1} x t \right),$$

and for portfolio P6,

$$\hat{\varphi}_{x}(t, 6) = \exp \left( \hat{\alpha} + \hat{\gamma}_{1} + \hat{\beta} \log \hat{\varphi}_{x}^{\text{ref}}(t) + (\hat{\delta}_{1} + \hat{\delta}_{4}) x + (\hat{\kappa}_{1} + \hat{\kappa}_{5}) t + \hat{\lambda}_{1} x t \right).$$

We observe that the final model (5.6) only includes the baseline age calendar year mixed effect, meaning that there is no significant difference of the age calendar year mixed effect between the portfolios. Portfolio P2 differs significantly from P1 only by the age pattern of the forces of mortality. The time trends are then similar to P1. Conversely, Portfolios P4 and P5 have similar age pattern but differ significantly from P1 by the time trends. P3 and P6 behave differently than P1 in age and calendar year, while the behavior of P7 is similar and only the overall level of mortality is significantly different. Similarly, the overall level of mortality is significantly different for portfolios P6, and P8. In addition, the age effect is also significant for P8.

The derivation of the portfolio specific prospective tables can sometimes lead to unrealistic estimates at the highest ages for long-term projections. Therefore, in a similar fashion as the reference table obtained in the first step, we apply Model (6) in Thatcher (1999, p9) to the forces of mortality to adapt the data at the highest ages. Figure 5.9 shows the estimated regression parameters $\alpha_{t}$, Figure 5.9a, and $\beta_{t}$, Figure 5.9b.

**Figure 5.9:** Estimated regression parameters of model 6 of Thatcher (1999), male population.
The linear component, Figure 5.9b, is much higher for portfolio P5 indicating that the forces of mortality increase more rapidly than the other portfolios at the highest ages. Conversely, portfolios P3 and P8 have smaller estimated $\beta_t$'s illustrating that those portfolios lead to a less pronounced increase.

Figure 5.10 displays the forces of mortality derived for each portfolio by age and calendar years. Since we have incorporated interactions in the model, we see that the portfolio specific prospective mortality tables show different patterns with age and calendar year. As noted in Table 5.6, P3, P5 and P1 yield the highest mortality experience, while P7 and P2 lead to the lowest mortality experience.

5.6 Summary and outlook

In this chapter, we illustrated the construction and the validation of entity specific prospective mortality tables by a two steps approach. From portfolios of several insurance companies we constructed, in a first step, a global prospective reference table summarizing the mortality experience of these portfolios. We used a non-parametric method, the local kernel-weighted log-likelihood, and semi-parametric relational models to graduate and extrapolate the surfaces. The extrapolations of the smoothed surface, obtained by local likelihood methods, were performed by identifying the mortality components and their importance over time using functional principal components analysis. Then time series methods were used to extrapolate the time-varying coefficients, while semi-parametric relational models had the advantage of integrated estimation and forecasting.

We investigated the divergences in the mortality surfaces generated by a number of proposed models. We found that the model risk is present. The overall model risk associated with a prospective mortality table was assessed by taking into account two factors, the adjustment according to the past mortality and the extrapolation of the future mortality. We have carried out a number of tests to assess the impact of model choices on the adjustment of the past mortality. We find that even for those models satisfying our criteria, there are significant differences among the smoothed forces of mortality at different ages. Moreover, selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

To measure the divergence in the extrapolation of the future mortality, we used single figure indices summarizing the lifetime probability distribution that utilize those forecasts, such as the survivor index, or partial life expectancy (which is, in turn, derived from the survivor index). We visualized those differences by a bivariate HDR boxplot of the first two robust principal
(a) Calendar year 2007

(b) Calendar year 2020

(c) Calendar year 2035

(d) Attained age 30

(e) Attained age 60

(f) Attained age 90

Figure 5.10: Comparisons of the forces of mortality by age and calendar years, log scale, male population.
component scores of the partial life expectancies for cohorts at several ages in 1996.

We found that the models have the following features in common: the overall level of mortality has been declining over time and these improvements have been greater at lower ages than at higher ages. However the models diverge in the level and speed of the improvement.

We therefore need to weigh the strengths and weaknesses of each model to validate the mortality table. It is up to potential users of the table to decide the weights they place on the different criteria. The validation of the mortality table involved many judgmental decisions. It has been driven by the trade-off between how the model smooths the historical data and the concept of biological reasonableness leading us to question the plausibility of the forecasts produced.

Then, we switched our attention to the construction of a portfolio specific prospective mortality table. The validated table is used in a second step to adjust the mortality to each portfolios by a Poisson generalized linear model including age and calendar year interactions. The estimated baseline forces of mortality are used in the regression analysis as if they were known with certainty. This approach has shown to be very simple and convenient in practical applications.

Another approach would be to use a generalized additive model (GAM) with $p$-splines to perform the mortality analysis in a one step approach. A GAM combines both continuous and categorical model components in one model and $p$-splines would have the advantage of integrated estimation and forecasting.