

DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 41, ARTICLE 38, PAGES 1091–1130
PUBLISHED 24 OCTOBER 2019

<http://www.demographic-research.org/Volumes/Vol41/38/>

DOI: 10.4054/DemRes.2019.41.38

Research Article

Smooth constrained mortality forecasting

Carlo G. Camarda

© 2019 Carlo G. Camarda.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.

See <https://creativecommons.org/licenses/by/3.0/de/legalcode>

Contents

1	Introduction	1092
2	<i>P</i> -splines for mortality data	1095
2.1	Forecasting with <i>P</i> -splines	1099
3	The <i>CP</i> -spline model	1102
3.1	Addressing infant mortality	1102
3.2	Enforcing mortality patterns over age and time	1104
3.2.1	Incorporating prior knowledge into the model	1110
3.2.2	Confidence intervals by bootstrap	1111
4	Application	1113
5	Conclusions	1118
6	Acknowledgments	1121
	References	1122

Smooth constrained mortality forecasting

Carlo G. Camarda¹

Abstract

BACKGROUND

Mortality can be forecast by means of parametric models, principal component methods, and smoothing approaches. These methods either impose rigid modeling structures or produce implausible outcomes.

OBJECTIVE

We propose a novel approach for forecasting mortality that combines a well established smoothing model and prior demographic information. We constrain future smooth mortality patterns to lie within a range of valid age profiles and time trends, both computed from observed patterns.

METHODS

Within a *P*-spline framework, we enforce shape constraints through an asymmetric penalty approach on forecast mortality. Moreover, we properly integrate infant mortality in a smoothing framework so that the mortality forecast covers the whole age range.

RESULTS

The proposed model outperforms the plain smoothing approach as well as commonly used methodologies while retaining all the desirable properties that demographers expect from a forecasting method, e.g., smooth and plausible age profiles and time trends. We illustrate the proposed approach to mortality data for Danish females and US males.

CONCLUSIONS

The proposed methodology offers a new paradigm in forecasting mortality, and it is an ideal balance between pure statistical methodology and traditional demographic models. Prior knowledge about mortality development can be conveniently included in the approach, leading to large flexibility. The combination of powerful statistical methodology and prior demographic information makes the proposed model suitable for forecasting mortality in most demographic scenarios.

¹ Institut National d'Études Démographiques, Paris, France. Email: carlo-giovanni.camarda@ined.fr.

1. Introduction

Mortality modeling and forecasting are crucial in epidemiology and population studies, as well as in the insurance and pensions industries. In recent decades, several methodologies have been proposed, and many demographers, actuaries, and statisticians have suggested approaches for projecting mortality. (See Booth and Tickle (2008) and Cairns, Blake, and Dowd (2008) and the references therein.)

In addition to scenario- and expert-based approaches in this growing body of methods, we can broadly distinguish three classes based on their assumptions. A traditional procedure relies on using parametric models to describe mortality age patterns and then extrapolating the estimated parameters to reconstruct future mortality (see, among others, Tabeau, Willekens, and van Poppel (2002)). When dealing with adult mortality, parametric models are extremely parsimonious, but a large number of parameters are often necessary when the whole age range is considered. Although these models present the advantage of a clear-cut interpretation of their parameters, this feature is not particularly useful in a forecasting approach. Moreover, when dealing with the whole age range, it is generally hard to disentangle the meaning attached to each parameter and simultaneously forecast plausible mortality patterns based on their time series.

Overparametrized models such as that of Lee and Carter (1992) and its variants have been widely used and have become a benchmark for many newly proposed methodologies. They describe mortality development over age and time using their principal components. They reduce a two-dimensional problem to a fixed-age effect with a univariate time index. The time index condenses the mortality changes in past years and is thus used to forecast future mortality. This class of models presents several drawbacks, however. A simple univariate time series results in mortality improvements at all ages being perfectly correlated. Due to a fixed-age effect over time, lack of smoothness in the estimated mortality pattern is evident, especially in the forecast years. A large number of parameters are implicitly assumed in this class of model. Given the regular structure of human mortality development, this overparameterization may seem unnecessary. Various solutions for these issues have been proposed in successive papers, including a generalization with multiple time indices (Currie 2011, 2013; D'Amato, Piscopo, and Russolillo 2011; Delwarde, Denuit, and Eilers 2007; Hyndman and Ullah 2007; Li, Lee, and Gerland 2013; Renshaw and Haberman 2003). In a Bayesian setting, Girosi and King (2007) enforce smoothness in the age pattern by informative prior. However, they also underline that “almost no matter what one’s prior is for a reasonable age profile, Lee–Carter forecasts although they may be reasonable over the short run will eventually violate it as time passes” (p. 17).

A combination of Lee–Carter and parametric approaches could be found within recently developed Bayesian methods for probabilistic population projections (Raftery et al. 2013; Ševčíková et al. 2016). First, they performed a Bayesian hierarchical model

for forecasting life expectancy at birth, considering available data for all countries in the world. A second step consists of converting overall levels by means of a variant of the Lee–Carter model (Li, Lee, and Gerland 2013) and an adjusted parametric model (Thatcher, Kannisto, and Vaupel 1998). Additionally, coherence in forecasting more populations simultaneously is accounted for. These methods have been used to produce the most recent “World Population Prospects” released by the United Nations (Gerland et al. 2014).

Another class of models used for forecasting mortality is the Age–Period–Cohort (APC) approach. Developed to separate the changes of incidence data with the three demographic coordinates, APC models have been largely adopted to forecast rates in epidemiological instances (Knorr-Held and Rainer 2001; Lopez et al. 2006; Martínez-Miranda, Nielsen, and Nielsen 2014; Riebler and Held 2017; Smith and Wakefield 2016; Tzeng and Lee 2015; Wong et al. 2013). Actuaries and demographers have implemented several variants within this class of models for forecasting mortality (Cairns et al. 2009a, 2011; Haberman and Renshaw 2009; Mammen, Martínez-Miranda, and Nielsen 2015). In this framework, smoothness of the outcomes could be obtained within a Generalized Additive Model framework within a Bayesian setting (Hilton et al. 2019; Ogata et al. 2000), or in a frequentist context (Carstensen 2007; Currie 2013; Heuer 1997). Additional statistical aspects of the APC models in relation to forecasting have been investigated (see, among others, Holford (2006); Kuang, Nielsen, and Nielsen (2008); Nielsen and Nielsen (2014)). A widely recognized issue in APC models concerns identifiability. Since their introduction by Clayton and Schifflers (1987), it is known that estimated coefficients depend on the particular constraints that are used to force a unique solution. Consequently, estimated age, period, and cohort effects cannot be clearly interpretable (Grotenhuis et al. 2016). Nevertheless, forecast values are invariant with respect to the choice of constraint systems when an autoregressive integrated moving average (ARIMA) model is used to forecast period and cohort effects (Currie 2019).

However, both parametric, Lee–Carter and APC approaches are based on rigid modeling structures that are often unable to capture certain features of mortality change.

An alternative compromise method was proposed by Currie, Durbán, and Eilers (2004). They employed a two-dimensional penalized *B*-splines approach (*P*-splines) to smooth mortality over age and time without any specific model structure, allowing for a parsimonious description of mortality development. They treated forecasting of future values as a missing-value problem and estimated the fitted and forecast values simultaneously. Moreover, routines for estimating and forecasting mortality based on this approach are freely available (Camarda 2012).

P-splines have been extensively used for smoothing observed mortality developments in demographic, ecological, and epidemiological studies. See, for instance, Colchero et al. (2016); Goicoa et al. (2012); Jones et al. (2014); Lindahl-Jacobsen et al. (2016); Minton et al. (2017); Ouellette and Bourbeau (2011); Trias-Llimós, Bijlsma, and Janssen

(2016); Ugarte, Goicoa, and Militino (2010). Few studies, however, produced mortality forecasts using this methodology (Bohk-Ewald and Rau 2017; Carfora, Cutillo, and Orlando 2017; Ribeiro 2015; Ugarte et al. 2012). More extensive literature can be found in actuarial science, where smoothness is relevant to future mortality trends (Barrieu et al. 2012; Blake, Cairns, and Dowd 2006; Cairns, Blake, and Dowd 2006; Currie 2016; Djeundje and Currie 2011; Huang and Browne 2017; Lu, Wong, and Bajekal 2014; Pitacco, Denuit, and Haberman 2009; Richards, Kirkby, and Currie 2006; Richards, Currie, and Ritchie 2014; Wang, Yue, and Tsai 2016).

The reason for this mixed recognition of two-dimensional P -splines lies in their lack of robustness for forecasting mortality (Cairns et al. 2009b). Even though P -splines outperform all competitors in modeling mortality, this approach suffers from all the issues that encumber a purely data-driven approach when employed for forecasting purposes. Forecast mortality simply follows estimated trends with a blind adherence to extrapolation, and mortality structure over age is not fully considered in the forecast values. Moreover, the penalty structure, which ensures smoothness in the fitted values, critically affects future mortality forecasts. Unreasonable trends from a demographic perspective could then emerge: increasing mortality over time for specific ages and hence crossover of mortality trends for adjacent ages in future years (cf. Section 2.1).

This paper aims to enhance two-dimensional P -splines through incorporating demographic knowledge into the model, allowing for a better performance in forecasting mortality trends. We retain all features of two-dimensional penalized B -splines and, additionally, we ensure that future mortality over the age range follows a known and well-behaved profile, estimated from past years. This prior knowledge is incorporated by means of asymmetric penalties into the P -spline system (Bollaerts, Eilers, and van Mechelen 2006; Eilers 2005). Since we constrain as well as penalize splines, we call the proposed approach a CP -spline model.

Finally, in the original paper by Currie, Durbán, and Eilers (2004), no solution was proposed for modeling and forecasting mortality for the whole age range. In the following we will also propose a solution for smoothing and forecasting mortality from infancy to oldest old ages.

The remainder of this paper is structured as follows: Section 2 presents basic assumptions and describes the original P -spline methodology which lays the groundwork for further steps. Section 3 is then devoted to the proposed CP -spline approach for incorporating demographic knowledge into the model. Inferences on estimated patterns are also provided. Trends in age-specific death rates and in summary measures such as life expectancy at birth and lifespan variations are presented in Section 4. A critical discussion of the methodology and possible extensions concludes the paper. Throughout the paper and solely for illustrative purposes, two datasets are used: Danish females and US males. Additional supplementary materials will serve as a means to further assess the performance of the proposed method by out-of-sample forecast using eight populations and

in comparison with five alternative forecasting methods. We also validate the influence of the time window and robustness with respect to the parameters used in the model.

2. *P*-splines for mortality data

The proposed model requires two simple datasets as input data: deaths and exposures to the risk of death, arranged in two $m \times n_1$ matrices, $\mathbf{Y} = (y_{ij})$ and $\mathbf{E} = (e_{ij})$:

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n_1} \\ y_{21} & y_{22} & \cdots & y_{2n_1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn_1} \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n_1} \\ e_{21} & e_{22} & \cdots & e_{2n_1} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mn_1} \end{bmatrix}. \quad (1)$$

Rows and columns are classified by single age at death, \mathbf{a} , $m \times 1$ and single year of death, \mathbf{t}_1 , $n_1 \times 1$, respectively.

We assume that the number of deaths y_{ij} at age i in year j is Poisson-distributed with mean $\mu_{ij} e_{ij}$ (Keiding 1990):

$$y_{ij} \sim \mathcal{P}(e_{ij} \mu_{ij}). \quad (2)$$

The value of μ_{ij} is commonly named force of mortality and its estimation is the object of all mortality models. For instance, the matrix of the empirical mortality rates, which are the fully nonparametric estimations of the force of mortality, can be easily computed as $\mu_{ij} \approx m_{ij} = y_{ij}/e_{ij}$. Forecasting approaches aim to reconstruct trends in μ_{ij} for n_2 future years, \mathbf{t}_2 , $n_2 \times 1$.

In the following we will illustrate the proposed method on two populations - Danish females and US males - which differ in terms of their epidemiological history. A more extensive application can be found in the supplementary materials. The left panel in Figure 10 plots observed life expectancy at birth for both populations. (It also plots estimates and forecasts, which should be ignored for now.) From the life expectancy trend, the Danish female population shows a clear mortality stagnation in the 1970s, followed by rising in the last two decades. Cohort effects and high smoking prevalence have been shown to be the main driving factors behind this peculiar pattern (among others, Jacobsen et al. 2004; Lindahl-Jacobsen et al. 2016). US males have shown a clear stagnation of mortality in the 1960s, a rapid improvement in the 1970s, and again a stagnation in recent available years. Several determinants stood out in the literature: smoking behavior (Thun et al. 2013), obesity (Masters et al. 2013), and the performance of the health care system (Muennig and Glied 2010), as well as the recently observed drug epidemic among young adults (Dowell, Noonan, and Houry 2017). A good performance on these

two populations for a newly proposed forecasting method is a clear sign of the robustness and flexibility of the approach. Moreover, using data for both females and males serves as a challenge for the proposed methodology on diverse mortality age patterns. For both populations, we use data from the Human Mortality Database (2019), from ages 0 to 105 over the period 1960–2016, forecasting up to 2050.

We will now give an overview of the P -spline approach for Poisson-distributed data in both one- and two-dimensional settings. A more extensive description of the method can be found in the seminal paper of Eilers and Marx (1996) as well as in the review article by Eilers, Marx, and Durbán (2015). A demographic perspective is provided in Camarda (2008).

In a simple one-dimensional setting, we extract either a column or a row of the original matrices of death counts and exposures, i.e., \mathbf{y} and \mathbf{e} . In modeling mortality, one aims to portray the expected values of the Poisson distribution as follows:

$$\ln[\mathbb{E}(\mathbf{y})] = \ln(\mathbf{e}) + \ln(\boldsymbol{\mu}) = \ln(\mathbf{e}) + \boldsymbol{\eta}, \quad (3)$$

where $\boldsymbol{\eta}$ is the linear predictor and, dealing with Poisson data, a logarithm is used as a link function. The logarithm of the exposures, $\ln(\mathbf{e})$, is commonly called “offset”.

In a parametric setting we would model the linear predictor by a simple structure. For instance, a Gompertz law over age can be written as follows:

$$\boldsymbol{\eta} = \mathbf{X} \boldsymbol{\alpha}, \quad (4)$$

where $\mathbf{X} = [\mathbf{1} : \mathbf{a}]$ and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$. Commonly, these two parameters are used to describe the starting level of mortality and rate of aging, respectively.

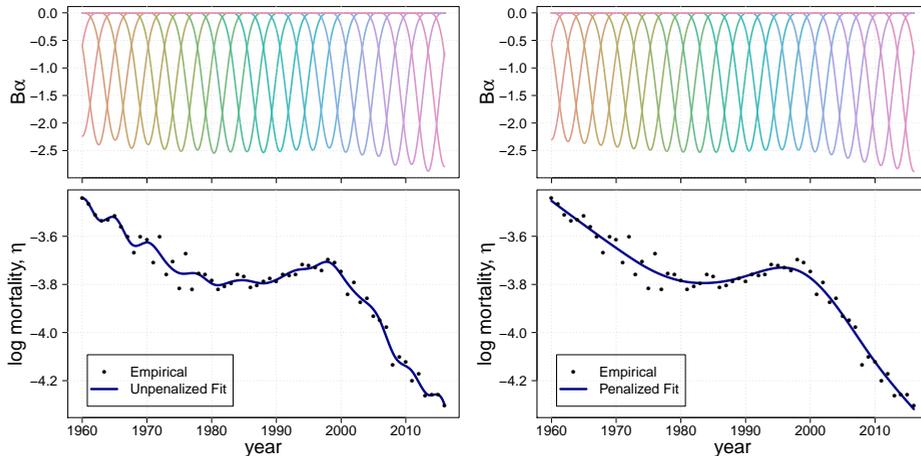
In a smoothing context, instead of deciding a prior mortality shape, we describe the log mortality as a linear combination of B -splines and associated coefficients:

$$\boldsymbol{\eta} = \mathbf{B} \boldsymbol{\alpha}, \quad (5)$$

where \mathbf{B} are k equally spaced B -spline bases: bell-shaped curves composed of smoothly joined polynomial pieces of degree q . In the following we will use cubic B -splines, $q = 3$. The positions on the horizontal axis, where the pieces come together, are called “knots.” Details on B -splines and related algorithms can be found in de Boor (1978), and examples of B -spline bases are provided on the top panels in Figure 1.

The basic idea of the P -splines is to combine (fixed knot) B -splines with a roughness penalty. A relatively large number of B -splines ensures enough flexibility to capture trends in the mortality patterns, and a roughness penalty acting on the associated coefficients enforces the desirable amount of smoothness. Specifically, the number of B -splines, as well as their degree, is irrelevant on final results (Eilers, Marx, and Durbán 2015).

Figure 1: Unpenalized (left) and penalized (right) *B*-splines on Danish female mortality at age 70 from 1960 to 2016



A penalized version of the iteratively reweighted least squares (IRLS) algorithm (McCullagh and Nelder 1989) is sufficient for estimating coefficients $\alpha \in \mathbb{R}^k$:

$$(B' \tilde{W} B + P) \tilde{\alpha} = B' \tilde{W} \tilde{z}, \tag{6}$$

where $\tilde{z} = (\mathbf{y} - \mathbf{e} * \tilde{\mu}) / \mathbf{e} * \tilde{\mu} + \tilde{\eta}$ is the working dependent variable.

The tilde symbol and * denote current approximations to the solution and element-wise product, respectively. \tilde{W} is a diagonal matrix of weights, $\tilde{W} = \text{diag}(\mathbf{e} * \tilde{\mu})$.

The only difference from the standard procedure for fitting a Generalized Linear Model (GLM) with *B*-splines as regressors is the modification of $B' \tilde{W} B$ by a penalty factor given by

$$P = \lambda D' D, \tag{7}$$

where the matrix *D* constructs differences in the coefficients over either ages or years.

As examples, when we have only six coefficients and we use second-order differences, *D* is given by:

$$D = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}. \tag{8}$$

When using a standard P -spline approach, the choice of the order of difference is crucial only for forecasting (cf. Section 2.1). Second-order difference will be used in the following: The smoothing parameter λ regulates the trade-off between goodness of fit and effective dimension used in the model. On the one hand, higher values will lead to a higher penalty term and, consequently, smoother fitted values. On the other, $\lambda = 0$ results in a straightforward GLM estimation with B -splines as regressors.

Figure 1 presents the P -splines logic applied to a one-dimensional example: Danish females aged 70 from 1960 to 2016. On the top panels we have $k = 25$ B -splines multiplied by associated coefficients that determine the height of each B -spline. The bottom panels present empirical and estimated log mortality. On the left panels we have the outcomes from a simple GLM estimation ($\lambda = 0$). The right panels show the penalized version in which the coefficients are forced to change smoothly. The heights of the associated B -splines therefore do not show wiggling behavior and, consequently, neither do the fitted values.

Our aim is to model and forecast mortality over both age and time, so we need to set up a P -splines model in a two-dimensional setting. For the purpose of regression, we arrange the complete matrices as a column vector, that is, $\mathbf{y} = \text{vec}(\mathbf{Y})$ and $\mathbf{e} = \text{vec}(\mathbf{E})$. Then we can directly use Equation (6) to estimate coefficients over age and years by generalizing both the basis and the penalty term.

Let \mathbf{B}_a , $m \times k_a$ and \mathbf{B}_{t_1} , $n_1 \times k_{t_1}$ be the B -splines over ages and years, respectively. The regression matrix for our two-dimensional model is given by

$$\mathbf{B} = \mathbf{B}_{t_1} \otimes \mathbf{B}_a, \quad (9)$$

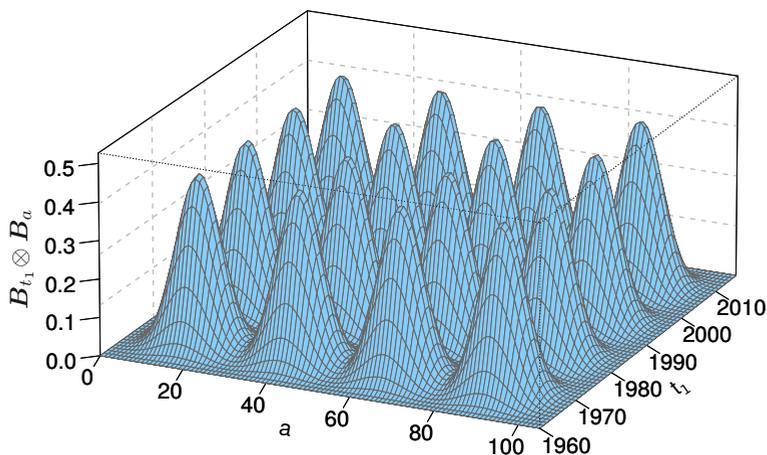
where \otimes denotes the Kronecker product of two matrices. Following the same idea as for the one-dimensional case, we will use a relatively large number of equally spaced B -splines over both domains ($k_a = 24$, $k_{t_1} = 14$). Figure 2 shows a subset of the Kronecker product basis for ages 0–105 and years 1960–2016. The age–year grid is populated by a dense set of overlapping hills that are placed at regular intervals as in an “egg carton.” Associated with each hill we will have a coefficient that will determine the height of its hill. Generalizing in 2D the concept illustrated in Figure 1, the fitted surface will be the sums of these two-dimensional heights.

Concerning the 2D generalization of the penalty term, from the definition of the Kronecker product, Currie, Durbán, and Eilers (2004) show that α can be independently penalized over ages and years. Let \mathbf{D}_a and \mathbf{D}_{t_1} be the difference matrices acting on the two domains. A two-dimensional penalty is given by

$$\mathbf{P} = \lambda_a (\mathbf{I}_{k_{t_1}} \otimes \mathbf{D}'_a \mathbf{D}_a) + \lambda_{t_1} (\mathbf{D}'_{t_1} \mathbf{D}_{t_1} \otimes \mathbf{I}_{k_a}), \quad (10)$$

where λ_a and λ_{t_1} are the smoothing parameters used for age and year, respectively. \mathbf{I}_{k_a} and $\mathbf{I}_{k_{t_1}}$ are identity matrices of dimension k_a and k_{t_1} , respectively.

Figure 2: A subset of two-dimensional Kronecker product cubic B -spline basis



By changing λ_a and λ_{t_1} , smoothness can be tuned to balance smoothness and model fidelity. In the following λ_a and λ_{t_1} will be selected by minimizing the Bayesian Information Criterion (BIC, Schwarz 1978) that penalizes model complexity more heavily and is more suitable for mortality data (Camarda 2008; Currie, Durbán, and Eilers 2004).

As in the one-dimensional setting, B -splines provide enough flexibility to capture surface trends. The additional penalty reduces the effective dimension, leading to a *wisely* parsimonious model with a smoothed fitted surface. The advantage of using two-dimensional P -splines lies also in the fact that different smoothing parameters can be chosen over ages and years, leading to considerable model flexibility. Furthermore, P -splines in 2D can be embedded in the class of Generalized Linear Array Models, saving computational time and reducing storage problems in the estimation of the model (Currie, Durbán, and Eilers 2006).

2.1 Forecasting with P -splines

In the original paper by Currie, Durbán, and Eilers (2004), forecasting is treated as a missing-value problem and the smooth surface is simply extrapolated into future years. Keeping the same age range and forecasting over the years, we augment data and B -spline bases as follows:

$$\check{Y} = [Y : Y_2], \quad \check{E} = [E : E_2], \quad \check{B} = [B_{t_1} : B_{t_2}] \otimes B_a,$$

where \mathbf{Y}_2 and \mathbf{E}_2 are $m \times n_2$ matrices filled with arbitrary future values. In this paper, the complete B -spline basis over years will extend the original basis with an additional eight B -splines.

Finally, let us denote by $\mathbf{1}_{m \times n_1}$ an all-ones matrix over ages and first n_1 observed years. Similarly $\mathbf{0}_{m \times n_2}$ is a zero matrix over ages and future n_2 years. If we define a weight matrix \mathbf{V} :

$$\mathbf{V} = \text{diag}(\text{vec}(\mathbf{1}_{m \times n_1} : \mathbf{0}_{m \times n_2})), \quad (11)$$

we can adapt the algorithm in (6) as follows:

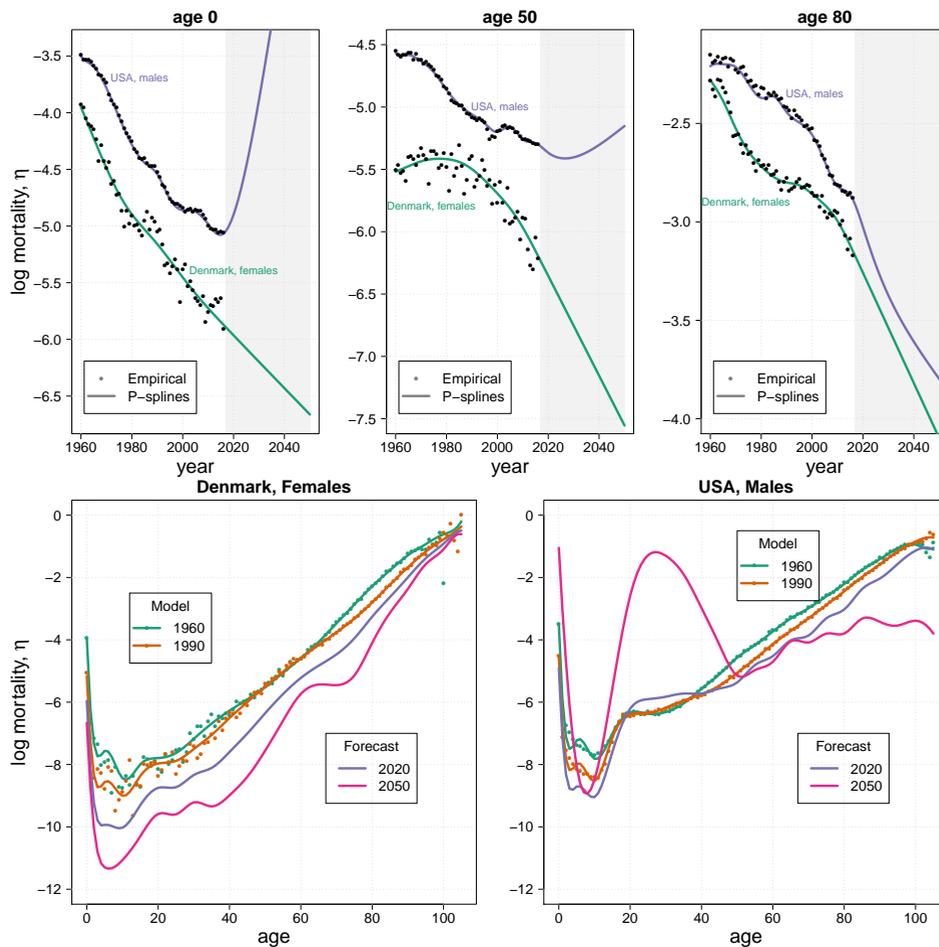
$$(\check{\mathbf{B}}' \mathbf{V} \check{\mathbf{W}} \check{\mathbf{B}} + \mathbf{P}) \check{\boldsymbol{\alpha}} = \check{\mathbf{B}}' \mathbf{V} \check{\mathbf{W}} \check{\mathbf{z}}, \quad (12)$$

with $\check{\mathbf{z}} = \mathbf{V}(\check{\mathbf{y}} - \check{\mathbf{e}} * \check{\boldsymbol{\mu}}) / \check{\mathbf{e}} * \check{\boldsymbol{\mu}} + \check{\boldsymbol{\eta}}$. The penalty is also augmented to account for future years by difference matrices \mathbf{D}_a and $\mathbf{D}_{t_1+t_2}$. This unified structure allows us to simultaneously model and forecast mortality. Moreover, the structure of \mathbf{V} makes it clear that, first, we use all observed data, and second, we assume nothing about the future, i.e., weights equal to one for the observed years and zero for the forecast horizon.

It is noteworthy that the form of the penalty determines the form of the forecast. Whereas the order of the penalty is negligible in the model section, it has a crucial role in the forecasting section. As suggested by Currie, Durbán, and Eilers (2004), second-order differences in $\mathbf{D}_{t_1+t_2}$ are preferable because the forecast coefficients will be approximately linear over the future years, and this choice is more appropriate when working with mortality data. However, this aspect will lose its relevance when prior demographic knowledge is incorporated into the model (cf. Section 3).

Figure 3 presents the outcomes of a two-dimensional P -splines approach in modeling and forecasting mortality data for Danish females and US males. The top panels show empirical and fitted as well as forecast trends for selected ages (0, 50, 80). Concerning estimation on observed data, P -splines show a rather good fit, though unsmooth patterns are visible around age 10 in both populations (bottom panels in Figure 3). Nevertheless, an odd mortality increase in future years is visible for some ages in the US male data, and the forecast mortality improvement at age 80 for Danish females is probably too fast. These outcomes are obviously implausible given the observed mortality trends of the past years. The bottom panels in Figure 3 mirror the above mentioned results from a different perspective. For the future years, mortality age profiles are obviously improbable given the knowledge we already have on the phenomenon: Unreasonable wiggling behavior is evident from age 20 onward in both populations. This outcome can be seen as a consequence of a low smoothing parameter selected by the BIC for the age domain. A small λ_a is necessary in any case to allow flexibility in properly describing the whole age pattern with its notable peak at age 0.

Figure 3: Empirical, model and forecast mortality. Two-dimensional P -spline approach. Selected ages over years (top panels) and selected years over ages (bottom panels)



Note: Danish females and US males, ages 0–105, years 1960–2016, forecast up to 2050.

3. The CP -spline model

In the last section, we noted that several issues may occur when two-dimensional P -splines are applied without any demographic consideration in a plain data-driven approach. Specifically, we need to deal with infant mortality preserving smoothness in the following ages. Moreover, previous outcomes call for inclusion in forecast years of prior knowledge on typical mortality age profiles and time trends.

3.1 Addressing infant mortality

The first year of life is usually treated in a different manner when life tables are constructed (Chiang 1984). Therefore, we decided to follow this practice by modifying the basis related to the age domain. The new basis will then be:

$$\mathbf{B}_a^0 = \begin{bmatrix} 1 & \mathbf{0}_{1 \times k_a} \\ \mathbf{0}_{(m-1) \times 1} & \mathbf{B}_a \end{bmatrix}, \quad (13)$$

where \mathbf{B}_a is now a $(m - 1) \times k_a$ matrix of B -splines.

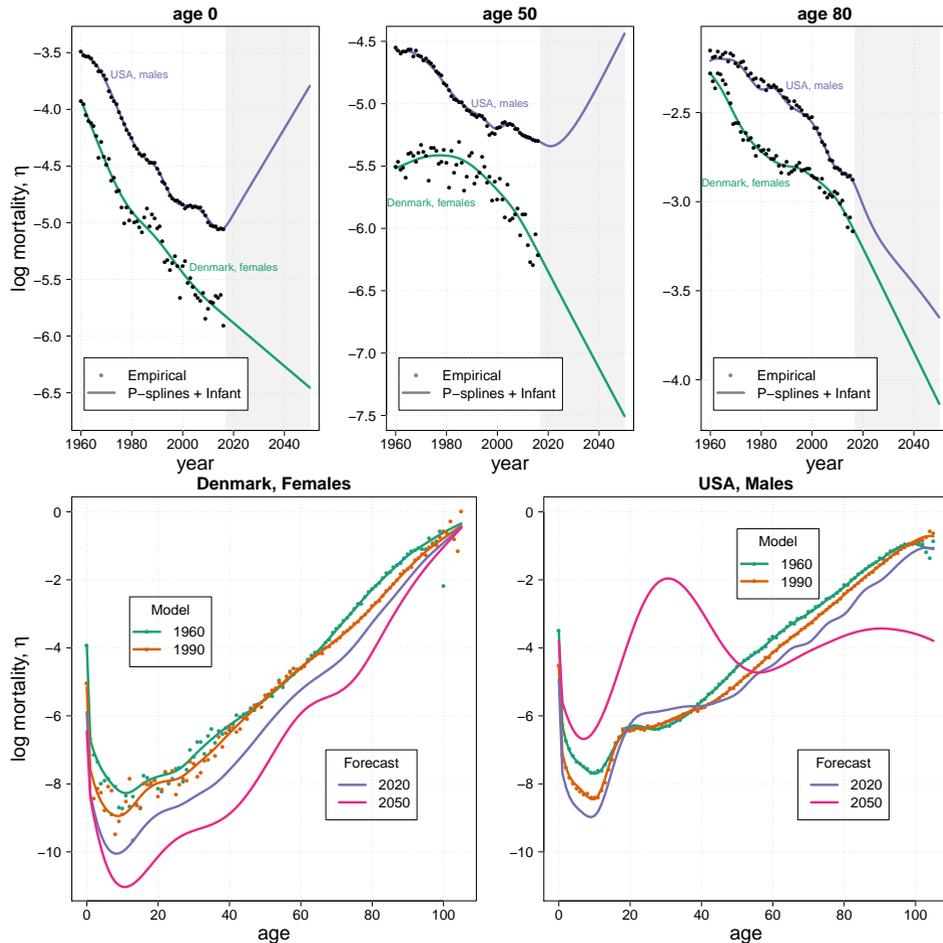
Moreover we replace the first cell in \mathbf{D}_a with a zero, which implies that infant mortality is not connected with variation in subsequent coefficients. In other words, we separate development of infant mortality from the remaining ages.

Using this new basis and a new penalty ensures that a single and specialized coefficient will be attached to infant mortality values. In a one-dimensional setting, this additional coefficient will be exactly the log of death rates at age 0. In a two-dimensional framework, we allow for a smooth change in infant mortality over time.

On the one hand, we increase the number of coefficients in the model. On the other, we allow a certain freedom in describing mortality at age 0 via its specific series of coefficients over years. This ensures that the smoothness of the surface from age 1 onward will not be affected by the evident disruption due to infant mortality.

Figure 4 shows the outcomes for our Danish and US datasets when mortality levels at age 0 are explicitly considered. Although the forecast trends are still unreasonable, especially for US males, considering infant mortality as a peculiar phenomenon helps to improve goodness of fit during the observed period: BIC reduces from 79,308 to 34,795 for US males and from 8,135 to 7,701 for Danish females. The specifically optimized smoothing parameter for the age domain (λ_a) becomes larger with respect to the original approach. As a result, we achieve smoother outcomes over ages, avoiding wiggling behavior around age 10 in both populations.

Figure 4: Empirical, model and forecast mortality. Two-dimensional *P*-spline approach including a specialized coefficient for infant mortality. Selected ages over years (top panels) and selected years over ages (bottom panels)



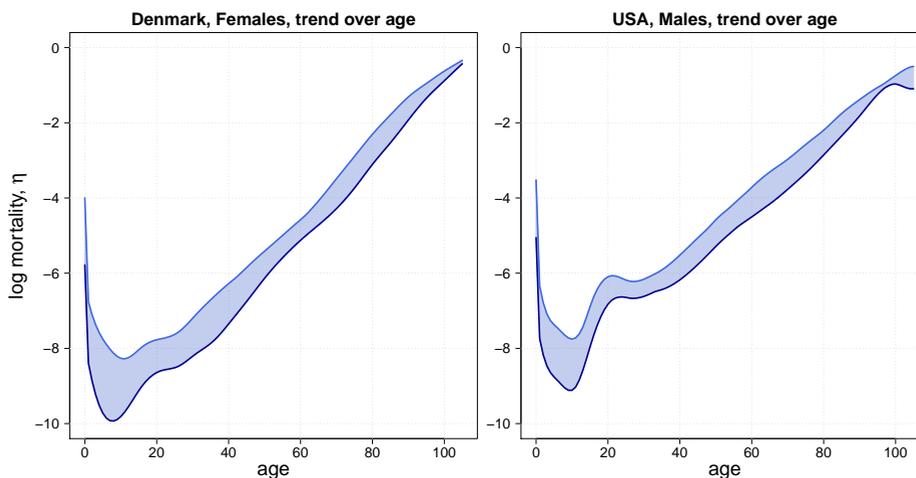
Note: Danish females and US males, ages 0–105, years 1960–2016, forecast up to 2050.

3.2 Enforcing mortality patterns over age and time

As we have seen in Section 2.1, the original P -spline approach is purely data-driven and the extrapolated trends are based on the last estimated coefficients solely constrained by a certain amount of smoothness. However, following Gompertz (1825), demographers started observing well-defined regularities in the shape of mortality over ages. Moreover, past mortality trends present a certain predictability that must guide any model. It would be unreasonable to disregard information on mortality patterns over ages and time in a forecasting method.

Regarding the age dimension, Figure 5 shows the 95% confidence interval over age based on the mortality surface, smoothed by P -splines with the additional specialized coefficient for infant mortality (cf. Section 3.1). A rather stable mortality behavior over ages is evident. In general, at infant ages, mortality decreases steeply, dropping rapidly within the first few years (Levitis 2011). A minimum is commonly reached at about ages 10–15. Afterward, especially for men, mortality rates show a hump at young adult ages (Goldstein 2011; Remund 2015). Mortality then rises exponentially after approximately age 30 and levels off at ages above 80 (Preston 1976; Thatcher, Kannisto, and Vaupel 1998).

Figure 5: 95% pointwise confidence intervals of mortality age profiles from fitted two-dimensional P -splines approach



Note: Model includes a specialized coefficient for infant mortality. Danish females (left panel) and US males (right panel), ages 0–105, years 1960–2016.

We aim to incorporate the information about this stable profile within the forecast

period without modifying fitted values which are based on observed and past data. Instead of borrowing mortality profiles from model life tables or parametric models, we constrain forecast age profiles lie within the 95% confidence interval of the fitted age profiles.

Since we aim to carry out our analysis referring to the mortality shape and regardless of its level, our constraints must be based on the relative derivatives of the age mortality profile, commonly named rate of aging. Given the estimated linear predictor $\hat{\eta} = B\hat{\alpha}$, the rate of aging for each year can be computed by a linear combination of a modified version of the B -splines and the estimated coefficients:

$$\frac{\frac{\partial}{\partial \mathbf{a}} \hat{\mu}}{\hat{\mu}} = \frac{\partial}{\partial \mathbf{a}} \ln(\hat{\mu}) = \frac{\partial}{\partial \mathbf{a}} \hat{\eta} = D_a^{t_1} \hat{\alpha}, \quad (14)$$

where the matrix $D_a^{t_1}$ computes the first difference of the estimated coefficients for each year and simultaneously multiplies them by B -splines of lower degree. In the formula,

$$D_a^{t_1} = B_{t_1} \otimes C_a, \quad (15)$$

where

$$C_a = \frac{1}{h} [q^{-1} B_a^\nu - q^{-1} B_a^{\nu-1}], \quad (16)$$

with h , q and ν being knot distance, degree, and positions of the original B -spline basis, B_a .

In this way, using directly estimated coefficients, we can compute the instantaneous rate of aging over all ages and for each year. This allows us to modify the algorithm in (12) for incorporating possible constraints. Moreover, when working on smooth mortality surfaces, estimated relative derivatives over age will not show the wiggling behavior produced by simple differentiation of observed death rates.

Figure 6 presents 95% confidence intervals of the instantaneous rate of aging over all ages above 0 for Danish females and US males. We denote by δ_L^a and δ_U^a the lower and upper bounds of these confidence intervals, respectively.

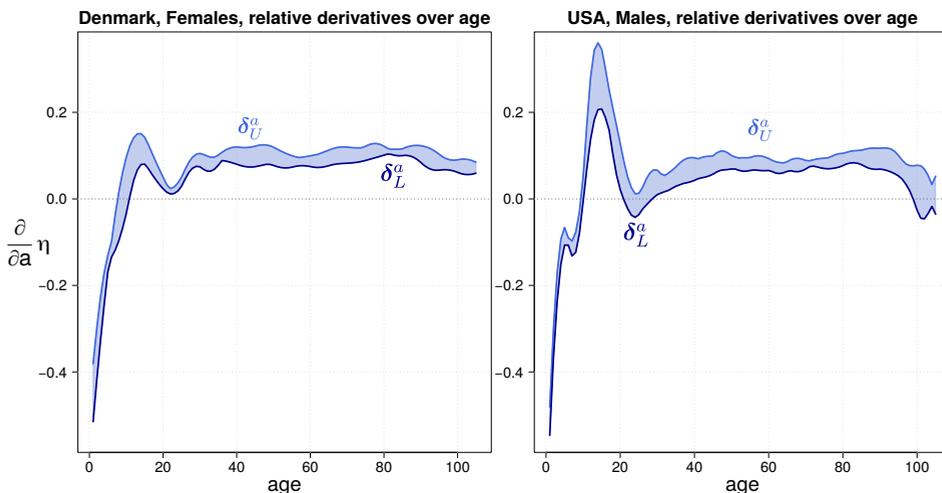
For better readability of the graph, relative derivatives for infant mortality are not displayed. A steep decrease in mortality at age 0 will enormously expand the limits of the ordinate of the associated rate of aging. Specifically, the 95% confidence interval of the relative derivatives with respect to age 0 is $[-2.76, -2.39]$ for Danish females and $[-2.82, -2.56]$ for US males.

In Figure 6 we can read the mortality age patterns of our data without referring to their level. In general, values above zero correspond to mortality increase and, conversely, ages with mortality reduction coincide with negative values of the rate of aging.

The rather constant values, about 0.1 after age 30 in both datasets, correspond to the exponential mortality increase in adult and old ages, commonly described by a single

parameter in the Gompertz model, α_2 in Equation (4). However, we do not predefine a constant rate of aging for adult mortality, and can also capture a possible leveling-off of mortality at oldest old ages. See the declining relative derivatives above age 80.

Figure 6: 95% pointwise confidence intervals of relative derivatives of force of mortality with respect to age, i.e., rate of aging, (δ_L^a, δ_U^a) . Fitted values computed from fitted two-dimensional P -splines approach



Note: Shown only for ages 1–105. Model included a specialized coefficient for infant mortality. Danish females (left panel) and US males (right panel), ages 0–105, years 1960–2016.

The rate of aging also shows clear sex differences in young adult mortality: Danish females present a U-shape pattern about age 20 that is much less pronounced than for US males. This latter population also reaches relative derivatives equal to zero around age 25, which means a corresponding constant mortality. This sex difference points out the young adult excess mortality that is particularly evident for males (see Figure 5). In both populations, the decreasing mortality trend during childhood is clear: The associated rate of aging shows a steeply increasing pattern with negative values.

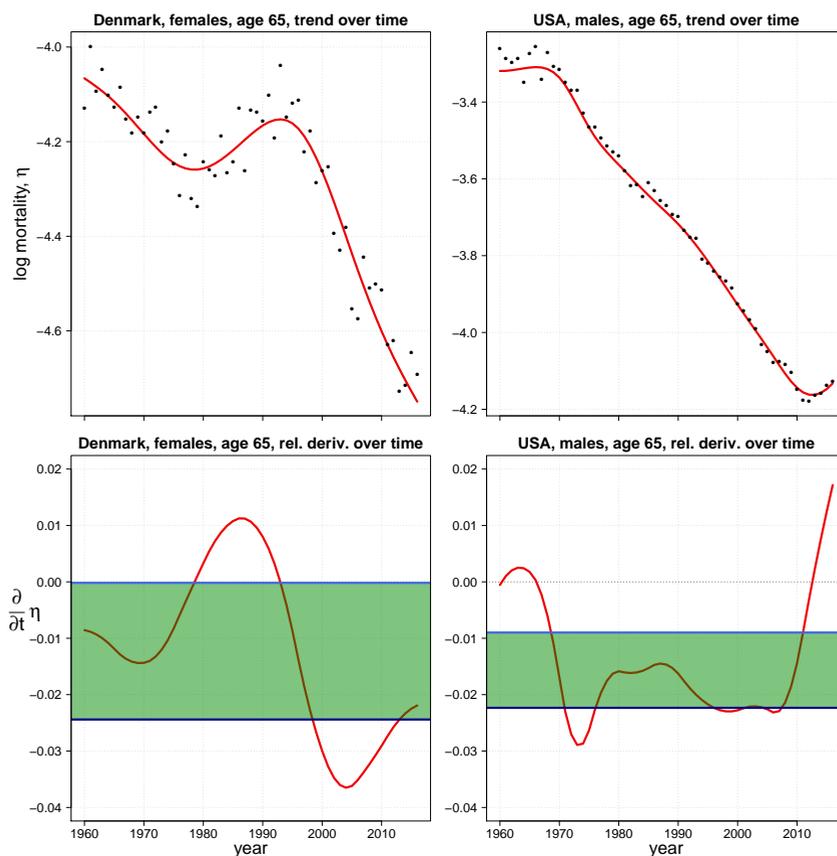
Concerning time trends, we can compute relative derivatives with respect to time as we did for the age domain:

$$\frac{\frac{\partial}{\partial t_1} \hat{\mu}}{\hat{\mu}} = \frac{\partial}{\partial t_1} \ln(\hat{\mu}) = \frac{\partial}{\partial t_1} \hat{\eta} = \mathbf{D}_{t_1}^{t_1} \hat{\alpha}. \quad (17)$$

Again matrix $\mathbf{D}_{t_1}^{t_1}$ computes first difference of estimated coefficients by differentiating the associated B -spline basis over years for each age. The rate of change over time

fluctuates more than the age profile, as one can see by looking at a particular age (65) for both Danish females and US males in Figure 7. Whereas the trend is generally smooth and downward, there are periods of mortality increase: during the 1980s for Danish females and the most recent years for US males. Mortality stagnation in US males is also evident during the 1960s.

Figure 7: Empirical and smooth log mortality for age 65 over time (top panels) and associated rate of change (bottom panels). Horizontal blue lines depict 50% confidence intervals of the rate of change. Fitted values computed by two-dimensional P -splines approach



Note: Model includes a specialized coefficient for infant mortality. Danish females (left panel) and US males (right panel), ages 0–105, years 1960–2016.

These features are immediately visible in the associated relative derivatives with respect to time for mortality at age 65 in both populations (bottom panels in Figure 7), which express mortality improvement over time regardless of the actual level. Likewise for the rate of aging, negative values correspond to downward mortality trends, which represents the majority of the observed relative derivatives for this specific age. Positive values are observed when mortality stagnates and/or deteriorates.

Although values for mortality rate of change over time are smaller than the observed rate of aging, variation is wider: Every mortality fluctuation over time - even minor - is amplified in the associated relative derivatives. Note, for example, the time trend in US males aged 65 in the most recent years: A rather small estimated increase translates into a disproportionate jump in the rate of change (right panels in Figure 7). This is a well-known issue in statistics: Derivatives will always show undersmooth behavior with respect to the associated estimated function (Erickson, Fabian, and Marik 1995).

On the one hand, we aim to constrain future mortality developments to lie within observed mortality rates-of-change. On the other, we intend to avoid in future years the peculiar past mortality trends that we assume to be solely due to specific and unlikely events. For this reason we decide to use only 50% confidence intervals of the observed mortality rate of change, i.e., the interquartile range of past experienced mortality development.

We recommend the interquartile range based on several experiments conducted on numerous populations from the Human Mortality Database (2019) (not shown here). However, it is important to note that this value (50%) is simply a way to express the forecaster's prior knowledge of how past mortality should inform future development. As a result, different attitudes toward the future or the peculiar mortality history of a specific population may guide forecasters to different values in computing $\delta_L^{t_1}$ and $\delta_U^{t_1}$. In the Supplementary Materials C we evaluate this choice for Danish females and US males. Outcomes do not change markedly, as long as extreme mortality fluctuations over time are not considered. Moreover, we present an out-of-sample forecast exercise which could be eventually used to select population-specific percentages for computing $\delta_L^{t_1}$ and $\delta_U^{t_1}$.

The horizontal green stripes in the bottom panels of Figure 7 depict the 50% confidence intervals of the observed mortality rate of change for age 65 in both populations.

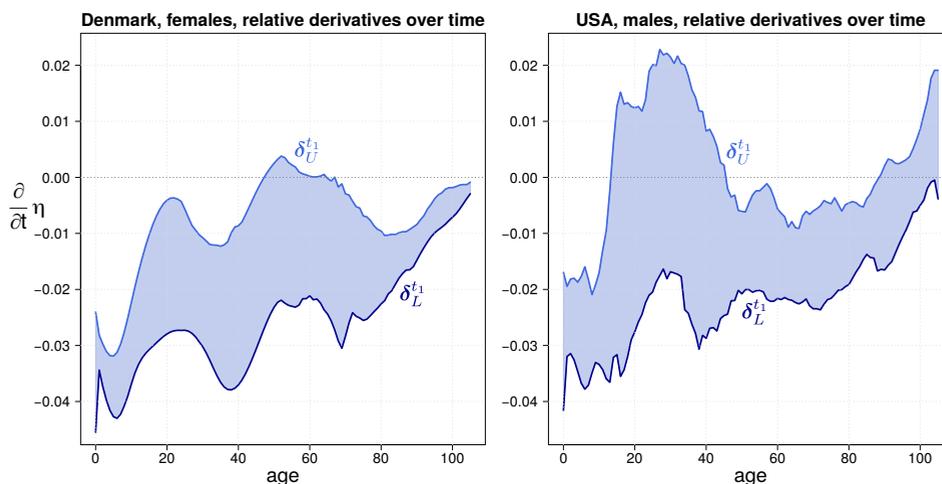
Figure 8 shows the 50% confidence intervals of the relative derivatives of the force of mortality with respect to time for each age. The lower and upper levels of these confidence intervals are denoted by $\delta_L^{t_1}$ and $\delta_U^{t_1}$, respectively.

In Figure 8 we can easily see which ages have experienced greater improvements (larger negative values) as well as with greater variations in terms of observed mortality changes (broader confidence bounds).

The idea is to inform our model about rate of aging and mortality changes over time, i.e., constrain future mortality to lie within a range of plausible age patterns and

time trends expressed by estimated values and portrayed in Figures 6 and 8. Hence, we propose the Constrained Penalized spline (*CP*-spline) model.

Figure 8: 50% pointwise confidence intervals of relative derivatives of force of mortality with respect to time for each age, i.e., age-specific rate of change. Fitted values computed from fitted two-dimensional *P*-splines approach



Note: Model included a specialized coefficient for infant mortality. Danish females (left panel) and US males (right panel), ages 0–105, years 1960–2016.

Although specific values are suggested for computing δ , forecasters can adapt the model to their needs and to prior knowledge about future mortality development. In general, the lower the percentages in estimating the vectors of δ , the closer future mortality will be to mean age patterns and time trends, as observed in past years. In other words, extremely low percentages for δ lead to a fixed age profile, along with an invariant age-specific rate of mortality improvement, i.e., something similar to the Lee–Carter model, but with a nonlinear time index based on the amount of smoothness. Conversely, large confidence levels leads to an extremely flexible *CP*-spline model without much prior knowledge about future mortality, i.e., a plain *P*-splines approach. Additionally, different levels of mortality improvement can be set for different ages, making the model highly flexible. Finally, the choice of the percentage for computing δ_L^a and δ_U^a is less important than for $\delta_L^{t_1}$ and $\delta_U^{t_1}$. Derivatives are computed on previously smooth rate values, and commonly, human mortality shows a rather regular pattern over ages. Consequently, val-

ues higher than 95% for constraining the age profile will only slightly modify the final outcomes.

We must warn the forecaster on two important issues. First, allowing high flexibility may lead to unreasonable outcomes (see Figure 3), whereas the other, a rigid model will forecast patterns that slavishly reproduce the structure of the model in future years. Second, recommended levels for computing δ in CP -splines should be adopted with care: Although they have been tested for many datasets in the Human Mortality Database (2019), specific populations might need distinct confidence levels to either account for or neglect unique patterns over age and/or years.

3.2.1 Incorporating prior knowledge into the model

Once the constraints are set, we retain them for all years by augmenting the values in δ over both dimensions:

$$\begin{aligned} g_L^a &= \mathbf{1}_{n_1+n_2} \otimes \delta_L^a & \text{and} & & g_L^t &= \mathbf{1}_{n_1+n_2} \otimes \delta_L^{t_1} \\ g_U^a &= \mathbf{1}_{n_1+n_2} \otimes \delta_U^a & & & g_U^t &= \mathbf{1}_{n_1+n_2} \otimes \delta_U^{t_1} \end{aligned} \quad (18)$$

We enforce our shape constraints by adding asymmetric penalties within the system introduced in (12):

$$(\check{B}'V\check{W}\check{B} + P + P^a + P^t)\check{\alpha} = \check{B}'V\check{W}\check{z} + p^a + p^t, \quad (19)$$

where

$$\begin{aligned} P^a &= P_L^a + P_U^a & \text{and} & & p^a &= p_L^a + p_U^a \\ P^t &= P_L^t + P_U^t & & & p^t &= p_L^t + p_U^t \end{aligned} \quad (20)$$

As an example, we present the penalty terms for the lower bounds over ages. The other terms are constructed in a similar fashion. In formulas

$$\begin{aligned} P_L^a &= \kappa D_a^{t_1+t_2'} \text{diag}(s \tilde{v}_L^a) D_a^{t_1+t_2} \\ p_L^a &= \kappa D_a^{t_1+t_2'} \text{diag}(s \tilde{v}_L^a) g_L^a \end{aligned} \quad \text{with } v_L^a = \begin{cases} 0 & \text{if } D_a^{t_1+t_2} \tilde{\alpha} \geq g_L^a \\ 1 & \text{if } D_a^{t_1+t_2} \tilde{\alpha} < g_L^a \end{cases},$$

and s is a 0/1 vector equal to 1 when the constraint is to be applied (future years).

Note that, being asymmetric, these penalties act on both the left and right sides of the system of equations, and values in v_L^a , v_U^a , v_L^t and v_U^t are computed iteratively. In other words, for each new value of $\tilde{\alpha}$ during the iteration (19), the shape penalties exert an influence only when the shape constraint is violated. The size of κ regulates how strictly the constraints are enforced. In this paper, we chose $\kappa = 10^4$, which is an intermediate value in this setting. We inform the model about our shape constraints, but meanwhile their effects should not overwhelm the smoothness behavior of the fitted

values as expressed in the roughness penalty P . More details on asymmetric penalties and their statistical properties can be found in Eilers (2005) and Bollaerts, Eilers, and van Mechelen (2006).

3.2.2 Confidence intervals by bootstrap

No forecasting method can be completely satisfactory without an estimation of the uncertainty affecting the forecast quantities. Estimation of confidence intervals is thus particularly necessary. It is noteworthy that CP -splines intrinsically constrain future log mortality within defined ranges of possible outcomes, and these constraints hold within the estimation of the associated standard errors. The consequences are twofold: On the one hand, confidence intervals for future years will always contain values that lie within the defined shape constraints and, on the other, the interval will often be narrower with respect to alternative methods. Relaxation of the shape constraints by smaller values of κ and larger confidence levels for computing δ , or accounting for the overdispersion that is generally found in mortality data may mitigate this drawback. These generalizations of the current approach will be avenues for future research.

In practice, without shape constraints, plain two-dimensional P -splines are a straightforward extension of a regression model, and methods for computing the covariance matrix and associated standard errors can be borrowed from regression theory. However, with our CP -spline model, we depart from this setting with the inclusion of asymmetric penalties in the system of equations.

In the absence of analytical solutions for the estimation of uncertainty in our model, we opt for confidence intervals constructed via a bootstrap approach. We are thus able to combine all sources of uncertainty in the model and simultaneously compute confidence intervals for summary measures, which are complicated nonlinear functions of the estimated coefficients. Details on bootstrap methods in general can be found in Efron and Tibshirani (1993). While bootstraps have been used by Brouhns, Denuit, and Van Keilegom (2005) and Koissi, Shapiro, and Högnäs (2006) in the Lee–Carter context, Ouellette, Bourbeau, and Camarda (2012) have adapted this methodology to a two-dimensional P -spline model.

Following Koissi, Shapiro, and Högnäs (2006) and Ouellette, Bourbeau, and Camarda (2012), we carried out a residual bootstrap of our fitted model. Specifically, deviance residuals are the standard measures to assess the discrepancy between empirical and fitted data (McCullagh and Nelder 1989: 39–40):

$$r = \text{sign}(\mathbf{y} - \hat{\mathbf{y}}) \sqrt{2 \left[\mathbf{y} \ln \left(\frac{\mathbf{y}}{\hat{\mathbf{y}}} \right) - \mathbf{y} + \hat{\mathbf{y}} \right]}. \quad (21)$$

These residuals should be independent and identically distributed (provided the model

is well specified). Hence, we sample from them (with replacement) an entire new set of residuals $\mathbf{r}_{(b)}$ called the bootstrapped residuals. Replacing deviance residuals \mathbf{r} with bootstrapped residuals $\mathbf{r}_{(b)}$ in (21) and rearranging the equation leads to

$$\hat{\mathbf{y}} - \mathbf{y} \ln(\hat{\mathbf{y}}) + \mathbf{r}_{(b)}^2 + \mathbf{y} - \mathbf{y} \ln(\mathbf{y}) = 0. \quad (22)$$

Given $\mathbf{r}_{(b)}$ and actual death \mathbf{y} , equation (22) can be solved numerically with respect to $\hat{\mathbf{y}}$, thus obtaining a new matrix of bootstrapped deaths $\hat{\mathbf{y}}_{(b)}$. Together with the original exposures \mathbf{e} , we can estimate the proposed *CP*-spline model on the bootstrapped deaths $\hat{\mathbf{y}}_{(b)}$, obtaining new bootstrapped coefficients $\hat{\boldsymbol{\alpha}}_{(b)}$.

The procedure described above, starting with the residual sampling step, was repeated 1,000 times. This led to a bootstrapped distribution of constrained and penalized coefficients. From this distribution, we extract empirical percentiles and compute confidence intervals for force of mortality $\boldsymbol{\mu}$ and linear predictor $\boldsymbol{\eta}$, as well as for any desirable summary measure, e.g., life expectancy at birth.

Whereas common approaches focus on the variability in the (univariate) time index (see both Lee–Carter and Hyndman–Ullah variants), residual bootstrap incorporates the variability from all model parameters. This approach is more suitable in a nonparametric framework, and it allows us to account for uncertainty due to the Poisson stochastic process in (2), i.e., larger populations would tend to have narrower confidence intervals in the fitting periods.

When we apply any distinct forecasting model we do not question uncertainty due to model misspecification. Similarly, within a *CP*-spline framework, we specify a certain model structure: Values of optimized smoothing parameters and confidence levels for computing $\boldsymbol{\delta}$ are kept fixed in the bootstrapping procedure. However, unlike other approaches, complexity of the model structure is data-driven in the estimation procedure. See columns with effective dimension (ED) in Table A-1 and A-2 in Supplementary Materials.

To sum up, a forecaster that intends to estimate and forecast mortality by *CP*-splines needs to adopt the following procedure:

1. Collect as in (1) deaths and exposures in two matrices over age and years;
2. Estimate a two-dimensional *P*-spline model over the observed period with the algorithm in (6) and a specialized basis for infant mortality as in (13);
3. Optimize smoothing parameters in the penalty term (10);
4. Evaluate rate of aging and rate of change for each age using (14) and (17);
5. Portray previous relative derivatives as in Figures 6 and 8 to decide on level of confidence for future rate of aging (δ_L^a, δ_U^a) and rate of change ($\delta_L^{t_1}, \delta_U^{t_1}$). Here, eventual prior knowledge about mortality developments in a specific population could modify the recommended 95% and 50% levels;

6. Solve the system of equations in (19), which adds in the forecasting algorithm (12) the asymmetric penalty terms;
7. Carry out the residual bootstrap presented above to obtain confidence intervals of the fitted model.

Routines for running all previous steps and forecast mortality with CP -splines were implemented in \mathbb{R} (R Development Core Team 2019). Available on the journal's website are routines that do not require stringent package installations. In terms of computational time, for a single population, fitting the model takes about 20 seconds, and it takes less than 80 minutes to run the bootstrap procedure with 1,000 simulations (portable personal computer, Intel i5-6300U processor, 2.4 GHz \times 4 and 16 Gbytes random-access memory). Concerning computer data storage, a file of about 125 MB can store the \mathbb{R} workspace with all outcomes for a single population.

4. Application

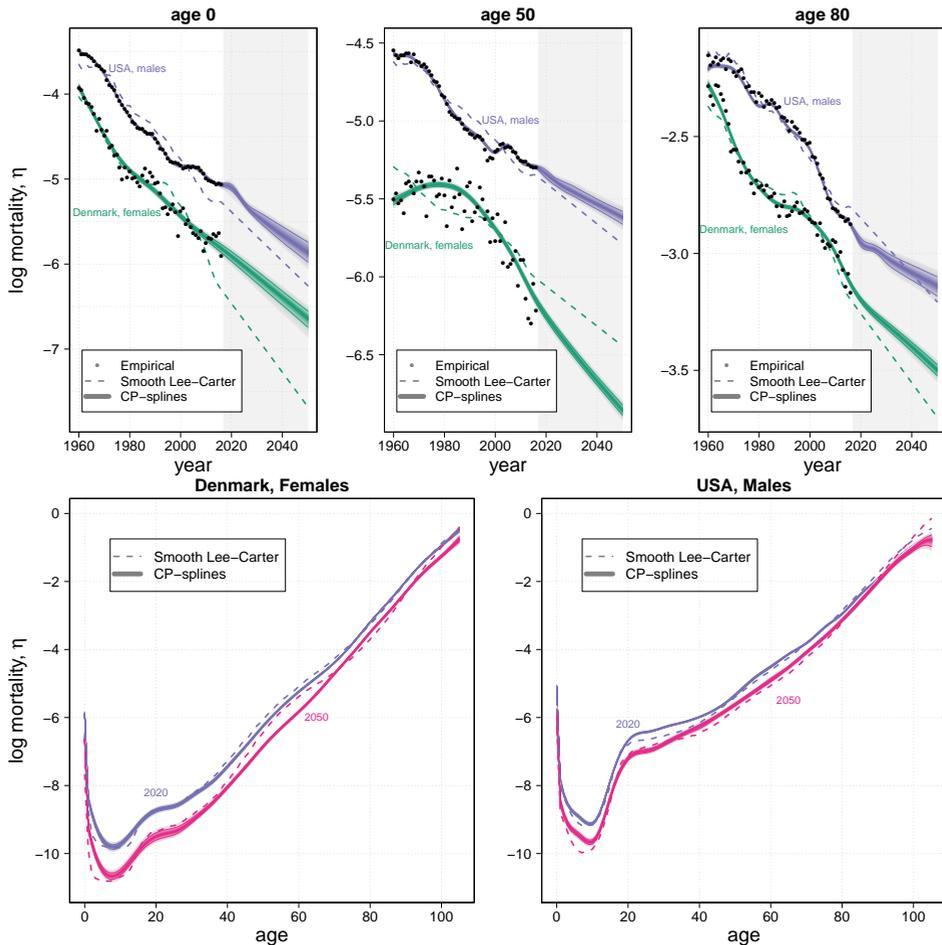
Figure 9 shows the outcomes of the proposed smooth constrained forecasting model. In order to better visualize uncertainty around estimated values, we portray outcomes by means of `fanplot` (Abel et al. 2013). Colored bands are limited by 10% and 90% of the empirical percentiles.

Moreover, in this section, we compare the outcomes from the suggested CP -splines with a variant of the Lee–Carter model. Specifically, we apply the Lee–Carter model estimated in a smooth setting as described in Delwarde, Denuit, and Eilers (2007). Both the CP -spline model and this Lee–Carter variant are embedded in a Poisson framework, and smoothing of the coefficients is ensured in both settings. Hence, differences between models will be solely due to differences in model structure.

It is evident in Figure 9 how fitted values from CP -splines follow empirical patterns adequately and forecast values from CP -splines present reasonable trends over both ages and years. Specifically, there is no wiggling behavior because we enforce a specific shape via the asymmetric penalties, which also ensures no crossover of adjacent ages in the long term. Moreover, future mortality of US males no longer shows increasing trends.

In addition to yielding an accurate modeling performance due to its flexibility, the proposed approach ensures smooth mortality age patterns for future years. This is extremely important in both demographic and actuarial studies because it means that forecast mortality can be treated in a continuous setting.

Figure 9: Empirical, model and forecast mortality along with their bootstrapped distributions. Constrained Penalized spline model including a specialized coefficient for infant mortality. Estimates from a smooth Lee–Carter variant are given for comparative reasons



Note: Colored bands depict 80% of the empirical percentiles. Selected ages over years (top panels) and selected years over ages (bottom panels). Danish females and US males, ages 0–105, years 1960–2016, forecast up to 2050.

At first glance, the boundary of the confidence intervals in Figure 9 is extremely close to the fitted values in the observed period (1960–2016). This is mainly due to the large expected values in the Poisson distribution (2) which are the products of force of mortality and population exposure, remarkably large for US males. Consequently, we see relatively larger confidence bands at oldest ages over the years. In this area, high force of mortality is compensated by exposures with moderate counts. Similarly, slightly wider variability is observed at young ages (see pattern at age 20). This results from low force of mortality along with a large exposure population. The relatively larger uncertainty at age 0 is due to its peculiar treatment within the model: Smoothness is enforced only over time and therefore larger variability is expected. Moreover, variability increases for all ages the further we move toward future years.

In Figure 9 it is evident how the proposed *CP*-spline approach clearly outperforms the Lee–Carter model in fitting past trends: For US males, the deviance that measures goodness of fit in a Poisson setting is equal to 32,906 for *CP*-splines and 175,726 for the smooth Lee–Carter variant. For Danish females, the deviance is equal to 7,728 (12,831) for *CP*-splines (smooth Lee–Carter). A larger comparison with other forecasting methods is available in the supplementary materials to this paper.

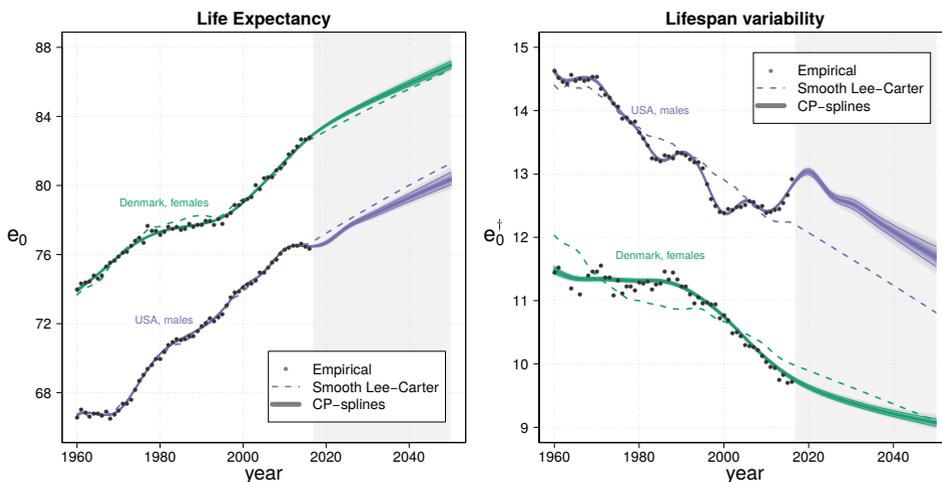
Concerning patterns for future years, the Lee–Carter model is not able to capture all observed mortality changes of the past decades, and its forecast trends are often unreasonable, i.e., a simple linear extrapolation of fitted values, which are a poor description of empirical trends. It is noteworthy that, although embedded in a smooth setting, the rigidity of the Lee–Carter structure leads to unsmooth fitted values. On the contrary, the proposed model does not suffer from this drawback and is able to accommodate diverse trends.

A direct consequence of these differences is visible in the age profiles in 2050 (bottom panels of Figure 9). Whereas the *CP*-spline model yields smooth and plausible future mortality age-patterns, the Lee–Carter produces unrealistic wiggling behavior of the age-profile in 2050 for both populations.

We decided to assess the performance of the proposed *CP*-spline model by means of two complementary summary measures. Life expectancy at birth is used as a classic measure of average lifespan. Lifespan variation describes differences in the length of life across members of a population and, among the large number of possible measures, we opt for the average number of life years lost at birth (Vaupel and Canudas-Romo 2003; Zhang and Vaupel 2009), commonly denoted by e_0^\dagger . Easy to interpret as a potential increase in life expectancy at birth, this measure has already been used to evaluate the performance of mortality forecasts (Bohk-Ewald, Ebeling, and Rau 2017).

The results in terms of these summary measures are presented in Figure 10. As in the case of log mortality, the *CP*-spline model fits the development of e_0 over the past years rather well in both populations. The largest difference in e_0 between empirical and fitted values is equal to 0.26 and 0.65 years for US males and Danish females, respectively.

Figure 10: Empirical, model, and forecast values for life expectancy at birth (left panel) and a measure of lifespan variability (e_0^\dagger , right panel). Constrained Penalized spline model including a specialized coefficient for infant mortality



Note: Colored bands depict 80% of the empirical percentiles. Estimates from a smooth Lee-Carter variant are given for comparative reasons. Danish females and US males, ages 0–105, years 1960–2016, forecast up to 2050.

In 2050 the proposed method results in a life expectancy at birth of 86.97 years with a 95% confidence interval (86.66–87.30) for Danish females, and of 80.36 years for US males (79.88–80.92). We compare our results to those released by the United Nations in the 2019 World Population Prospects, medium variant (United Nations 2019). For Danish females, for the periods 2045–2050 and 2050–2055, the results provide a life expectancy at birth equal to 86.13 and 86.68, respectively. These values are slightly lower than our forecast 95% confidence intervals. On the contrary, we produce more pessimistic outcomes for US males: United Nations forecast e_0 equal to 81.40 and 82.17 for the two periods, about 2050.

Lifespan variability measured by e_0^\dagger is captured extremely well by the *CP*-spline model for the period 1960–2016: The largest errors in estimating the average number of life years lost at birth is equal to 0.12 (0.24) years for US males (Danish females). Forecast trends also appear to be reasonable. For US males, we forecast an average of 11.69 life years lost at birth in 2050 with a 95% confidence interval (11.46–11.92). Danish females already start with a lower level of e_0^\dagger and smoothly move to a value of 9.07 in 2050 (8.99–9.17), with a leveling-off trend. The relationships between both summary

measures are also adequately described and forecast by the proposed model (not shown here).

Despite a similar final value in 2050, the development of future life expectancy at birth is different from that of the Lee–Carter model, which simply extrapolates a linear trend from erroneously estimated values. The proposed CP -spline model, on the other hand, is able to capture curved trends over time and consequently to provide more reliable forecasts.

Unlike the proposed approach, which is capable of reproducing all changes in lifespan variability measures, the Lee–Carter approach poorly estimates observed trends in e_0^\dagger , and then linearly extrapolates its fitted values.

In the supplementary materials, we further assess the performance of the proposed model in additional ways. We compare CP -splines with alternative methods (Supplementary Material A), and we test stability with respect to the time window over which the model is estimated, an important and often neglected choice in all forecasting methods (Supplementary Material B).

Specifically, we present a large and detailed comparison study in which CP -splines have been analyzed with five alternative forecasting methods: the mentioned smooth variant of the Lee–Carter model (Delwarde, Denuit, and Eilers 2007); the Lee–Carter variants proposed Lee and Miller (2001) and by Booth, Maindonald, and Smith (2002), and two of the approaches proposed within a functional data framework by Hyndman and Ullah (2007). We model four countries (United States, Denmark, France, and Japan), both males and females, for the period 1960–2016 and forecast up to 2050. For the observed years, we assess goodness of fit (balanced with model complexity) using the Bayesian Information Criterion. In all datasets the proposed CP -splines outperformed the alternative approaches.

To evaluate the performance against observed mortality trends, we performed an out-of-sample forecast study. We estimate all models on the period 1960–2006, forecast up to 2016 and compare to observed values in 2005–2016. Models are compared using three different measures computed on e_0 , e_0^\dagger and the whole mortality surface (η). Given four populations, two sexes, three accuracy measures, and three demographic indicators, we compare six alternative approaches over 72 values. The proposed CP -splines outperformed its competitors 39 times (54%), followed by a variant of Hyndman–Ullah model with only 16 times. A subset of the whole out-of-sample forecast exercise is presented here in Table 1. Additional details are shown in Supplementary Materials A (Table A-1 and A-2).

Table 1: Root Mean Square Errors from the out-of-sample forecast exercise computed on e_0 , e_0^\dagger and η . Lower values (in bold) correspond to greater accuracy. Fitting period 1960–2006, forecast up to 2016 and compared to observed values in 2007–2016

		<i>CP-S</i>	<i>LCsmo</i>	<i>LM</i>	<i>BMS</i>	<i>HU</i>	<i>HUrob</i>
USAf	e_0	0.154	0.203	0.184	0.200	0.369	0.302
	e_0^\dagger	0.224	0.229	0.204	0.200	0.313	0.276
	η	0.079	0.128	0.123	0.123	0.103	0.086
DNKf	e_0	0.690	1.420	0.793	1.299	1.237	1.033
	e_0^\dagger	0.235	0.850	0.658	0.708	0.516	0.248
	η	0.324	0.420	0.367	0.373	0.344	0.328
JPNf	e_0	0.404	1.049	0.777	0.986	0.319	0.601
	e_0^\dagger	0.105	0.317	0.505	0.455	0.298	0.126
	η	0.121	0.287	0.241	0.265	0.123	0.122
FRAf	e_0	0.368	0.431	0.358	0.438	0.172	0.258
	e_0^\dagger	0.047	0.263	0.244	0.267	0.101	0.048
	η	0.107	0.180	0.157	0.157	0.120	0.109
USAm	e_0	0.192	0.269	0.297	0.264	0.306	0.304
	e_0^\dagger	0.349	0.614	0.570	0.600	0.469	0.399
	η	0.092	0.132	0.126	0.128	0.110	0.099
DNKm	e_0	1.069	1.440	1.457	3.804	1.047	1.345
	e_0^\dagger	0.308	0.247	0.409	0.861	0.097	0.250
	η	0.358	0.372	0.427	0.809	0.364	0.373
JPNm	e_0	0.152	0.391	0.311	0.389	0.226	0.300
	e_0^\dagger	0.095	0.088	0.092	0.098	0.260	0.196
	η	0.104	0.154	0.163	0.168	0.101	0.097
FRAm	e_0	0.297	0.463	0.305	0.617	0.166	0.540
	e_0^\dagger	0.081	0.130	0.089	0.160	0.205	0.053
	η	0.101	0.209	0.455	0.433	0.134	0.138

Note: Populations: United States, Denmark, Japan, and France; females and males. Models: *CP*-splines (*CP-S*), smooth Lee–Carter (*LCsmo*), Lee–Miller (*LM*), Booth–Maindonald–Smith (*BMS*), Hyndman–Ullah (*HU*), and a robust version of Hyndman–Ullah (*HUrob*). Larger comparison is shown in the Supplementary Materials.

5. Conclusions

Our paper starts from a simple consideration about established forecasting methods for mortality. We recognized that widely used methodologies are either too rigid to properly

capture mortality developments over age and time, or too flexible to impose certain well-known structures in absence of observed data.

Among the rigid models, one can certainly list parametric models such as Gompertz and Heligman-Pollard. These models predefine a mortality law over age, and forecast estimated parameters will reproduce with a blind adherence these laws in future years. However, several Lee-Carter variants suffer from equivalent drawbacks describing all the mortality developments within a bilinear model and fixing an age-specific rate of mortality improvement. To free mortality models from rigid structures, nonparametric methods have been suggested. Superior in describing observed patterns, these approaches do not account for demographic knowledge in guiding forecast mortality developments.

Our study bridges the gap between these seemingly distant approaches. We enhance a powerful nonparametric statistical methodology, incorporating observed demographic information from the past years into the model. Under a P -spline approach we obtain good fit, flexibility, and smooth outcomes. Nevertheless, we show that a plain smoothing approach results in unreasonable outcomes when used to forecast mortality. This purely data-driven approach is not able to reproduce past mortality experience because it extrapolates last estimated trends with blind adherence. A certain amount of smoothness is the only restriction integrated into the model, and it is not sufficient when no data are available, as in the case for future years.

Thus, our proposal accommodates the P -splines approach, constraining future mortality to lie within trends estimated from observed data. We thus propose a Constrained Penalized spline (CP -spline) model. Initially, we consider infant mortality as a specific age and attach to age 0 a specialized coefficient. This improves the estimation of past trends but is insufficient to correct inconsistent trends in future years. Consequently, we also incorporate information about past mortality experience. Instead of working on observed log mortality, we operate in terms of rate of aging and rate of change over time. In practice, we enforce shape constraints by asymmetric penalties based on observed relative derivatives of the force of mortality with respect to age and time. Uncertainty on estimated and forecast quantities is then computed by a residual bootstrap approach. In this way we are able to simultaneously smooth past trends and forecast mortality in a sensible manner.

Results on two distinct populations in terms of mortality development (Danish females and US males) show that the CP -spline method performs well. Low deviance indicates that we are able to accurately capture past trends. Moreover, forecast mortality patterns and age profiles are reasonable given the past observed development. Fitted and forecast summary measures are also presented, and these results confirm that the proposed forecasting model performs remarkably well with very small errors in the observed period. Moreover, we show how the suggested CP -splines outperforms a smooth variant of the Lee-Carter model and, in the supplementary materials, a wider comparison shows the better performance of CP -splines with respect to other forecasting methods.

In the paper, we specified levels of confidence for rate of aging and rate of change aiming to constrain future mortality to observed shapes. These levels can be adapted to express diverse prior knowledge about future mortality for a distinct population and/or for specific ages. Specifically, our model might provide a means for guiding expert-based forecast approaches. Whereas it is hard to explicitly foresee future values for age-specific mortality rates, experts in the field may have more precise opinions on (un)feasible ranges of mortality improvements for ages, or groups of ages, looking to what has been observed in the past. This is obviously a population-specific procedure, and it will help to forecast populations that experienced exceptional patterns such as HIV epidemics, wars, and general mortality crises. We plan a generalization of the proposed *CP*-splines for these peculiar situations. Moreover, we envisage a large use of the *CP*-splines on all available and comparable mortality data. This will allow researchers to calibrate the choice of possible future age patterns and time trends based on broader past experiences, and eventually guide forecasters toward a more informed selection of possible constraints. In a Bayesian framework, and for forecasting cohort fertility, a similar approach was proposed by Schmertmann et al. (2014).

Specific cohort effects can be observed in certain populations and, in these situations, forecasters search for a procedure to incorporate these effects in the forecast horizon. Shape constraints over the diagonal of the mortality surface can be employed for accommodating peculiar cohort behaviors without influencing neighboring cohorts. Similarly, the proposed model may provide a flexible and rigorous approach for deriving the age pattern of mortality given a predicted level of life expectancy and/or lifespan variability measures. We plan to extend our model along both lines.

Finally, we think that the *CP*-spline model can be generalized to produce coherent forecasts for both sexes, or more populations, following a recent strand of research on mortality forecasting (Ahmadi and Li 2014; Bergeron-Boucher et al. 2016; Hyndman, Booth, and Yasmeen 2013; Li and Lee 2005; Shang 2016; Ševčíková et al. 2016). Constraining future mortality to observed age profiles and time trends can also be viewed as a way of constraining different subpopulations to behave analogously in terms of mortality shape and trend. This idea can be also adapted to estimate, and eventually forecast, mortality patterns for small areas where prior knowledge is often necessary to obtain reasonable outcomes. *CP*-splines would also be a useful tool for creating future mortality scenarios that borrow known mortality experiences from other populations and/or periods, and for eventually answering to the often relevant question “What if?” We shall pursue these ideas in future research.

6. Acknowledgments

I am grateful to Paul Eilers for his valuable suggestions about asymmetric penalties. I am also grateful to Iain Currie, Vladimir Canudas-Romo, Ugofilippo Basellini, Jutta Gampe, and Jim Oeppen for their insightful comments on earlier versions of the model and manuscript. I would also like to thank the anonymous reviewers for their valuable suggestions during the revision process. Presenting an earlier version of this paper to a P-splines Fiesta helped me to clarify and improve it.

References

- Abel, G., Bijak, J., Forster, J.J., Raymer, J., Smith, P.W., and Wong, J.S. (2013). Integrating uncertainty in time series population forecasts: An illustration using a simple projection model. *Demographic Research* 29(43): 1187–1226. doi:10.4054/DemRes.2013.29.43.
- Ahmadi, S.S. and Li, J.S.H. (2014). Coherent mortality forecasting with generalized linear models: A modified time-transformation approach. *Insurance: Mathematics and Economics* 59: 194–221. doi:10.1016/j.insmatheco.2014.09.007.
- Barrieu, P., Bensusan, H., El Karoui, N., Hillairet, C., Loisel, S., Ravanelli, C., and Salhi, Y. (2012). Understanding, modelling and managing longevity risk: key issues and main challenges. *Scandinavian Actuarial Journal* 2012(3): 203–231. doi:10.1080/03461238.2010.511034.
- Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J., and Vaupel, J.W. (2016). Coherent forecasts of mortality with compositional data analysis. *Demographic Research* 37(17): 527–566. doi:10.4054/DemRes.2017.37.17.
- Blake, D., Cairns, A.J.G., and Dowd, K. (2006). Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal* 12(1): 153–197. doi:10.1017/S1357321700004736.
- Bohk-Ewald, C., Ebeling, M., and Rau, R. (2017). Lifespan disparity as an additional indicator for evaluating mortality forecasts. *Demography* 54(4): 1559–1577. doi:10.1007/s13524-017-0584-0.
- Bohk-Ewald, C. and Rau, R. (2017). Probabilistic mortality forecasting with varying age-specific survival improvements. *Genus* 73(1): 1–37. doi:10.1186/s41118-016-0017-8.
- Bollaerts, K., Eilers, P.H.C., and van Mechelen, I. (2006). Simple and multiple P-splines regression with shape constraints. *British Journal of Mathematical and Statistical Psychology* 59: 451–469. doi:10.1348/000711005X84293.
- Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science* 3(1–2): 3–43. doi:10.1017/S1748499500000440.
- Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee–Carter under conditions of variable mortality decline. *Population Studies* 56: 325–336. doi:10.1080/00324720215935.
- Brouhns, N., Denuit, M., and Van Keilegom, I. (2005). Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal* 3: 212–224. doi:10.1080/03461230510009754.

- Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A., and Balevich, I. (2009a). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal* 13: 1–35. doi:10.1080/10920277.2009.10597538.
- Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics* 48: 355–367. doi:10.1016/j.insmatheco.2010.12.005.
- Cairns, A.J.G., Blake, D., and Dowd, K. (2006). Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin: The Journal of the IAA* 36(1): 79–120. doi:10.1017/S0515036100014410.
- Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I. (2009b). 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. doi:10.1080/10920277.2009.10597538.
- Cairns, A.J., Blake, D., and Dowd, K. (2008). Modelling and management of mortality risk: A review. *Scandinavian Actuarial Journal* 2008(2-3): 79–113. doi:10.1080/03461230802173608.
- Camarda, C.G. (2008). *Smoothing Methods for the Analysis of Mortality Development*. Ph.D. thesis, Programa de Doctorado en Ingeniería Matemática. Universidad Carlos III, Departamento de Estadística, Madrid.
- Camarda, C.G. (2012). MortalitySmooth: An R package for smoothing Poisson Counts with *P*-Splines. *Journal of Statistical Software* 50: 1–24. doi:10.18637/jss.v050.i01.
- Carfora, M.F., Cutillo, L., and Orlando, A. (2017). A quantitative comparison of stochastic mortality models on Italian population data. *Computational Statistics and Data Analysis* 112: 198–214. doi:10.1016/j.csda.2017.03.012.
- Carstensen, B. (2007). Age-Period-Cohort models for the Lexis diagram. *Statistics in Medicine* 26: 3018–3045. doi:10.1002/sim.2764.
- Chiang, C. (1984). *The Life Table and its Application*. Malabar: Krieger.
- Clayton, D. and Schifflers, E. (1987). Models for temporal variation in cancer rates. II. Age-period-cohort models. *Statistics in Medicine* 6: 469–481. doi:10.1002/sim.4780060406.
- Colchero, F., Rau, R., Jones, O.R., Barthold, J.A., Conde, D.A., Lenart, A., Nemeth, L., Scheuerlein, A., Schoeley, J., Torres, C., Zarulli, V., Altmann, J., Brockman, D.K., Bronikowski, A.M., Fedigan, L.M., Pusey, A.E., Stoinski, T.S., Strier, K.B., Baud-

- isch, A., Alberts, S.C., and Vaupel, J.W. (2016). The emergence of longevous populations. *Proceedings of the National Academy of Sciences* 113(48): E7681–E7690. doi:10.1073/pnas.1612191113.
- Currie, I. (2019). Constraints, the identifiability problem and the forecasting of mortality [forthcoming]. *The Annals of Actuarial Science* .
- Currie, I.D. (2011). Modelling and forecasting the mortality of the very old. *ASTIN Bulletin: The Journal of the IAA* 41: 419–427.
- Currie, I.D. (2013). Smoothing constrained generalized linear models with an application to the Lee–Carter model. *Statistical Modelling* 13: 69–93. doi:10.1177/1471082X12471373.
- Currie, I.D. (2016). On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal* 2016(4): 356–383. doi:10.1080/03461238.2014.928230.
- Currie, I.D., Durbán, M., and Eilers, P.H.C. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling* 4: 279–298. doi:10.1191/1471082X04st080oa.
- Currie, I.D., Durbán, M., and Eilers, P.H.C. (2006). Generalized linear array models with applications to multidimensional smoothing. *Journal of the Royal Statistical Society. Series B* 68: 259–280. doi:10.1007/978-1-4899-3242-6.
- D’Amato, V., Piscopo, G., and Russolillo, M. (2011). The mortality of the Italian population: Smoothing techniques on the Lee–Carter model. *The Annals of Applied Statistics* 5(2A): 705–724. doi:10.1214/10-AOAS394.
- de Boor, C. (1978). *A Practical Guide to Splines*. New York: Springer. doi:10.1007/978-1-4612-6333-3.
- Delwarde, A., Denuit, M., and Eilers, P.H.C. (2007). Smoothing the Lee–Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood approach. *Statistical Modelling* 7: 29–48. doi:10.1177/1471082X0600700103.
- Djeundje, V.A.B. and Currie, I.D. (2011). Smoothing dispersed counts with applications to mortality data. *Annals of Actuarial Science* 5: 33–52. doi:10.1017/S1748499510000047.
- Dowell, D., Noonan, R.K., and Houry, D. (2017). Underlying factors in drug overdose deaths. *Journal of the American Medical Association* 318(23): 2295–2296. doi:10.1001/jama.2017.15971.
- Efron, B. and Tibshirani, R.J. (1993). *An introduction to the bootstrap*. Chapman and Hall. doi:10.1007/978-1-4899-4541-9.

- Eilers, P.H.C. (2005). Unimodal Smoothing. *Journal of Chemometrics* 19: 317–328. doi:10.1002/cem.935.
- Eilers, P.H.C., Marx, B.D., and Durbán, M. (2015). Twenty years of P -splines. *SORT. Statistics and Operations Research Transactions* 39(2): 149–186.
- Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B -splines and penalties (with discussion). *Statistical Science* 11: 89–102. doi:10.1214/ss/1038425655.
- Erickson, R.V., Fabian, V., and Marik, J. (1995). An optimum design for estimating the first derivative. *The Annals of Statistics* 23: 1234–1247. doi:10.1214/aos/1176324707.
- Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J.L., Lalic, N., Bay, G., Buettner, T., Heilig, G.K., and Wilmoth, J. (2014). World population stabilization unlikely this century. *Science* 346: 234–237. doi:10.1126/science.1257469.
- Giroi, F. and King, G. (2007). Understanding the Lee–Carter mortality forecasting method. Santa Monica: RAND Corporation.
- Goicoa, T., Ugarte, M.D., Etxeberria, J., and Militino, A.F. (2012). Comparing CAR and P -spline models in spatial disease mapping. *Environmental and Ecological Statistics* 19(4): 1–27. doi:10.1007/s10651-012-0201-8.
- Goldstein, J.R. (2011). A secular trend toward earlier male sexual maturity: Evidence from shifting ages of male young adult mortality. Rostock: Max Planck Institute for Demographic Research (WP-2010-022). doi:10.1371/journal.pone.0014826.
- Gompertz, B. (1825). *On the nature of the function expressive of the law of human mortality*. London: Philosophical Transactions Royal Society. doi:10.1098/rstl.1825.0026.
- Grotenhuis, M.t., Pelzer, B., Luo, L., and Schmidt-Catran, A.W. (2016). The intrinsic estimator, alternative estimates, and predictions of mortality trends: A comment on Masters, Hummer, Powers, Beck, Lin, and Finch. *Demography* 53: 1245–1252. doi:10.1007/s13524-016-0476-8.
- Haberman, S. and Renshaw, A. (2009). On age–period–cohort parametric mortality rate projections. *Insurance: Mathematics and Economics* 45: 255–270. doi:10.1016/j.insmatheco.2009.07.006.
- Heuer, C. (1997). Modeling of time trends and interactions in vital rates using restricted regression splines. *Biometrics* 53: 161–177. doi:10.2307/2533105.
- Hilton, J., Dodd, E., Forster, J.J., and Smith, P.W.F. (2019). Projecting UK mortality by using Bayesian generalized additive models. *Journal of the Royal Statistical Society. Series A* 68: 29–49. doi:10.1111/rssc.12299.

- HMD (2019). Human Mortality Database [electronic resource]. Berkeley: University of California, Rostock: Max Planck Institute for Demographic Research. www.mortality.org.
- Holford, T.R. (2006). Approaches to fitting age–period–cohort models with unequal intervals. *Statistics in Medicine* 25: 977–993. doi:10.1002/sim.2253.
- Huang, F. and Browne, B. (2017). Mortality forecasting using a modified Continuous Mortality Investigation Mortality Projections Model for China I: Methodology and country-level results. *Annals of Actuarial Science* 11(1): 20–45. doi:10.1017/S1748499516000142.
- Hyndman, R.J. and Ullah, M.S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis* 51: 4942–4956. doi:10.1016/j.csda.2006.07.028.
- Hyndman, R.J., Booth, H., and Yasmeen, F. (2013). Coherent mortality forecasting: The product–ratio method with functional time series models. *Demography* 50: 261–283. doi:10.1007/s13524-012-0145-5.
- Jacobsen, R., Von Euler, M., Osler, M., Lyngge, E., and Keiding, N. (2004). Women’s death in Scandinavia: What makes Denmark different? *European Journal of Epidemiology* 19(2): 117–121. doi:10.1023/B:EJEP.0000017834.35943.bd.
- Jones, O.R., Scheuerlein, A., Salguero-Gomez, R., Camarda, C.G., Schaible, R., Casper, B.B., Dahlgren, J.P., Ehrlén, J., Garcia, M.B., Menges, E.S., Quintana-Ascencio, P.F., Caswell, H., Baudisch, A., and Vaupel, J.W. (2014). Diversity of ageing across the tree of life. *Nature* 505(7482): 169–173. doi:10.1038/nature12789.
- Keiding, N. (1990). Statistical inference in the Lexis diagram. *Philosophical Transactions: Physical Sciences and Engineering* 332: 487–509. doi:10.1098/rsta.1990.0128.
- Knorr-Held, L. and Rainer, E. (2001). Projections of lung cancer mortality in West Germany: A case study in Bayesian prediction. *Biostatistics* 2: 109–129. doi:10.1093/biostatistics/2.1.109.
- Koissi, M.C., Shapiro, A.F., and Högnäs, G. (2006). Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics* 38: 1–20.
- Kuang, D., Nielsen, B., and Nielsen, J. (2008). Forecasting with the age–period–cohort model and the extended chain-ladder model. *Biometrika* 95: 987–991. doi:10.1093/biomet/asn038.
- Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association* 87: 659–671. doi:10.1080/0162

[1459.1992.10475265](https://doi.org/10.1459.1992.10475265).

- Lee, R.D. and Miller, T. (2001). Evaluating the performance of the Lee–Carter method for forecasting mortality. *Demography* 38: 537–549. doi:[10.1353/dem.2001.0036](https://doi.org/10.1353/dem.2001.0036).
- Levitits, D.A. (2011). Before senescence : The evolutionary demography of ontogenesis. *Proceedings of the Royal Society B: Biological Sciences* 278(1707): 801–809. doi:[10.1098/rspb.2010.2190](https://doi.org/10.1098/rspb.2010.2190).
- Li, N. and Lee, R.D. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee–Carter method. *Demography* 42: 575–594.
- Li, N., Lee, R.D., and Gerland, P. (2013). Extending the Lee–Carter method to model the rotation of age patterns of mortality-decline for long-term projection. *Demography* 50: 2037–2051. doi:[10.1007/s13524-013-0232-2](https://doi.org/10.1007/s13524-013-0232-2).
- Lindahl-Jacobsen, R., Rau, R., Jeune, B., Canudas-Romo, V., Lenart, A., Christensen, K., and Vaupel, J.W. (2016). Rise, stagnation, and rise of Danish women’s life expectancy. *Proceedings of the National Academy of Sciences* 113(15): 4015–4020. doi:[10.1073/pnas.1602783113](https://doi.org/10.1073/pnas.1602783113).
- Lopez, A., Shibuya, K., Rao, C., Mathers, C., Hansell, A., Held, L., Schmid, V., and Buist, S. (2006). Chronic obstructive pulmonary disease: Current burden and future projections. *European Respiratory Journal* 27: 397–412. doi:[10.1183/09031936.06.00025805](https://doi.org/10.1183/09031936.06.00025805).
- Lu, J.L.C., Wong, W., and Bajekal, M. (2014). Mortality improvement by socio-economic circumstances in England (1982 to 2006). *British Actuarial Journal* 19(1): 1–35. doi:[10.1017/S1357321712000359](https://doi.org/10.1017/S1357321712000359).
- Mammen, E., Martínez-Miranda, M.D., and Nielsen, J.P. (2015). In-sample forecasting applied to reserving and mesothelioma mortality. *Insurance: Mathematics and Economics* 61: 76–86. doi:[10.1016/j.insmatheco.2014.12.001](https://doi.org/10.1016/j.insmatheco.2014.12.001).
- Martínez-Miranda, M.D., Nielsen, B., and Nielsen, J.P. (2014). Inference and forecasting in the age–period–cohort model with unknown exposure with an application to mesothelioma mortality. *Journal of the Royal Statistical Society. Series A* 178: 29–55. doi:[10.1111/rssa.12051](https://doi.org/10.1111/rssa.12051).
- Masters, R.K., Reither, E.N., Powers, D.A., Yang, Y.C., Burger, A.E., and Link, B.G. (2013). The impact of obesity on us mortality levels: The importance of age and cohort factors in population estimates. *American Journal of Public Health* 103(10): 1895–1901. doi:[10.2105/AJPH.2013.301379](https://doi.org/10.2105/AJPH.2013.301379).
- McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models*. Monographs on Statistics Applied Probability. London: Chapman and Hall, 2nd ed. doi:[10.1111/](https://doi.org/10.1111/)

[j.1467-9868.2006.00543.x](#).

- Minton, J., Shaw, R., Green, M.A., Vanderbloemen, L., Popham, F., and McCartney, G. (2017). Visualising and quantifying ‘excess deaths’ in Scotland compared with the rest of the UK and the rest of Western Europe. *Journal of Epidemiology and Community Health* 71(5): 461–467. doi:[10.1136/jech-2016-207379](#).
- Muennig, P.A. and Glied, S.A. (2010). What changes in survival rates tell us about US health care. *Health Affairs* 29(11): 2105–2113. doi:[10.1377/hlthaff.2010.0073](#).
- Nielsen, B. and Nielsen, J.P. (2014). Identification and forecasting in mortality models. *The Scientific World Journal* 2014: 1–24. doi:[10.1155/2014/347043](#).
- Ogata, Y., Katsura, K., Keiding, N., Holst, C., and Green, A. (2000). Empirical Bayes age–period–cohort analysis of retrospective incidence data. *Scandinavian Journal of Statistics* 27: 415–432. doi:[10.1111/1467-9469.00198](#).
- Ouellette, N. and Bourbeau, R. (2011). Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach. *Demographic Research* 25(19): 595–628. doi:[10.4054/DemRes.2011.25.19](#).
- Ouellette, N., Bourbeau, R., and Camarda, C.G. (2012). Regional disparities in canadian adult and old-age mortality: A comparative study based on smoothed mortality ratio surfaces and age-at-death distributions. *Canadian Studies in Population* 39(3-4): 79–106. doi:[10.25336/P61P53](#).
- Pitacco, E., Denuit, M., and Haberman, S. (2009). *Modelling longevity dynamics for pensions and annuity business*. Oxford: Oxford University Press.
- Preston, S.H. (1976). *Mortality patterns in national populations. With special reference to recorded causes of death*. Cambridge: Academic Press.
- R Development Core Team (2019). R: A Language and Environment for Statistical Computing [electronic resource]. Vienna: R Foundation for Statistical Computing. [R-project.org](#).
- Raftery, A.E., Chunn, J., Gerland, P., and Ševčíková, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography* 50: 777–801. doi:[10.1007/s13524-012-0193-x](#).
- Remund, A. (2015). Jeunesses vulnérables? Mesures, composantes et causes de la surmortalité des jeunes adultes [PhD Thesis]. Geneva: Université de Genève.
- Renshaw, A.E. and Haberman, S. (2003). Lee–Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* 33: 255–272. doi:[10.1016/S0167-6687\(03\)00138-0](#).

- Ribeiro, F. (2015). Statistical analysis and forecasting of cause of death data: Novel approaches and insights [PhD Thesis]. Évora: Universidade de Évora.
- Richards, S.J., Kirkby, J., and Currie, I.D. (2006). The importance of year of birth in two-dimensional mortality data. *British Actuarial Journal* 12: 5–61. doi:10.1017/S1357321700004682.
- Richards, S.J., Currie, I.D., and Ritchie, G.P. (2014). A value-at-risk framework for longevity trend risk. *British Actuarial Journal* 19(1): 116–139. doi:10.1017/S1357321712000451.
- Riebler, A. and Held, L. (2017). Projecting the future burden of cancer: Bayesian age–period–cohort analysis with integrated nested Laplace approximations. *Biometrical Journal* 59: 531–549. doi:10.1002/bimj.201500263.
- Schmertmann, C., Zagheni, E., Goldstein, J.R., and Myrskylä, M. (2014). Bayesian forecasting of cohort fertility. *Journal of the American Statistical Association* 109: 500–513. doi:10.1080/01621459.2014.881738.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* 6: 461–464. doi:10.1214/aos/1176344136.
- Shang, H.L. (2016). Mortality and life expectancy forecasting for a group of populations in developed countries: A multilevel functional data method. *The Annals of Applied Statistics* 10: 1639–1672. doi:10.1353/dem.2005.0021.
- Smith, T.R. and Wakefield, J. (2016). A review and comparison of age–period–cohort models for cancer incidence. *Statistical Science* 31: 591–610. doi:10.1214/16-STS580.
- Tabeau, E., Willekens, F., and van Poppel, F. (2002). Parameterisation as a tool in analysing age, period and cohort effects on mortality: A case study of the Netherlands. In: Wunsch, G., Mouchart, M., and Duchene, J. (eds.). *The life table: Modelling survival and death*. Dordrecht: Kluwer Academic Publishers: 141–169.
- Thatcher, R., Kannisto, V., and Vaupel, J.W. (1998). *The force of mortality at ages 80 to 120*, vol. 5 of *Monographs on Population Aging*. Odense: Odense University Press.
- Thun, M.J., Carter, B.D., Feskanich, D., Freedman, N.D., Prentice, R., Lopez, A.D., Hartge, P., and Gapstur, S.M. (2013). 50-year trends in smoking-Related Mortality in the United States. *New England Journal of Medicine* 368(4): 351–364. doi:10.1056/NEJMsa1211127.
- Trias-Llimós, S., Bijlsma, M.J., and Janssen, F. (2016). The role of birth cohorts in long-term trends in liver cirrhosis mortality across eight European countries. *Addiction* 112: 250–258. doi:10.1111/add.13614.
- Tzeng, I.S. and Lee, W.C. (2015). Forecasting hepatocellular carcinoma mortality in

- Taiwan using an age–period–cohort model. *Asia-Pacific Journal of Public Health* 27: NP65–NP73. doi:10.1177/1010539511422941.
- Ugarte, M.D., Goicoa, T., Etxeberria, J., and Militino, A.F. (2012). Projections of cancer mortality risks using spatio-temporal P -spline models. *Statistical Methods in Medical Research* 21(5): 545–560. doi:10.1177/0962280212446366.
- Ugarte, M.D., Goicoa, T., and Militino, A.F. (2010). Spatio-temporal modeling of mortality risks using penalized splines. *Environmetrics* 21(3–4): 270–289. doi:10.1002/env.1011.
- United Nations (2019). *World Population Prospects 2019: Data Booklet*. Department of Economic and Social Affairs, Population Division.
- Vaupel, J.W. and Canudas-Romo, V. (2003). Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. *Demography* 40: 201–216. doi:10.1353/dem.2003.0018.
- Ševčíková, H., Li, N., Kantorova, V., Gerland, P., and Raftery, A.E. (2016). Age-specific mortality and fertility rates for probabilistic population projections. In: Schoen, R. (ed.). *The Springer series on demographic methods and population analysis: Vol. 39. Dynamic demographic analysis*. Berlin: Springer: 285–310.
- Wang, H.C., Yue, J.C., and Tsai, Y.H. (2016). Marital status as a risk factor in life insurance: An empirical study in Taiwan. *ASTIN Bulletin: The Journal of the IAA* 46(2): 487–505. doi:10.1017/asb.2016.3.
- Wong, I.O.L., Cowling, B.J., Leung, G.M., and Schooling, C.M. (2013). Age-period-cohort projections of ischaemic heart disease mortality by socio-economic position in a rapidly transitioning Chinese population. *PLOS-One* 8: e61495. doi:10.1371/journal.pone.0061495.
- Zhang, Z. and Vaupel, J.W. (2009). The age separating early deaths from late deaths. *Demographic Research* 20(29): 721–730. doi:10.4054/DemRes.2009.20.29.