An age-at-death distribution approach to forecast cohort mortality

Ugofilippo Basellini*[1,2], Søren Kjærgaard[2], and Carlo Giovanni Camarda[1]

[1] *Institut national d'études démographiques (INED), Paris*
[2] *Interdisciplinary Centre on Population Dynamics (CPop), Department of Public Health, University of Southern Denmark, Odense*

October 21, 2019

Abstract

Mortality forecasting has received increasing interest during recent decades due to the negative financial effects of continuous longevity improvements on public and private institutions' liabilities. However, little attention has been paid to forecasting mortality from a cohort perspective. In this article, we introduce a novel methodology to forecast adult cohort mortality from age-at-death distributions. We propose a relational model that associates a time-invariant standard to a series of fully and partially observed distributions. Relation is achieved via a transformation of the age-axis. We show that cohort forecasts can improve our understandings of mortality developments by capturing distinct cohort effects, which might be overlooked by a conventional age-period perspective. Moreover, mortality experience of partially observed cohorts are routinely completed. We illustrate our methodology on adult female mortality for cohorts born between 1835 and 1970 in two high-longevity countries using data from the Human Mortality Database.

Keywords: Mortality forecasting · Mortality modelling · Relational models · Cohort life table · Smoothing

1 Introduction

Continuous and widespread gains in life expectancy (Riley, 2001; Oeppen and Vaupel, 2002) are increasingly challenging governments and insurance companies to provide adequate pension products and elderly health care in ageing societies. Mortality forecasting has thus gained relevant prominence during the last decades, as relatively small differences in the expected lifetimes of pensioners can have significant effects on financial institutions' liabilities.

A growing number of models have recently been proposed to forecast human mortality using stochastic methodologies that produce probabilistic assessments of the future. For comprehensive reviews, see Booth (2006) and Shang et al. (2011). The vast majority of these techniques are based on *period* mortality: financial institutions are typically interested in the mortality developments of groups of individuals that comprise different birth cohorts. In addition, cohort data can be outdated, unavailable or incomplete, hence period life tables

*Corresponding author: `ugofilippo.basellini@ined.fr`
Address: 133 boulevard Davout, 75020 Paris

have been developed to analyse a hypothetical cohort if its age-specific death rates pertained throughout its life (Preston et al., 2001).

The completion of the mortality experience of non-extinct cohorts is nonetheless interesting in the actuarial domain. Insurance companies are indeed interested in the future longevity of groups of people born in specific cohorts. In such settings, cohort forecasts are typically obtained by first forecasting mortality in a period fashion, and then extracting cohort mortality patterns from the diagonals of the projected Lexis surface. Although widely used, this approach seems rather counter-intuitive and inefficient. In this article, we propose a more direct and alternative approach to cohort mortality forecasting that is solely based on cohort data.

More generally, analysis and forecasts of cohort mortality are interesting and worth exploring for two main reasons. First, survival in real birth cohorts is different from survival in the hypothetical situation of unchanged period mortality rates because of: (i) tempo effects, (ii) cohort effects and (iii) selection (for a full discussion, see Borgan and Keilman, 2019, pp. 90–92). Second, cohort mortality developments are *actually* observed, and they may differ from those of the synthetic cohorts assumed in period life tables. In other words, cohort life tables record information about what happened to real birth cohorts, whereas period life tables are models of what would happen to hypothetical cohorts if their age-specific period death rates remained constant throughout the life of the cohorts (Preston et al., 2001).

Analyses of age-cohort data have indeed provided different insights on mortality developments than studies based on the age-period perspective. For example, Shkolnikov et al. (2011) have shown that best-practice cohort life expectancy for women born between 1870 and 1920 increased almost twice as fast as best-practice period life expectancy since 1840. Moreover, Goldstein and Wachter (2006) suggested that, in populations experiencing steady mortality declines, period life expectancy can be regarded as a lagged indicator of cohort life expectancy, and that the lag amounts to about 40–50 years for women in the USA. Finally, Borgan and Keilman (2019) have shown that the differences in period life expectancy between Japanese and Italian women on one side, and Scandinavian ones on the other, disappear or even reverse when considering cohort data: as such, they contend that these differences are caused by the distortion that period life tables imply in times of changing mortality. Given these considerations, forecasting cohort mortality from age-cohort data seems a more reasonable approach than extracting cohort patterns from age-period projections.

Models to forecast cohort mortality are relatively few in the literature. Among the firsts to use a cohort perspective, the Continuous Mortality Investigation (2007) employed the two-dimensional P-spline model of Currie et al. (2004) to complete the mortality experience of cohorts in England & Wales. Furthermore, Chiou and Müller (2009) proposed a functional data analysis approach for forecasting cohort log-hazard functions using Swedish mortality data. More recently, the combination of an EM algorithm with an eight-parameter model for the age-at-death distribution was suggested by Zanotto and Mazzuco (2017) for estimating deaths of non-extinct generations. Finally, Rizzi et al. (2019) proposed to complete partially observed cohort age-at-death distributions using a penalized composite link model (Eilers, 2007), assuming a smooth underlying distribution over age.

In this article, we introduce a novel methodology to forecast adult cohort mortality. Rather than modelling mortality rates (the standard approach in mortality forecasting, as in, for example, the Lee and Carter model and its variants), our model is based on the distribution of deaths. Age-at-death distributions have recently received increasing attention in mortality forecasting (Oeppen, 2008; Bergeron-Boucher et al., 2017; Basellini and Camarda, 2019a,b; Pascariu et al., 2019), as they provide a different and rather unexplored perspective on mor-

tality developments that can be leveraged by forecasters. For this reason, we extend a newly introduced methodology to model and forecast adult age-at-death distributions (Basellini and Camarda, 2019b) with the aim of analyzing and forecasting mortality developments across cohorts.

This paper is organized as follows. In Section 2, we describe the structure of cohort mortality data, and we illustrate mortality developments for one population analysed in this article from the age-cohort perspective. In Section 3, we review the methodology proposed in this article, and the data that we employ for the analyses. Section 4 presents two applications of our model to Swedish and Danish female adult mortality for the cohorts 1835–1970. In Section 5, we discuss the results of our methods and conclude.

2 Background

The structure of cohort mortality data differs from the conventional one of age-period mortality analysis. Before introducing our methodology, it is therefore convenient to describe the structure of the data employed in this article, and to analyse mortality developments from the age-cohort perspective considered here.

Let $d_{x,c}$ and $e_{x,c}$ denote observed death counts and central exposures to the risk of death at age x for the birth cohort c, respectively. Data are arranged into two matrices $\boldsymbol{D} = (d_{x,c})$ and $\boldsymbol{E} = (e_{x,c})$, each of dimensions $m \times n$, where rows are classified by m single ages at death $\boldsymbol{x}' = [1, \ldots, m]$, and columns by n single cohorts $\boldsymbol{c}' = [1, \ldots, n]$. Unlike the case of age-period data, where data are generally fully observed, here the matrices $\boldsymbol{D}$ and $\boldsymbol{E}$ contain missing data, corresponding to periods (i.e. calendar years) beyond the last available year of data collection.

Displaying the matrix $\boldsymbol{D}$ provides a clear understanding of the structure of cohort mortality data. Let $\breve{x}$ and $\breve{c}$ be the last age and cohort with fully observed data, respectively. Then, death counts are available as follows, where "na" denotes the unobserved data:

$$\boldsymbol{D} = (d_{x,c}) = \begin{bmatrix} d_{1,1} & \ldots & d_{1,\breve{c}} & d_{1,\breve{c}+1} & d_{1,\breve{c}+2} & \ldots & d_{1,n} \\ d_{2,1} & \ldots & d_{2,\breve{c}} & d_{2,\breve{c}+1} & d_{2,\breve{c}+2} & \ldots & d_{2,n} \\ \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ d_{\breve{x},1} & \ldots & d_{\breve{x},\breve{c}} & d_{\breve{x},\breve{c}+1} & d_{\breve{x},\breve{c}+2} & \ldots & d_{\breve{x},n} \\ d_{\breve{x}+1,1} & \ldots & d_{\breve{x}+1,\breve{c}} & d_{\breve{x}+1,\breve{c}+1} & d_{\breve{x}+1,\breve{c}+2} & \ldots & na \\ \vdots & \ldots & \vdots & \vdots & \vdots & \iddots & \vdots \\ d_{m-2,1} & \ldots & d_{m-2,\breve{c}} & d_{m-2,\breve{c}+1} & d_{m-2,\breve{c}+2} & \iddots & na \\ d_{m-1,1} & \ldots & d_{m-1,\breve{c}} & d_{m-1,\breve{c}+1} & na & \ldots & na \\ d_{m,1} & \ldots & d_{m,\breve{c}} & na & na & \ldots & na \end{bmatrix}.$$

Fully observed — Partially observed

The matrix clearly shows that, for the cohorts after $\breve{c}$, data are increasingly missing from the last age downwards. Exposure data in $\boldsymbol{E}$ are unobserved for the same elements of $\boldsymbol{D}$.

The analysis of mortality developments over ages and cohorts will therefore inevitably display this data structure. Figure 1 shows an example of such structure: observed adult age-at-death distributions for Swedish females are illustrated for selected cohorts between 1835–1970 (data retrieved from the Human Mortality Database, 2019). For this population, the last cohort with fully observed data is 1906; more recent cohorts are thus characterized by a decreasing amount of observed data, as shown by the graphs.

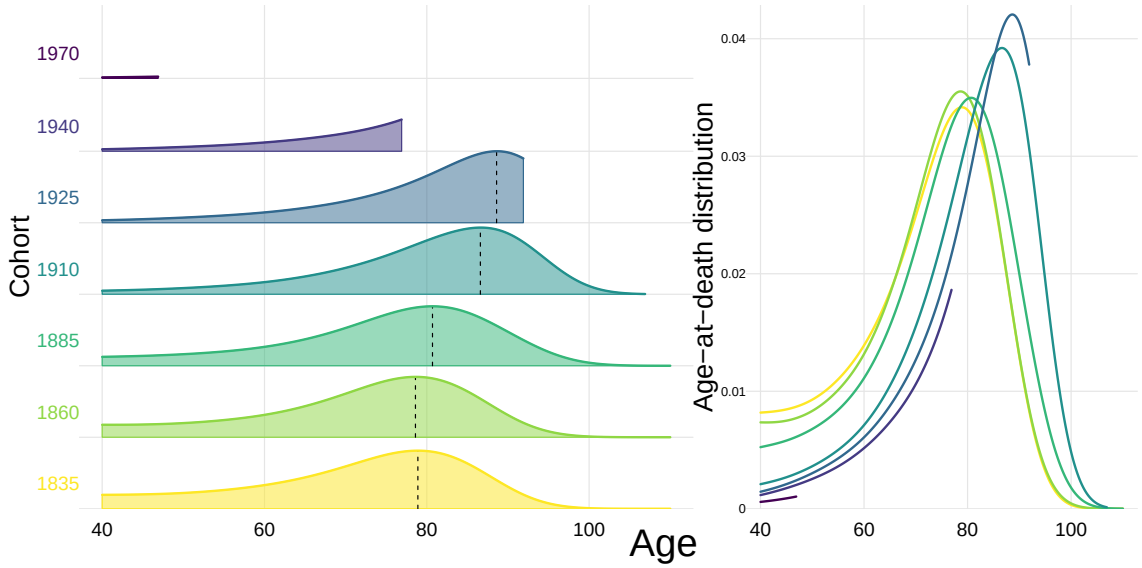

Figure 1. Observed age-at-death distributions for Swedish females aged 40–110+ for selected cohorts between 1835–1970. Dashed black lines correspond to modal ages at death. Data have been smoothed for illustrative purposes.
Source: authors' elaborations on data from the Human Mortality Database (2019).

Importantly, Figure 1 further motivates the development of a methodology that can leverage the features and the changes of age-at-death distributions with the purpose of modelling and forecasting human mortality. Clearly intelligible and demographically meaningful mortality developments readily emerge from this Figure. The left panel shows the increase of the modal age at death for the cohorts born after the 1860. In the right panel, the same distributions of the left panel are plotted over a common y-axis: decreasing age-at-death variability is evident for more recent cohorts.

Age-at-death distributions thus provide informative insights on mortality patterns and developments of the population, and they further allow to study the shifting and compression dynamics of mortality changes (see, e.g., Fries, 1980; Kannisto, 2001; Bongaarts, 2005; Janssen and de Beer, 2019). We therefore propose an approach to model and forecast adult cohort mortality that is based on age-at-death distributions. A relational model inspired by the seminal contribution of Brass (1971) serves our purposes well. Specifically, the combination of a reference distribution and its transformation over different cohorts allows us to simultaneously: (i) fit the observed data, and (ii) obtain an estimation for the unobserved data, thereby completing the mortality experience of partially observed cohorts. In the following Section, we introduce our proposed methodology.

3 Methods & Data

3.1 The C-STAD model

Suppose we have two adult age-at-death distributions defined on the age range $x \geq 40$. Specifically, let $f(x)$ be a "standard", or reference, distribution and $g(x)$ an observed distribution. Let $t(x; \boldsymbol{\theta})$ be a transformation function of the age axis and a vector of parameters $\boldsymbol{\theta}$ such

that:

$$g(x) = f\left[t(x;\boldsymbol{\theta})\right] , \tag{1}$$

i.e. the distribution $g(x)$ is derived from a warping transformation of the age axis of $f(x)$. In particular, the term "warping" originates in Functional Data Analysis (Ramsay and Silverman, 2005) and refers to the transformation of a time axis to achieve close alignment of functions.

We propose a parsimonious yet flexible transformation function $t(x;\boldsymbol{\theta})$ that captures adult mortality developments across cohorts rigorously. Let $\boldsymbol{\theta}' = [s, b_L, c_L, d_L, b_U]$ be a vector containing the model's parameters, where $s = M^g - M^f$ is the difference between the modal ages at death of the unimodal distributions $g(x)$ and $f(x)$. The proposed *Cohort Segmented Transformation Age-at-death Distributions* (C-STAD) model can be written as:

$$t(x;\boldsymbol{\theta}) = \begin{cases} M^f + b_L\,\tilde{x} + c_L\,\tilde{x}^2 + d_L\,\tilde{x}^3 & \text{if } x \leq M^g \\ M^f + b_U\,\tilde{x} & \text{if } x > M^g \end{cases} \tag{2}$$

where $\tilde{x} = x - s - M^f$, and the subscripts $_L$ and $_U$ refer to the *Lower* and *Upper* parts of the segmented transformation (i.e. before and after M^g), respectively.

The warping function $t(x;\boldsymbol{\theta})$ takes the form of a segmented transformation model which breaks at the value of M^g. Below M^g, the transformation function is cubic, while it is linear above M^g. Although acting on $t(x;\boldsymbol{\theta})$, the model's parameters are directly related to the summary measures of the age-at-death distributions. Specifically, while s captures the difference in modal ages between $g(x)$ and $f(x)$, b_L and b_U measure the change in variability before and after the modal ages of the two distributions. For the ages below M^g, c_L and d_L further measure differences in terms of asymmetry and heaviness of the left tail between $g(x)$ and $f(x)$, respectively. In terms of mathematical moments, the parameters b_L and b_U can be related to the variance of the distribution before and after the mode, while c_L and d_L to the skewness and kurtosis of the distribution.

Figure 2 provides a graphical illustration of the C-STAD model. For ease of presentation, we start from the simpler case in which $b_L = b_U = 1$ and $c_L = d_L = 0$. Substitution of these parameters in Eq. (2) yields a unique transformation function $t(x;\boldsymbol{\theta}) = x - s$, and a corresponding distribution $g_1(x) = f(x - s)$ via Eq. (1). In the left panel of Fig. 2, the standard distribution (grey line) is shifted to the right by an amount equal to s, and $g_1(x)$ (blue line) maintains the original shape of $f(x)$. The right panel shows the transformation function related to this plain shifting scenario. Note that a left-shift could be simply obtained with a negative value of s.

Different parameters' values allow to capture broader mortality developments than the shifting scenario described above. While b_L and b_U modify the variability of the distribution $g_2(x) = f\left[t(x;\boldsymbol{\theta})\right]$ (orange line, left panel) w.r.t. $f(x)$ before and after the modal age at death, c_L and d_L affect the asymmetry and heaviness of the left tail of $g_2(x)$ as compared to $f(x)$. In the example shown in Fig. 2, $b_L > 1$ reduces the variability of $g_2(x)$ before M^g w.r.t. $f(x)$, while $b_U < 1$ increases the variability of $g_2(x)$ after M^g w.r.t. $f(x)$. The effects of c_L and d_L are difficult to discern from the left panel. However, the right panel shows the warping transformation $t(x;\boldsymbol{\theta})$ applied to $f(x)$ to derive $g_2(x)$; the transformation (orange line) is composed by a cubic function (due to non-zero values of c_L and d_L) before the cut-off point M^g, and by a linear function above M^g.

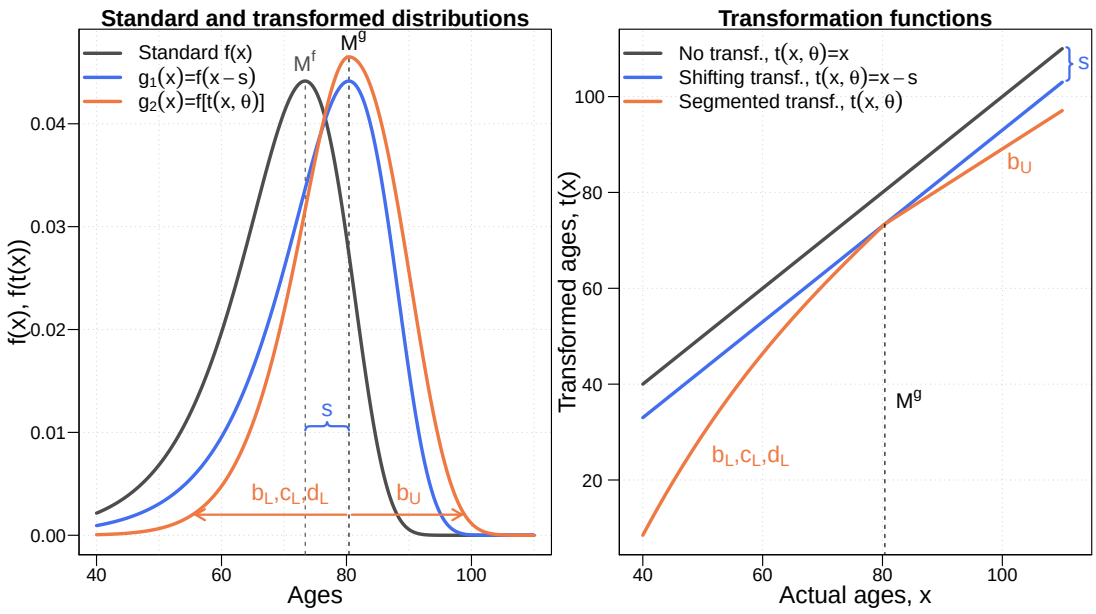

Figure 2. A schematic overview of the *Cohort Segmented Transformation Age-at-death Distributions* model.
Source: authors' own elaborations.

3.2 Data

For illustrative purposes we present outcomes from the proposed model on adult cohort mortality for females in two high-longevity countries, namely Sweden and Denmark. Long-term series of high quality data are available for both countries, even at the very old ages (Vaupel and Lundström, 1994; Wilmoth and Lundström, 1996; Andreev, 2002), and the two countries display different mortality developments. We therefore test the goodness-of-fit and forecast accuracy of our model with respect to different mortality trajectories (Christensen et al., 2010). The data are derived from the Human Mortality Database (HMD, 2019), which provides free access to detailed, consistent and high quality historical mortality data for 41 different areas and countries (Barbieri et al., 2015).

Our interest in this article is restricted to the senescent component of mortality, hence we start our analyses from age 40. We therefore cover the age range that is of greater interest for pension and insurance funds. Specifically, we work with two $m \times n$ matrices $\boldsymbol{D} = (d_{x,c})$ and $\boldsymbol{E} = (e_{x,c})$, defined for ages $\boldsymbol{x'} = [40, \ldots, 110+]$ and cohorts $\boldsymbol{c'} = [1835, \ldots, 1970]$. Figure 3 offers a schematic overview of the data structure by means of two Lexis diagrams: the first (left panel) shows the conventional age-period structure, while the second (right panel) illustrates the age-cohort perspective that we adopt in this paper. On one hand, we select 1835 as starting cohort of analysis for both populations because it is the first cohort with observed data at all ages in Denmark, and to compare the results for the two countries using the same fitting period. On the other hand, 1970 is the final cohort because it contains enough observed data points in both countries (seven in Sweden, six in Denmark) to estimate the three low parameters accurately (cf. Subsection 3.4).

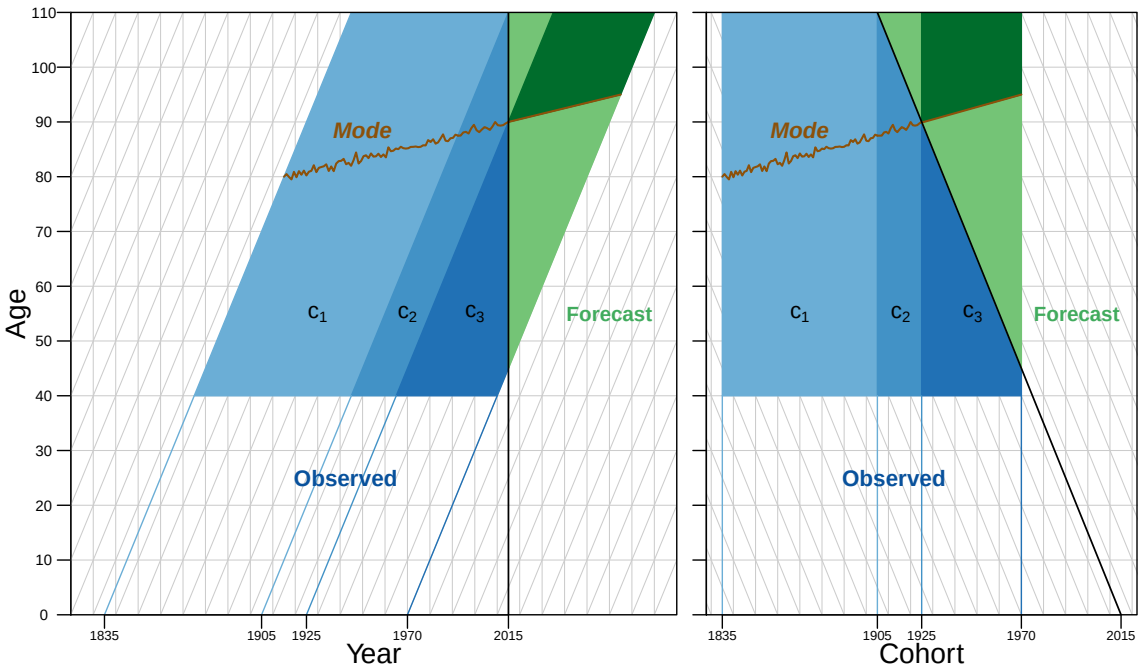

Figure 3. Conventional age-period (left panel) and age-cohort (right panel) Lexis diagrams illustrating the data structure and the division of cohorts into three groups. Here we assume that: (i) 2015 is the most recent year of data collection, (ii) $\check{c} = 1905$, and (iii) $\tilde{c} = 1925$. The three groups are then $\boldsymbol{c}_1' = [1835, \ldots, 1905]$, $\boldsymbol{c}_2' = [1906, \ldots, 1925]$ and $\boldsymbol{c}_3' = [1926, \ldots, 1970]$. The two colours in the forecast years correspond to different parameters' derivations (cf. Subsection 3.4): estimation with missing data (light green) and forecasting (dark green).
Source: authors' own elaborations.

Estimation and forecasting of the C-STAD parameters (Subsection 3.4) is performed on three different groups of cohorts:

$$\boldsymbol{c}' = [\underbrace{1835, \ldots, \check{c}}_{\boldsymbol{c}_1}, \underbrace{\check{c}+1, \ldots, \tilde{c}}_{\boldsymbol{c}_2}, \underbrace{\tilde{c}+1, \ldots, 1970}_{\boldsymbol{c}_3}] \,. \tag{3}$$

Therefore data are partitioned as follows:

$$\boldsymbol{D} = [\boldsymbol{D}_1 : \boldsymbol{D}_2 : \boldsymbol{D}_3] \qquad \boldsymbol{E} = [\boldsymbol{E}_1 : \boldsymbol{E}_2 : \boldsymbol{E}_3] \,. \tag{4}$$

The first group, denoted by $\boldsymbol{c}_1$, contains the fully observed cohorts $1835, \ldots, \check{c}$, where $\check{c}$ corresponds to the last cohort for which all data have been observed. As such, $\boldsymbol{D}_1$ and $\boldsymbol{E}_1$ have been observed at all ages x for all cohorts in $\boldsymbol{c}_1$. The second group, denoted by $\boldsymbol{c}_2$, is composed by cohorts $\check{c}+1, \ldots, \tilde{c}$, where $\tilde{c}$ corresponds to the last cohort for which two age-groups above the adult modal age at death have been observed. In other words, $\boldsymbol{D}_2$ and $\boldsymbol{E}_2$ are incomplete, i.e. data are not available for higher ages and more recent cohorts. However this group of cohorts is selected such that associated $d_{x,c}$ and $e_{x,c}$ have been observed for at least two data points above the modal age $x = M$ for all cohorts in $\boldsymbol{c}_2$. The choice of having two age groups above M is imposed by the estimation of the parameter above the mode (cf. Subsection 3.4). Finally, the third group $\boldsymbol{c}_3$ is composed by the remaining cohorts $\tilde{c}+1, \ldots, 1970$, in which data are only partially available and modal age at death is not observed. An illustration of the divisions of cohorts into the three groups is provided in Figure 3. Figure 4 shows an example of the observed and missing data for three age-at-death distributions belonging to the different groups of cohorts.

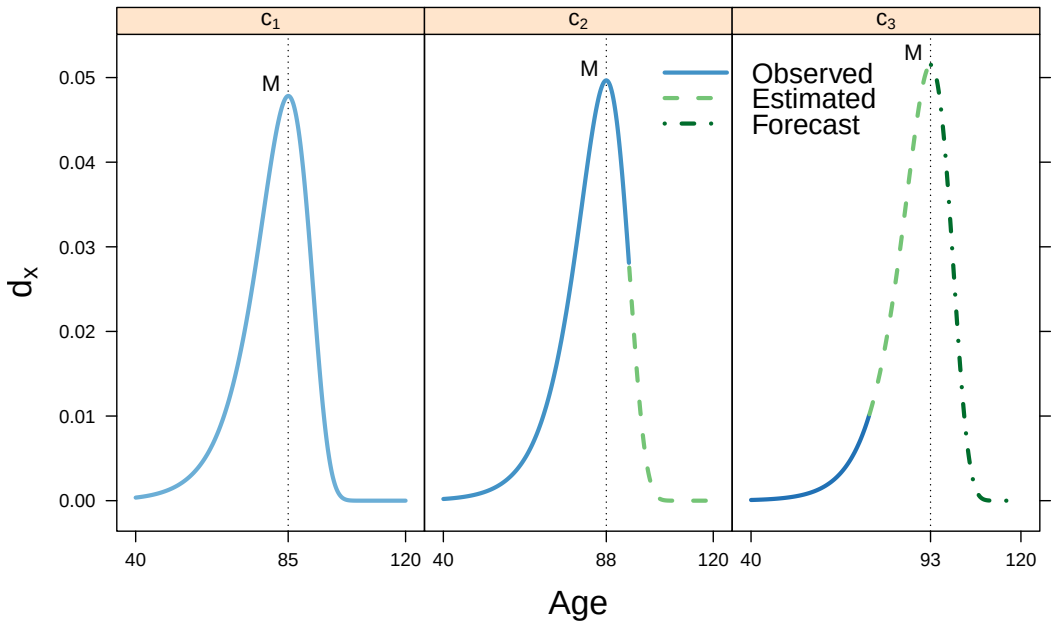

Figure 4. Example of observed, estimated and forecast data for three age-at-death distributions belonging to the groups of cohorts $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ and $\boldsymbol{c}_3$.
Source: authors' own elaborations.

3.3 The standard distribution

The first step in the estimation of the C-STAD model is the derivation of the standard distribution $f(x)$. The C-STAD can be interpreted as a relational model (Brass, 1971), hence it is desirable to include the representative features of the observed data for all cohorts in the computation of $f(x)$. Meanwhile, we also wish to remove the small random fluctuations that characterize the mortality pattern of age-at-death distributions. To achieve both goals at the same time, we derive the age-at-death distribution for each cohort 1835–1970 by a two-dimensional (2D) P-splines smoothing approach of cohort mortality (Eilers and Marx, 1996; Currie et al., 2004). Specifically, we assume that observed death counts $d_{x,c}$ at given age x and cohort c are realizations of the random variable $D_{x,c}$ which follows a Poisson distribution (Brillinger, 1986):

$$D_{x,c} \sim \mathcal{P}(e_{x,c}\,\mu_{x,c})\,, \tag{5}$$

where exposure-to-risk $e_{x,c}$ are given and $\mu_{x,c}$ denotes the hazard or force of mortality (such as in, for example, Brouhns et al., 2002). We smooth observed death counts using a tensor product of B-splines bases over ages and cohorts, and exposures as an offset. To account for uncompleted cohorts and still preserve the rectangular structure in the data, we include regression weights $\boldsymbol{W} = (w_{x,c})$ whose elements are equal to one if the corresponding death counts $d_{x,c}$ and exposures $e_{x,c}$ have been observed, and zero otherwise. Smoothing parameters over ages and cohorts are chosen by Bayesian Information Criterion minimization. The `R` package `MortalitySmooth` (Camarda, 2012) provides a direct implementation of this procedure.

The estimated smooth mortality surface allows us to derive smooth (partial) distributions. To compute the standard, we first employ a landmark registration procedure, a technique often used in Functional Data Analysis with the aim of aligning important features, or landmarks, of the curves analysed (Ramsay and Silverman, 2005). For age-at-death distributions, the

modal age at death is an obvious landmark. As such, we align the distributions so that their modal ages are equal to the mode of the first observed distribution (the 1835 cohort), maintaining their variability unchanged. In practice, the alignment is achieved by a plain shifting transformation of the observed distributions, which preserve all their features except the modal age-at-death.

Having aligned the observed distributions to a common modal age, we derive the standard $f(x)$ as the mean of the aligned distributions. Figure B.1 in Appendix B illustrates the alignment procedure and the derivation of the standard distribution for the Swedish female population analysed in our article. This landmark registration procedure has been suggested elsewhere as it enhances the representativeness of $f(x)$ while improving the goodness-of fit of the model (for additional details, see Basellini and Camarda, 2019b). Importantly, it should be noted that for cohorts in $\boldsymbol{c}_2$ and $\boldsymbol{c}_3$, we only use the part of the aligned distribution corresponding to the observed data (i.e. where regression weights are not zero).

Finally, we express the standard $f(x)$ as a linear combination of equally spaced B-spline bases $\boldsymbol{B}(x)$ over ages x and coefficients $\boldsymbol{\beta}_f$ specific to the standard:

$$f(x) = \exp\left[\boldsymbol{B}(x)\,\boldsymbol{\beta}_f\right]. \tag{6}$$

In this last step, we chose a generous number of B-splines with a very small penalty term. This whole procedure allows us to preserve all important features in the standard distribution (embodied in $\boldsymbol{\beta}_f$) after having removed unnecessary random fluctuations from the original data. In addition, the smoothing approach further allows us to evaluate mortality at any finer scale of the age axis, practically at a continuous level.

3.4 Estimation and forecast of the C-STAD parameters

Given the estimated standard distribution, we can derive the C-STAD parameters $\boldsymbol{\theta}' = [s, b_L, c_L, d_L, b_U]$ for each cohort in 1835–1970. As anticipated in Subsection 3.2, we use three different approaches to estimate $\boldsymbol{\theta}$, depending on the data available for each cohort; we thus divide cohorts into three groups, as shown in Figure 3.

For the first group of fully observed cohort $\boldsymbol{c}_1$, we start by estimating the parameters vector $\boldsymbol{s}$. To properly capture cohort-specific mortality fluctuations, we employ a one-dimensional P-spline approach, i.e. we smooth mortality for each cohort independently, numerically compute the corresponding density and extract the modal ages at death for each cohort (for a similar approach in a period perspective, see Ouellette and Bourbeau, 2011). From the modal ages we estimate the parameter $\hat{\boldsymbol{s}} = \left(\hat{s}_c = M_c - M_f\right)$ over cohorts in $\boldsymbol{c}_1$, where M_f denotes the mode of the standard distribution, which by construction corresponds to the modal age at death for the cohort born in 1835.

Having derived an estimate of the shifting parameter $\hat{\boldsymbol{s}}$, we can estimate the remaining parameters $\boldsymbol{\alpha}' = [b_L, c_L, d_L, b_U]$. We take advantage of the Poisson assumption in Eq. (5) and maximise the following log-likelihood function:

$$\ln \mathcal{L}_{\boldsymbol{\alpha}}\left(\boldsymbol{\alpha} \,|\, d_{x,c}, e_{x,c}, w_{x,c}, \hat{s}_c, \boldsymbol{\beta}_f\right) \propto \sum_x w_{x,c}\left[d_{x,c} \ln\left(\mu_{x,c}^{\texttt{C-STAD}}\right) - e_{x,c}\,\mu_{x,c}^{\texttt{C-STAD}}\right] \tag{7}$$

for each cohort in $\boldsymbol{c}_1' = [1835, \ldots, \check{c}]$, where regression weights $w_{x,c}$ are zero in the case of unobserved data at the highest ages, and $\mu_{x,c}^{\texttt{C-STAD}}$ denotes the estimated hazard of the C-STAD model. In words, the optimization procedure looks for a combination of parameters $\hat{\boldsymbol{\alpha}}$ that produces, for each cohort, an age-at-death distribution whose corresponding hazard

maximises the log-likelihood in Eq. (7). The associated age-at-death distribution $\hat{g}_c(x)$ can be written as follows:

$$\hat{g}_c(x) = \exp\left[B(x_t)\boldsymbol{\beta}_f\right] \qquad \text{where} \quad x_t = t(x; \hat{s}_c, \hat{\boldsymbol{\alpha}})\,. \tag{8}$$

The hazard $\hat{\mu}_c(x)$ corresponding to $\hat{g}_c(x)$ is computed using standard life-table and survival analysis formulas:

$$\hat{\mu}_c(x) = \frac{\hat{g}_c(x)}{\hat{\ell}_c(x)} = \frac{\hat{g}_c(x)}{\int_x^\infty \hat{g}_c(t)\,dt}\,, \tag{9}$$

where $\hat{\ell}_c(x)$ denotes the life-table probability of surviving to age x (i.e. the survival function, Preston et al., 2001; Klein and Moeschberger, 2003), which can be computed numerically from the smooth distribution $\hat{g}_c(x)$, evaluated at an extremely fine level.

For the second group of partially observed cohorts $\boldsymbol{c}_2$, we start again from the shifting parameter $\boldsymbol{s}$. We use the same estimation approach used in $\boldsymbol{c}_1$: data are available until the ages above the mode, therefore the smoothing approach produces an estimate of M_c and $\hat{\boldsymbol{s}}$ over cohorts in $\boldsymbol{c}_2$. With respect to the remaining parameters, we also follow the same approach: we maximize Eq. (7) for each cohort in $\boldsymbol{c}_2' = [\check{c}+1, \ldots, \tilde{c}]$, the only difference being that zero regression weights correspond to the missing data above the mode of the partially observed cohorts. It should be noted here that the missing data only influence the estimation of b_U, as complete data are observed below the mode for all cohorts in this group.

For the third group of partially observed cohorts $\boldsymbol{c}_3$, we employ a mixture of forecasting and estimation to determine the C-STAD parameters. The lack of data above the modal age at death makes it impossible to estimate the parameter $\boldsymbol{s}$ and compute the log-likelihood in Eq. (7). Hence, we start from the time-series of the estimated parameters $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{b}}_U$ over cohorts in $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ to compute their forecasts for cohorts $\boldsymbol{c}_3$.

From a theoretical perspective, these two parameters are related by the fact that only mortality changes occurring above the mode can modify its value (cf. Appendix B in Canudas-Romo, 2010). Correlation analyses for the two countries under study confirm the strong relation between the two series (Pearson correlation of 0.96 and 0.90 for the time-series in first differences for Sweden and Denmark, respectively). As such, we specify a vector autoregressive (VAR) model of order one with constant for the two (differenced) parameters, and we forecast their values for all cohorts $\boldsymbol{c}_3$. The `R` package `vars` allows us to perform model selection and estimation (Pfaff, 2008a,b).

Then, we take the forecast values of $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{b}}_U$ as given, and we estimate the remaining parameters $\check{\boldsymbol{\alpha}}' = [b_L, c_L, d_L]$ by maximizing the log-likelihood:

$$\ln \mathcal{L}_{\check{\boldsymbol{\alpha}}}\left(\check{\boldsymbol{\alpha}} \,|\, d_{x,c}, e_{x,c}, w_{x,c}, \hat{s}_c, \hat{b}_{U_c}, \boldsymbol{\beta}_f\right) \propto \sum_x w_{x,c}\left[d_{x,c}\ln\left(\mu_{x,c}^{\texttt{C-STAD}}\right) - e_{x,c}\,\mu_{x,c}^{\texttt{C-STAD}}\right] \tag{10}$$

for each cohort in $\boldsymbol{c}_3' = [\tilde{c}+1, \ldots, 1970]$. In contrast to the estimation procedure in $\boldsymbol{c}_2$, here the missing data influence the estimation of the parameters associated to the ages below the modal age at death.

The estimate $\hat{\boldsymbol{\theta}}$ for each cohort in 1835–1970 allows us to derive a complete set of age-specific mortality measures, i.e. we can complete the mortality experience for the partially observed cohorts of our analysis. In order to derive the C-STAD confidence intervals (CI)[1], we employ a bootstrapping procedure (Efron and Tibshirani, 1994). As suggested by Keilman and Pham

[1] to avoid confusion, we use the general term CI for all cohorts analysed, even when intervals are constructed from the mixture of forecast and estimated parameters (i.e. cohorts $\boldsymbol{c}_3$).

(2006), we consider the uncertainty related to: (i) the estimated parameters, and (ii) the forecast values of $\boldsymbol{s}$ and $\boldsymbol{b}_U$. The first source of uncertainty is accounted for by generating bootstrap death counts from the C-STAD deviance residuals (as in, for example, Koissi et al., 2006; Renshaw and Haberman, 2008; Ouellette et al., 2012). Appendix A provides more details on the computation of deviance residuals and bootstrap death counts. The second source of uncertainty is considered by simulating future values of the VAR model. We employ 40 different matrices of bootstrap death counts, and for each of these, we refit the C-STAD model and simulate 40 future values of $\boldsymbol{s}$ and $\boldsymbol{b}_U$. From the 1600 resulting simulations, we take the lower and higher deciles to construct 80% pointwise confidence intervals.

Finally, routines developed to fit and forecast the C-STAD model were implemented in `R` (R Development Core Team, 2018) and are publicly available, and all the results presented in the following Section are fully reproducible at `https://github.com/ubasellini/C-STAD` [this GitHub repository will be made public upon eventual acceptance of the manuscript].

4 Results

4.1 Out-of-sample validation of the C-STAD model

Before estimating the proposed C-STAD model to complete partially observed cohorts, we first assess the accuracy of the C-STAD model by performing six predictive out-of-sample validation exercises on Swedish and Danish adult females. We compare the forecast life expectancy at age 40 (e_{40}) and the Gini coefficient at age 40 (G_{40}) with the observed out-of-sample values. Both measures of longevity (the former) and lifespan inequality (the latter) are indeed useful to evaluate the accuracy of mortality forecasts (Bohk-Ewald et al., 2017). Formally, the Gini coefficient at age 40 is defined as:

$$G_{40} = 1 - \frac{1}{e_{40}\,(\ell_{40})^2} \int_{40}^{\omega} [\ell(x)]^2\, dx\,, \tag{11}$$

where ℓ_{40} is the life-table radix, which we let equal to one without loss of generality, and ω is the highest age attained in the population (Hanada, 1983; Shkolnikov et al., 2003).

Originally proposed to measure income or wealth inequality (Gini, 1912, 1914), the Gini coefficient is nowadays one of the most common statistical indices employed for measuring concentration in the distribution of a positive random variable. In recent years, the coefficient has been used to measure lifespan inequality within and between populations (see, e.g., Shkolnikov et al., 2003; Smits and Monden, 2009; van Raalte and Caswell, 2013; Gigliarano et al., 2017) and to evaluate mortality forecasts (Diaz et al., 2018; Basellini and Camarda, 2019b). The coefficient takes values between 0 and 1, which correspond to the limit cases of perfect equality and perfect inequality, respectively. For an age-at-death distribution, Gini is equal to zero if all individuals die at the same age, and equal to 1 if all people die at age 0 and one individual dies at a positive age (Shkolnikov et al., 2003). In this article, we compute G_{40} in Equation (11) with the approximation formula proposed by Shkolnikov et al. (2003), and we multiply G_{40} by 100 in order to have a comparable magnitude with e_{40}.

Specifically, we pretend that the last year of collected data is $2015-h$, where $h = 10, 15, 20, 25, 30$ and 35 years. We then fit the C-STAD model to the fully observed cohorts $\boldsymbol{c}_1' = [1835, \ldots, 1905 - h]$, and we forecast mortality h years ahead. If the modal age is always observed for the cohorts to be completed, the exercise is restricted to the group $\boldsymbol{c}_2$. Otherwise $\boldsymbol{c}_3$ will be considered too.

An explicative example of the validation procedure is useful to clarify the out-of-sample exercises. Let us consider $h = 10$: then, the last year of fully observed data is 2005. We fit the C-STAD to the fully observed cohorts $\boldsymbol{c}_1' = [1835, \ldots, 1895]$, and we forecast 10-year ahead. By doing so, we complete the mortality experience of the partially observed cohorts $1896, \ldots, 1905$, and for each of these, we compute and compare the estimated e_{40} and G_{40} with the observed ones.

It is worth mentioning at this point that, for the lower values of h, forecasting is achieved simply by fitting the C-STAD on the partially observed cohorts $\boldsymbol{c}_2$. In the explicative example above, where the last data collection occurred in 2005, the cohort 1896, for instance, has been observed at all ages except 110. We thus take advantage of the nature of cohort data and consider all possible observations to complete the mortality experience of this partial cohort. Conversely, for higher values of h, forecasting is achieved by considering also the cohorts $\boldsymbol{c}_3$, which require the combination of forecasting and estimation of the C-STAD parameters.

In addition to the C-STAD, we perform the same out-of-sample exercises with the 2D P-splines approach of Currie et al. (2004). This is indeed the only model that, up to our knowledge, has been employed to forecast cohort mortality from a cohort perspective (Continuous Mortality Investigation, 2007) and can be implemented in the `R` software after data manipulation (in the `MortalitySmooth` package, Camarda, 2012).

Furthermore, we compare our results with the standard procedure of: (i) forecasting mortality in an age-period fashion, and (ii) extracting cohort patterns from the diagonals of the age-period mortality surface. For these comparisons, we employ the benchmark model of Lee and Carter (LC, 1992). For consistency, in each exercise we select the shortest fitting period that includes all the cohorts needed to assess the LC forecasts. Fitting periods for the six exercises are 1936–2005, 1931–2000, 1926–1995, 1921–1990, 1916–1985 and 1911–1980: the starting year is computed by adding 40 (the starting age of the analysis) to the first forecast cohort, while the last year is given by $2015 - h$. It should be noted that the fitting age-period surface for these exercises was derived from the original age-cohort data, to exclude potential bias related to differences in the computation of period vs. cohort mortality rates (see Wilmoth et al., 2017, pp. 29–33).

Table 1 presents the results of our analysis. The first and second columns contain the cohorts used for fitting and forecasting the C-STAD and 2D P-splines models, respectively. The third column contains the forecast horizon of the out-of-sample exercise, while the fourth column the measure analysed (e_{40} and G_{40}). Results are shown in the last six columns. We assess the accuracy of the point forecasts by computing the root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{h}\sum_{c=1}^{h}(\hat{y}_c - y_c)^2},$$

where h is the forecasting horizon, and $\hat{y}_c$ and y_c are the forecast and observed out-of-sample values of either e_{40} or G_{40}.

The table shows that the C-STAD forecasts are accurate in completing the mortality experience of partially observed cohorts. The RMSE values of both e_{40} and G_{40} are low across the six exercises, and they do not increase significantly with the forecasting horizon. Additionally, C-STAD forecasts are more accurate than those of the 2D P-spline model, and both age-cohort models significantly outperform the standard LC age-period based approach. Employing different fitting periods in the age-period exercises, such as starting the analysis from the same year (for example, 1911) in all cases, would result into even greater RMSE of

Fitting cohorts	Forecast cohorts	Horizon	Measure	Sweden			Denmark		
				C-STAD	2D P-spline	LC (period)	C-STAD	2D P-spline	LC (period)
1835–1895	1896–1905	10y	e_{40}	0.08	**0.08**	0.22	**0.08**	0.08	0.41
			G_{40}	**0.09**	0.10	0.32	**0.08**	0.08	0.38
1835–1890	1891–1905	15y	e_{40}	**0.07**	0.09	0.26	**0.07**	0.08	0.41
			G_{40}	**0.08**	0.10	0.40	**0.07**	0.12	0.44
1835–1885	1886–1905	20y	e_{40}	**0.05**	0.08	0.35	**0.06**	0.08	0.44
			G_{40}	**0.09**	0.11	0.53	**0.06**	0.12	0.45
1835–1880	1881–1905	25y	e_{40}	**0.04**	0.08	0.41	**0.03**	0.08	0.47
			G_{40}	**0.10**	0.11	0.57	**0.11**	0.14	0.47
1835–1875	1876–1905	30y	e_{40}	**0.06**	0.09	0.48	**0.03**	0.11	0.52
			G_{40}	**0.10**	0.14	0.63	**0.11**	0.19	0.52
1835–1870	1871–1905	35y	e_{40}	0.14	**0.08**	0.56	**0.05**	0.14	0.55
			G_{40}	**0.04**	0.13	0.59	**0.06**	0.22	0.49

Table 1. Root mean square error (RMSE) of the C-STAD, 2D P-spline and LC (period) forecasts of e_{40} and G_{40} for adult females in Sweden and Denmark in six out-of-sample validation exercises: forecast horizon of 10, 15, 20, 25, 30 and 35 years. Lower values of the RMSE (in bold, assessed using all available decimals) correspond to greater forecast accuracy.
Source: authors' elaborations on data from the Human Mortality Database (2019).

the LC forecasts for the shorter horizon exercises. Very similar results are obtained by employing different prediction accuracy measures, such as the MAPE and MAE (see Appendix B).

4.2 Mortality developments for Swedish and Danish females, cohorts 1835–1970

In this Subsection we show the results of employing the C-STAD model to estimate and forecast adult female cohort mortality in Sweden and Denmark for the cohorts 1835–1970. The estimated and forecast parameters are shown in Appendix B. Figure 5 shows the observed and fitted remaining life expectancies at age 40 (e_{40}) and Gini coefficient at age 40 (G_{40}) in the two population analysed for the fully observed cohorts $\boldsymbol{c}_1$ (1835–$\check{c}$, where $\check{c}$ is 1906 for Sweden and 1905 for Denmark). The two graphs provide evidence on the goodness-of-fit of the C-STAD model, whose estimates are very close to the observed values for both measures in the two populations. Inspection of the deviance residuals (shown in Appendix B) provides additional evidence on the adequacy of the C-STAD model.

Figure 6 shows the observed (cohorts $\boldsymbol{c}_1$) and completed ($\boldsymbol{c}_2$ and $\boldsymbol{c}_3$) e_{40} and G_{40} computed with the C-STAD (with 80% pointwise confidence intervals) and 2D P-spline model for the two population analysed. Despite sharing similar country trends in the fully observed cohorts $\boldsymbol{c}_1$, it is interesting to observe the different mortality developments in the partially observed cohorts $\boldsymbol{c}_2$: while Swedish adult females show continuous improvements in longevity and lifespan equality, Danish ones display a stagnation of e_{40} and an increase in lifespan inequality. The trends of the mortality measures for the partially observed cohorts are similar across the models, with the exception of Danish e_{40}: the increase of the 2D P-spline forecast e_{40} is much faster then for the C-STAD model, resulting in a crossover among the two populations. Moreover, it is interesting to observe that the C-STAD confidence intervals are rather narrow for both countries in $\boldsymbol{c}_2$ (as the great majority of data is observed for these cohorts), while they increase in the cohorts $\boldsymbol{c}_3$ proportionally to the amount of missing data. Note that $\tilde{c}$ is 1925 and 1927 for Sweden and Denmark, respectively.

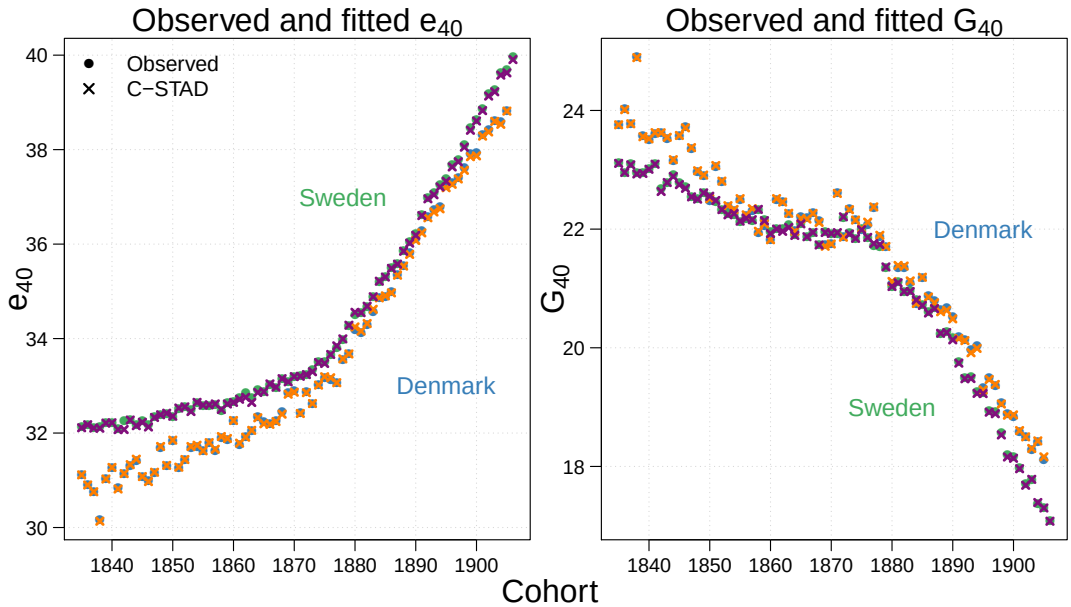

Figure 5. Observed and C-STAD estimated remaining life expectancies at age 40 (e_{40}, left panel) and Gini coefficient at age 40 (G_{40}, right panel) for adult females in Sweden and Denmark for the fully observed cohorts 1835–$\check{c}$ (where $\check{c}$ is 1906 for Sweden and 1905 for Denmark).
Source: authors' elaborations on data from the Human Mortality Database (2019).

The age-specific mortality rates analysis shown in Figure 7 offers additional insights on cohort mortality developments of the two populations. In the top panels, observed, fitted and forecast mortality rates over cohorts are shown for some selected ages. In addition to the goodness-of-fit of the C-STAD model, the graphs highlight diverse age-specific developments in the two countries: for example, mortality at ages 40 and 60 of Danish cohorts born at the beginning of the twentieth century did not improve, resulting in the atypical trends of the summary measures shown in Figure 6 (stagnation of e_{40} and increase of G_{40}). In the bottom panels, mortality rates over all ages are shown for some selected cohorts. This second perspective shows how the shape of the mortality curve, appropriately captured by the C-STAD model, changed over time: for example, mortality at young adult ages was still relatively high in both countries for the 1835 cohort, with the curve being rather flat in the age range 40–50. The subsequent mortality decline at all ages, mainly attributable to improvements in sanitary environment, public hygiene and nutrition (McKeown, 1976), clearly emerges from Figure 7. An additional interesting observation is that the confidence intervals of the C-STAD widens as expected: for example, variability increases with age for the completed cohorts, as fewer age-specific data have been observed at higher ages. Finally, Figure 8 shows the observed and C-STAD age-at-death distributions for the three cohorts analysed in the previous panels.

5 Discussion

Mortality forecasting has drawn considerable interest in recent decades among academics and financial sector practitioners due to the increasing challenges posed by population ageing. Advances in the field have almost exclusively been made on period mortality, as the most recent and innovative techniques are based on modelling and forecasting different functions of period life tables (see, for example, Lee and Carter, 1992; Cairns et al., 2006; Raftery

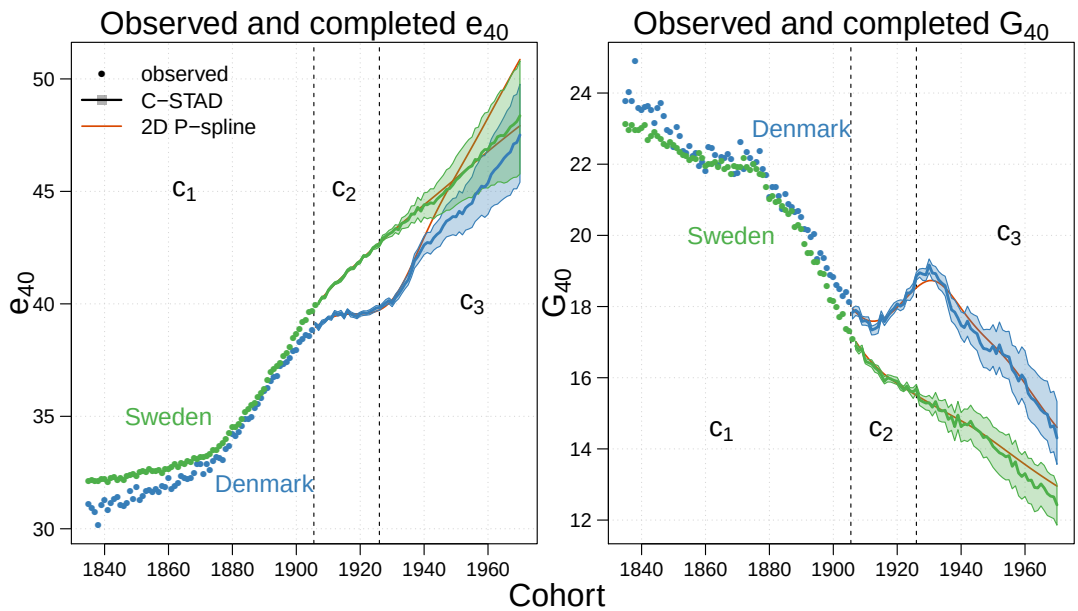

Figure 6. Observed (cohorts $\boldsymbol{c}_1$) and completed ($\boldsymbol{c}_2$ and $\boldsymbol{c}_3$) remaining life expectancies at age 40 (e_{40}, left panel) and Gini coefficient at age 40 (G_{40}, right panel) for the C-STAD (with 80% confidence intervals) and 2D P-spline models for adult females in Sweden and Denmark for the cohorts 1835–1970.
Source: authors' elaborations on data from the Human Mortality Database (2019).

et al., 2013). When considered, cohort effects in mortality modelling and forecasting are typically analysed within an age-period perspective (Renshaw and Haberman, 2006; Cairns et al., 2009; Plat, 2009; Dokumentov et al., 2018).

In this article, we take an alternative perspective and introduce a new methodology to model and forecast mortality from cohort data. An important advantage of cohort forecasts is that they allow to complete the mortality experience of non-extinct cohorts, thus enabling the derivation of their mortality developments. Our approach focuses on cohort age-at-death distributions: specifically, we propose a warping transformation of the age-axis of a standard distribution to describe and forecast adult mortality developments across cohorts. Since we focus on the cohort perspective, we denote our methodology *Cohort Segmented Transformation Age-at-death Distributions* (C-STAD) model. Warping transformations and skewing procedures have already been fruitfully employed to model distributional changes (see, e.g., Fernández and Steel, 1998; Camarda et al., 2008).

Our methodology is inspired by the Segmented Transformation Age-at-death Distributions (STAD) model recently proposed by Basellini and Camarda (2019b) to forecast adult age-at-death distributions. In addition to shifting the focus from period to cohort mortality, our methodology extends the STAD to a cubic transformation before the modal age at death. The additional parameters c_L and d_L are indeed necessary to adequately describe cohort mortality developments at young adult ages. A possible explanation for this is related to the significant improvements in mortality, especially at younger adult ages, across the cohorts that we analyse (cf. Fig. 7). Non-linear transformation functions above the mode were tested too, but they did not provide a better fit compared to a linear transformation function.

Only a handful of models have been proposed to directly forecast cohort mortality so far (Chiou and Müller, 2009; Zanotto and Mazzuco, 2017; Rizzi et al., 2019). One of the main reasons for the limited efforts in this direction is the heavy data demands that such models

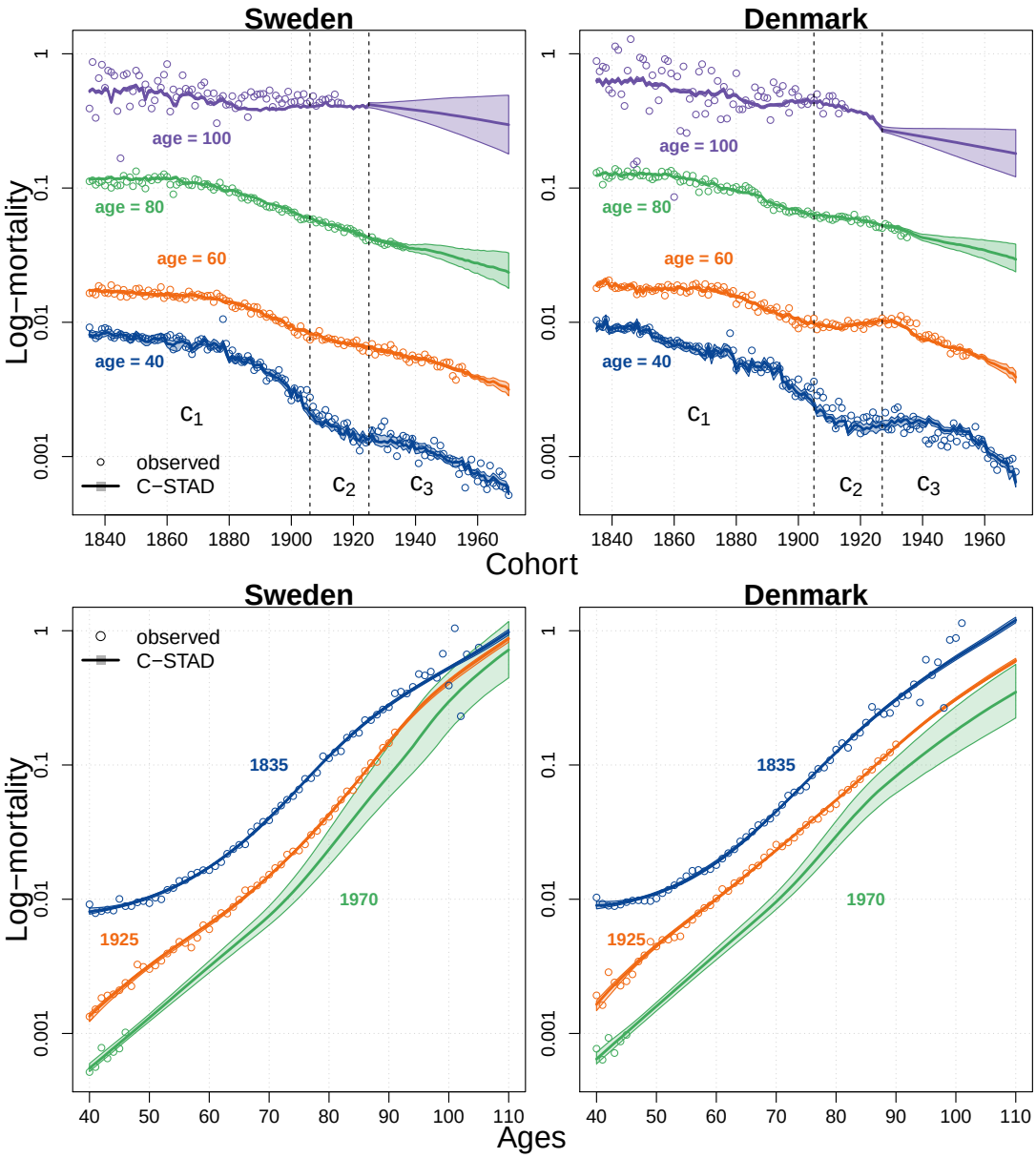

Figure 7. Observed, fitted and forecast age-specific mortality rates for selected ages (top panels) and for selected cohorts (bottom panels) with 80% confidence intervals for females in Sweden and Denmark aged 40–110+ for the cohorts 1835–1970.
Source: authors' elaborations on data from the Human Mortality Database (2019).

require. However, this problem is reduced when only adult mortality is considered (Booth, 2006). As such, the issue does not affect us to a great extent, as our interest in this article is restricted to adult mortality only.

We have shown the results of fitting and forecasting cohort mortality with the C-STAD model for Swedish and Danish adult females aged 40–110 for the cohorts 1835–1970. Our methodology is accurate from a point forecast perspective: for each population, we performed six out-of-sample validation exercises of different forecast horizons. The resulting point forecast errors are generally small, even for the longer forecast horizons. Additionally, the C-STAD forecasts are consistently more precise than those of the 2D P-spline model of Currie et al. (2004), which has been already used to directly forecast cohort mortality (Continuous

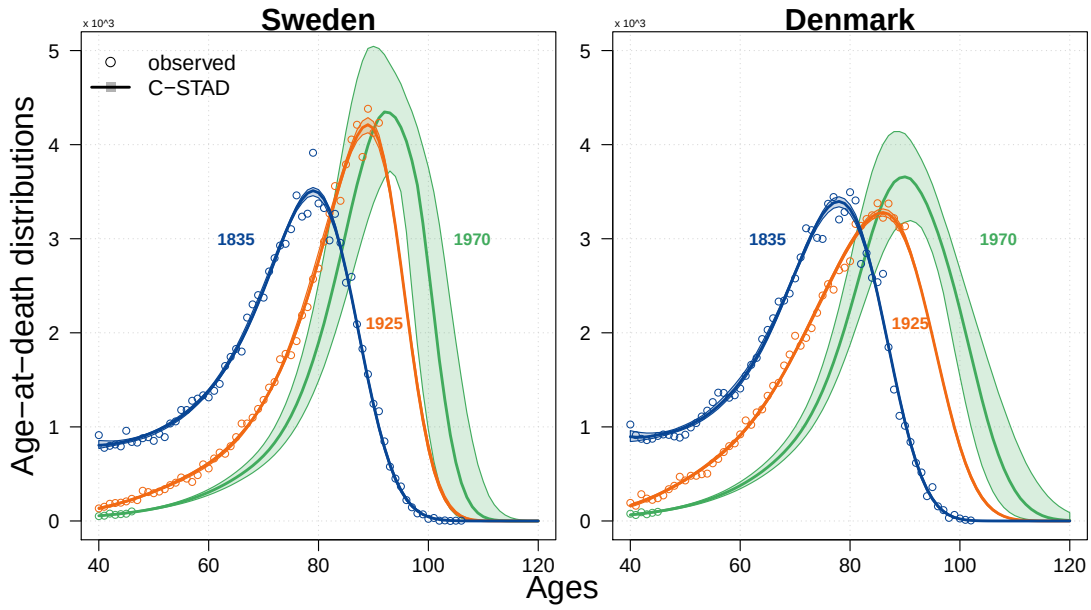

Figure 8. Observed, fitted and forecast age-at-death distributions for selected cohorts (bottom panels) with 80% confidence intervals for females in Sweden and Denmark aged 40–110+.
Source: authors' elaborations on data from the Human Mortality Database (2019).

Mortality Investigation, 2007).

Furthermore, both age-cohort approaches (the C-STAD and 2D P-spline) to forecast cohort mortality perform significantly better than the standard approach of extracting cohort patterns from the diagonals of a projected Lee and Carter (LC, 1992) age-period surface. More generally, age-cohort models are more efficient and parsimonious: looking at Figure 3 the objective of a cohort forecast is to project the green area, which is directly obtained by age-cohort approaches. Conversely, age-period models would also need to forecast all data below such green area as preliminary step, and then extract the cohort mortality of interest.

Our results allow us to derive age-specific and summary measures of mortality, such as remaining life expectancy and Gini coefficient at age 40 (e_{40} and G_{40}), for all cohorts of the population analysed. Although following similar trends to the 2D P-spline model, C-STAD estimates of e_{40} seem more coherent when considered together, lacking the rapid increase and cross-over of Danish forecast displayed by the 2D model. With respect to Danish forecasts, it is interesting to observe a stagnation of e_{40} and an increase of G_{40} for the cohorts 1910–1930. Such results are consistent with other findings in the literature (see, e.g., Fig. 4 in Jacobsen et al., 2002), which have been attributed to the smoking behaviour of Danish women (Jacobsen et al., 2006; Lindahl-Jacobsen et al., 2016).

Since Thiele (1871), actuaries and demographers have decomposed the age-pattern of mortality into three independent components that mainly operate at childhood, middle and old ages, respectively. Let us denote the three components Childhood, Early-adulthood and Senescence. Our proposed methodology is specifically designed to model and forecast the Senescent component of mortality. As such, applying the C-STAD from age 40 produces satisfactory results for the female populations analysed in this article because the Childhood and Early-adulthood components are negligible with respect to the Senescent one in the age range that we study.

Although theoretically and practically possible, we do not recommend to apply the model to

a (much) wider age range. This would result in a reduction of goodness-of-fit and forecast accuracy, because the C-STAD cannot capture and disentangle the combination of different components at younger ages. A more suitable approach for modelling and forecasting the entire age range would be to decompose and model the mortality pattern by specialized versions of the C-STAD on the component-specific distributions. See Basellini and Camarda (2019a) for an example of this procedure from the age-period perspective.

To conclude, the C-STAD model offers great prospects for mortality forecasting from the cohort perspective. Public and private institutions would benefit from employing our model, as it provides a direct approach to complete the mortality experience of non-extinct cohorts. The `R` code provided with this article allows a fast and freely available opportunity for this purpose.

References

Andreev, K. F. (2002). *Evolution of the Danish population from 1835 to 2000*. Odense Monographs on Population Aging. Odense University Press, Odense - DK.

Barbieri, M., Wilmoth, J. R., Shkolnikov, V. M., Glei, D., Jasilionis, D., Jdanov, D., Boe, C., Riffe, T., Grigoriev, P., and Winant, C. (2015). Data Resource Profile: The Human Mortality Database (HMD). *International Journal of Epidemiology*, 44(5):1549–1556.

Basellini, U. and Camarda, C. G. (2019a). A Three-component Approach to Model and Forecast Age-at-death Distributions. In Mazzuco, S. and Keilman, N., editors, *New Perspectives on Population Forecasting*. Springer. Forthcoming.

Basellini, U. and Camarda, C. G. (2019b). Modelling and forecasting adult age-at-death distributions. *Population Studies*, 73(1):119–138.

Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J., and Vaupel, J. W. (2017). Coherent forecasts of mortality with compositional data analysis. *Demographic Research*, 37(17):527–566.

Bohk-Ewald, C., Ebeling, M., and Rau, R. (2017). Lifespan Disparity as an Additional Indicator for Evaluating Mortality Forecasts. *Demography*, 54(4):1559–1577.

Bongaarts, J. (2005). Long-range trends in adult mortality: Models and projection methods. *Demography*, 42(1):23–49.

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

Borgan, Ø. and Keilman, N. (2019). Do Japanese and Italian Women Live Longer than Women in Scandinavia? *European Journal of Population*, 35(1):87–99.

Brass, W. (1971). On the Scale of Mortality. In Brass, W., editor, *Biological Aspect of Demography*, pages 69–110. Taylor & Francis, London.

Brillinger, D. R. (1986). A biometrics invited paper with discussion: The natural variability of vital rates and associated statistics. *Biometrics*, 42(4):693–734.

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

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

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

Camarda, C. G. (2012). MortalitySmooth: An `R` Package for Smoothing Poisson Counts with P-Splines. *Journal of Statistical Software*, 50:1–24. Available on `www.jstatsoft.org/v50/i01`.

Camarda, C. G., Eilers, P. H., and Gampe, J. (2008). A Warped Failure Time Model for Human Mortality. In *Proceedings of the 23rd International Workshop of Statistical Modelling*, pages 149–154.

Canudas-Romo, V. (2010). Three measures of longevity: Time trends and record values. *Demography*, 47(2):299–312.

Chiou, J.-M. and Müller, H.-G. (2009). Modeling hazard rates as functional data for the analysis of cohort lifetables and mortality forecasting. *Journal of the American Statistical Association*, 104(486):572–585.

Christensen, K., Davidsen, M., Juel, K., Mortensen, L., Rau, R., and Vaupel, J. W. (2010). *International Differences in Mortality at Older Ages*, chapter The Divergent Life-Expectancy Trends in Denmark and Sweden - and Some Potential Explanations. National Academies Press, Washington (DC).

Continuous Mortality Investigation (2007). Stochastic projection methodologies: Further progress and P-Spline model features, example results and implications. Revised Working Paper 20 (November 2007).

Currie, I. D., Durban, M., and Eilers, P. H. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, 4(4):279–298.

Diaz, G., Debón, A., and Giner-Bosch, V. (2018). Mortality forecasting in Colombia from abridged life tables by sex. *Genus*, 74(1):15.

Dokumentov, A., Hyndman, R. J., and Tickle, L. (2018). Bivariate smoothing of mortality surfaces with cohort and period ridges. *Stat*, 7(1):e199.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Eilers, P. H. (2007). Ill-posed problems with counts, the composite link model and penalized likelihood. *Statistical Modelling*, 7(3):239–254.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties (with discussion). *Statistical Science*, 11(2):89–102.

Fernández, C. and Steel, M. F. J. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine*, 303(3):130–135.

Gigliarano, C., Basellini, U., and Bonetti, M. (2017). Longevity and concentration in survival times: the log-scale-location family of failure time models. *Lifetime Data Analysis*, 23(2):254–274.

Gini, C. (1912). Variabilità e mutabilità. Contributi allo studio dele relazioni e delle distribuzioni statistiche. *Studi Economico-Giuridici della Università di Cagliari.*

Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73:1203–1248.

Goldstein, J. R. and Wachter, K. W. (2006). Relationships between period and cohort life expectancy: Gaps and lags. *Population Studies*, 60(3):257–269.

Hanada, K. (1983). A formula of Gini's concentration ratio and its application to life tables. *Journal of the Japan Statistical Society, Japanese Issue*, 13(2):95–98.

Human Mortality Database (2019). University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 8 May 2019).

Jacobsen, R., Keiding, N., and Lynge, E. (2002). Long term mortality trends behind low life expectancy of Danish women. *Journal of Epidemiology & Community Health*, 56(3):205–208.

Jacobsen, R., Keiding, N., and Lynge, E. (2006). Causes of death behind low life expectancy of Danish women. *Scandinavian Journal of Public Health*, 34(4):432–436.

Janssen, F. and de Beer, J. (2019). The timing of the transition from mortality compression to mortality delay in Europe, Japan and the United States. *Genus*, 75(1).

Kannisto, V. (2001). Mode and dispersion of the length of life. *Population: An English Selection*, 13:159–171.

Keilman, N. and Pham, D. (2006). Prediction intervals for Lee-Carter-based mortality forecasts. In *European Population Conference 2006, Liverpool, June 21–24, 2006.*

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer Science & Business Media, second edition.

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

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

Lindahl-Jacobsen, R., Oeppen, J., Rizzi, S., Möller, S., Zarulli, V., Christensen, K., and Vaupel, J. W. (2016). Why did Danish women's life expectancy stagnate? The influence of interwar generations' smoking behaviour. *European Journal of Epidemiology*, 31(12):1207–1211.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models.* London: Chapman and Hall.

McKeown, T. (1976). *The modern rise of population.* London: Edward Arnold.

Oeppen, J. (2008). Coherent forecasting of multiple-decrement life tables: a test using Japanese cause of death data. In *Compositional Data Analysis Conference.*

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

Ouellette, N. and Bourbeau, R. (2011). Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach. *Demographic Research*, 25:595–628.

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.

Pascariu, M. D., Lenart, A., and Canudas-Romo, V. (2019). The maximum entropy mortality model: forecasting mortality using statistical moments. *Scandinavian Actuarial Journal*, pages 1–25.

Pfaff, B. (2008a). *Analysis of integrated and cointegrated time series with R*. Springer Science & Business Media.

Pfaff, B. (2008b). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, 27(4):1–32.

Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics*, 45(3):393–404.

Preston, S. H., Heuveline, P., and Guillot, M. (2001). *Demography. Measuring and Modeling Population Processes*. Blackwell.

R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A. E., Chunn, J. L., Gerland, P., and Ševčíková, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography*, 50(3):777–801.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer-Verlag, 2nd edition.

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

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

Riley, J. C. (2001). *Rising life expectancy: a global history*. Cambridge University Press.

Rizzi, S., Kjærgaard, S., Bergeron-Boucher, M.-P., Lindahl-Jacobsen, R., and Vaupel, J. (2019). Forecasting mortality of not extinct cohorts. In *21st Nordic Demographic Symposium, Reykjavik (Iceland)*.

Shang, H. L., Booth, H., and Hyndman, R. (2011). Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods. *Demographic Research*, 25(5):173–214.

Shkolnikov, V. M., Andreev, E. E., and Begun, A. Z. (2003). Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples. *Demographic Research*, 8(11):305–358.

Shkolnikov, V. M., Jdanov, D. A., Andreev, E. M., and Vaupel, J. W. (2011). Steep Increase in Best-Practice Cohort Life Expectancy. *Population and Development Review*, 37(3):419–434.

Smits, J. and Monden, C. (2009). Length of life inequality around the globe. *Social Science & Medicine*, 68(6):1114–1123.

Thiele, T. N. (1871). On a mathematical formula to express the rate of mortality throughout the whole of life, tested by a series of observations made use of by the Danish Life Insurance Company of 1871. *Journal of the Institute of Actuaries and Assurance Magazine*, 16(5):313–329.

van Raalte, A. A. and Caswell, H. (2013). Perturbation analysis of indices of lifespan variability. *Demography*, 50(5):1615–1640.

Vaupel, J. and Lundström, H. (1994). Longer life expectancy? Evidence from Sweden of reductions in mortality rates at advanced ages. In *Studies in the Economics of Aging*, pages 79–102. University of Chicago Press.

Wilmoth, J., Andreev, K., Jdanov, D., Glei, D., Riffe, T., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., Vachon, P., Winat, C., and Barbieri, M. (2017). Methods Protocol for the Human Mortality Database. Last Revised: November 27, 2017 (Version 6).

Wilmoth, J. R. and Lundström, H. (1996). Extreme longevity in five countries. *European Journal of Population/Revue Européenne de Démographie*, 12(1):63–93.

Zanotto, L. and Mazzuco, S. (2017). Reconstruction of cohort data via EM algorithm. In *XXVIII International Population Conference*.

A Deviance residuals and bootstrap death counts

Model residuals are routinely analysed to explore the goodness-of-fit of a model as well as the adequacy of assumptions about error terms. Within a GLM setting (such as the Poisson considered here), deviance residuals are often used to measure discrepancy between fitted and actual data. For the Poisson distribution they are given by:

$$r_{\mathrm{D}} = \mathrm{sign}(d_{x,c} - \hat{d}_{x,c})\sqrt{2}\left[d_{x,c}\ln\left(\frac{d_{x,c}}{\hat{d}_{x,c}}\right) - \left(d_{x,c} - \hat{d}_{x,c}\right)\right]^{1/2} \tag{12}$$

where $d_{x,c}$ and $\hat{d}_{x,c}$ denote the observed and fitted death counts at age x and for cohort c, respectively (McCullagh and Nelder, 1989).

Deviance residuals can be further employed to take into account the uncertainty related to the estimation of a model parameters as suggested by Koissi et al. (2006). Specifically, bootstrap death counts can be computed by resampling deviance residuals with replacement and mapping them to corresponding death counts. We refer the interested reader to Renshaw and Haberman (2008) for details on the inverse formulas, which are based on the seminal work of Efron and Tibshirani (1994).

B Section 4: additional results

In this appendix, we present additional results related to Section 4.

First, Figure B.1 provides an illustration of the landmark registration procedure that we employ to compute the standard from the aligned distributions. The left panel shows the observed smooth distributions derived from the 2D P-spline model; in the right panel, the same distributions have been aligned to a common modal age at death, corresponding to

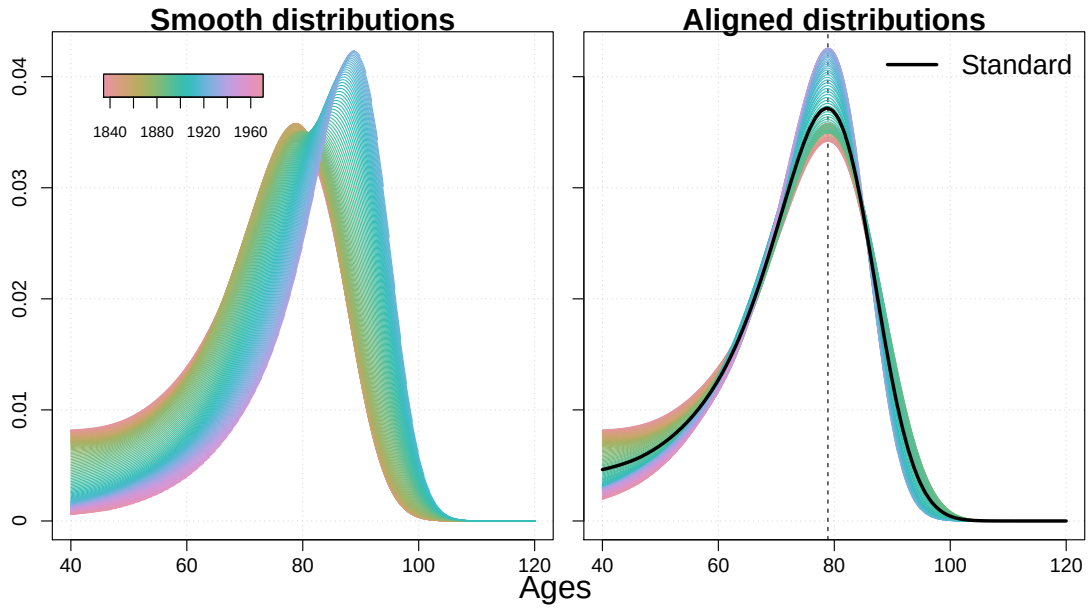

Figure B.1. Observed smooth (left panel) and aligned distributions for Swedish females aged 40–110+ for cohorts 1835–1970. The standard distribution (black line) is computed as mean of the aligned distributions.
Source: authors' elaborations on data from the Human Mortality Database (2019).

the mode of the first cohort (1835). The standard is computed as mean of the aligned distributions.

Next, we report the out-of-sample results of Subsection 4.1 derived from employing two different prediction accuracy measures. In addition to the root mean square error, we computed the mean absolute error (MAE) and the mean absolute percentage error (MAPE):

$$\text{MAE} = \frac{1}{h}\sum_{c=1}^{h} |\hat{y}_c - y_c| \ ,$$

$$\text{MAPE} = \frac{1}{h}\sum_{c=1}^{h} \left|\frac{\hat{y}_c - y_c}{y_c}\right| \ ,$$

where h is the forecasting horizon, and $\hat{y}_c$ and y_c are the forecast and observed out-of-sample values of either e_{40} or G_{40}.

Tables B.1 and B.2 show the out-of-sample results obtained using the MAE and the MAPE, respectively. The results are very similar to those obtained with the RMSE shown in Table 1: the C-STAD forecasts are accurate in completing the mortality experience of partially observed cohorts, with forecasts errors generally low and smaller than the 2D P-spline model. Moreover, both age-cohort models significantly outperform the standard LC age-period based approach.

Finally, we provide additional results with regard to Subsection 4.2. Figure B.2 shows the fitted and forecast C-STAD parameters with 80% confidence intervals for Swedish and Danish adult females for cohorts 1835–1970.

We performed diagnostic checks on the fitted C-STAD model for the two populations analysed in this paper by using Equation (12). Poisson deviance residuals for the two populations are shown in Figure B.3. No clear patterns emerge from this graphical analysis, with the exception of the years corresponding to the Spanish flu and the World War II.

Fitting cohorts	Forecast cohorts	Horizon	Measure	Sweden			Denmark		
				C-STAD	2D P-spline	LC (period)	C-STAD	2D P-spline	LC (period)
1835–1895	1896–1905	10y	e_{40}	0.08	**0.06**	0.20	0.07	**0.07**	0.37
			G_{40}	**0.07**	0.08	0.28	**0.06**	0.07	0.34
1835–1890	1891–1905	15y	e_{40}	**0.07**	0.07	0.22	**0.06**	0.07	0.35
			G_{40}	**0.07**	0.08	0.35	**0.06**	0.09	0.38
1835–1885	1886–1905	20y	e_{40}	**0.05**	0.07	0.29	**0.06**	0.07	0.38
			G_{40}	**0.06**	0.09	0.45	**0.05**	0.09	0.39
1835–1880	1881–1905	25y	e_{40}	**0.03**	0.06	0.37	**0.03**	0.07	0.42
			G_{40}	**0.07**	0.08	0.52	**0.08**	0.10	0.39
1835–1875	1876–1905	30y	e_{40}	**0.04**	0.07	0.41	**0.02**	0.08	0.46
			G_{40}	**0.08**	0.12	0.55	**0.07**	0.13	0.44
1835–1870	1871–1905	35y	e_{40}	0.09	**0.07**	0.48	**0.04**	0.10	0.49
			G_{40}	**0.03**	0.10	0.52	**0.05**	0.17	0.43

Table B.1. Mean absolute error (MAE) of the C-STAD, 2D P-spline and LC (period) forecasts of e_{40} and G_{40} for adult females in Sweden and Denmark in six out-of-sample validation exercises: forecast horizon of 10, 15, 20, 25, 30 and 35 years. Lower values of the MAE (in bold, assessed using all available decimals) correspond to greater forecast accuracy.
Source: authors' elaborations on data from the Human Mortality Database (2019).

Fitting cohorts	Forecast cohorts	Horizon	Measure	Sweden			Denmark		
				C-STAD	2D P-spline	LC (period)	C-STAD	2D P-spline	LC (period)
1835–1895	1896–1905	10y	e_{40}	0.20%	**0.16%**	0.51%	0.19%	**0.19%**	0.96%
			G_{40}	**0.42%**	0.45%	1.57%	**0.35%**	0.36%	1.83%
1835–1890	1891–1905	15y	e_{40}	**0.17%**	0.19%	0.58%	**0.16%**	0.19%	0.93%
			G_{40}	**0.36%**	0.45%	1.93%	**0.30%**	0.47%	2.04%
1835–1885	1886–1905	20y	e_{40}	**0.13%**	0.19%	0.77%	**0.15%**	0.18%	1.02%
			G_{40}	**0.34%**	0.46%	2.40%	**0.27%**	0.44%	2.02%
1835–1880	1881–1905	25y	e_{40}	**0.08%**	0.17%	0.98%	**0.08%**	0.18%	1.14%
			G_{40}	**0.40%**	0.44%	2.71%	**0.42%**	0.51%	2.05%
1835–1875	1876–1905	30y	e_{40}	**0.10%**	0.20%	1.12%	**0.07%**	0.23%	1.28%
			G_{40}	**0.42%**	0.59%	2.84%	**0.39%**	0.63%	2.21%
1835–1870	1871–1905	35y	e_{40}	0.22%	**0.19%**	1.33%	**0.12%**	0.29%	1.36%
			G_{40}	**0.14%**	0.51%	2.62%	**0.23%**	0.81%	2.09%

Table B.2. Mean absolute percentage error (MAPE) of the C-STAD, 2D P-spline and LC (period) forecasts of e_{40} and G_{40} for adult females in Sweden and Denmark in six out-of-sample validation exercises: forecast horizon of 10, 15, 20, 25, 30 and 35 years. Lower values of the MAPE (in bold, assessed using all available decimals) correspond to greater forecast accuracy.
Source: authors' elaborations on data from the Human Mortality Database (2019).

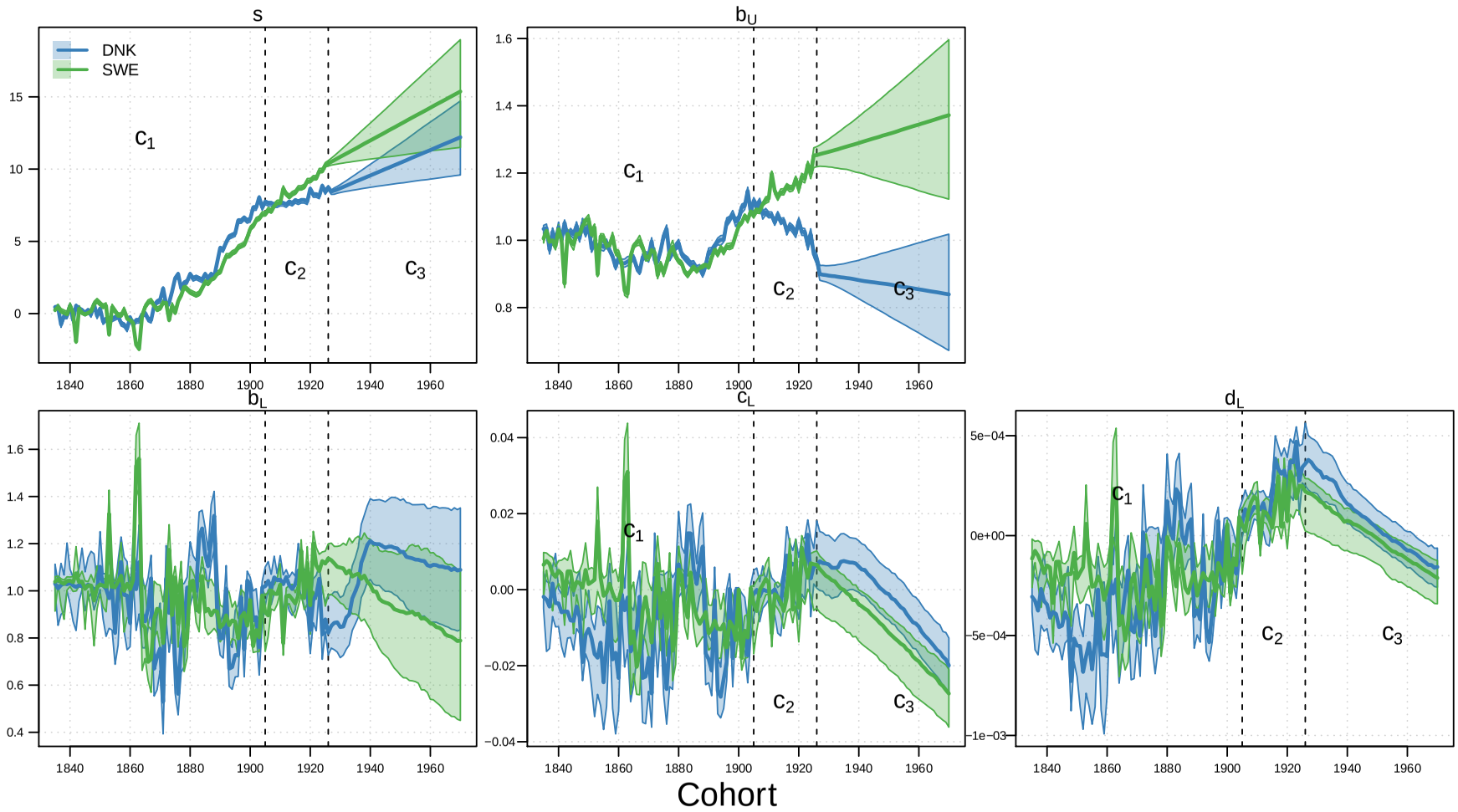

Figure B.2. Estimated and forecast C-STAD parameters for adult females in Sweden and Denmark for the cohorts 1835–1970.
Source: authors' elaborations on data from the Human Mortality Database (2019).

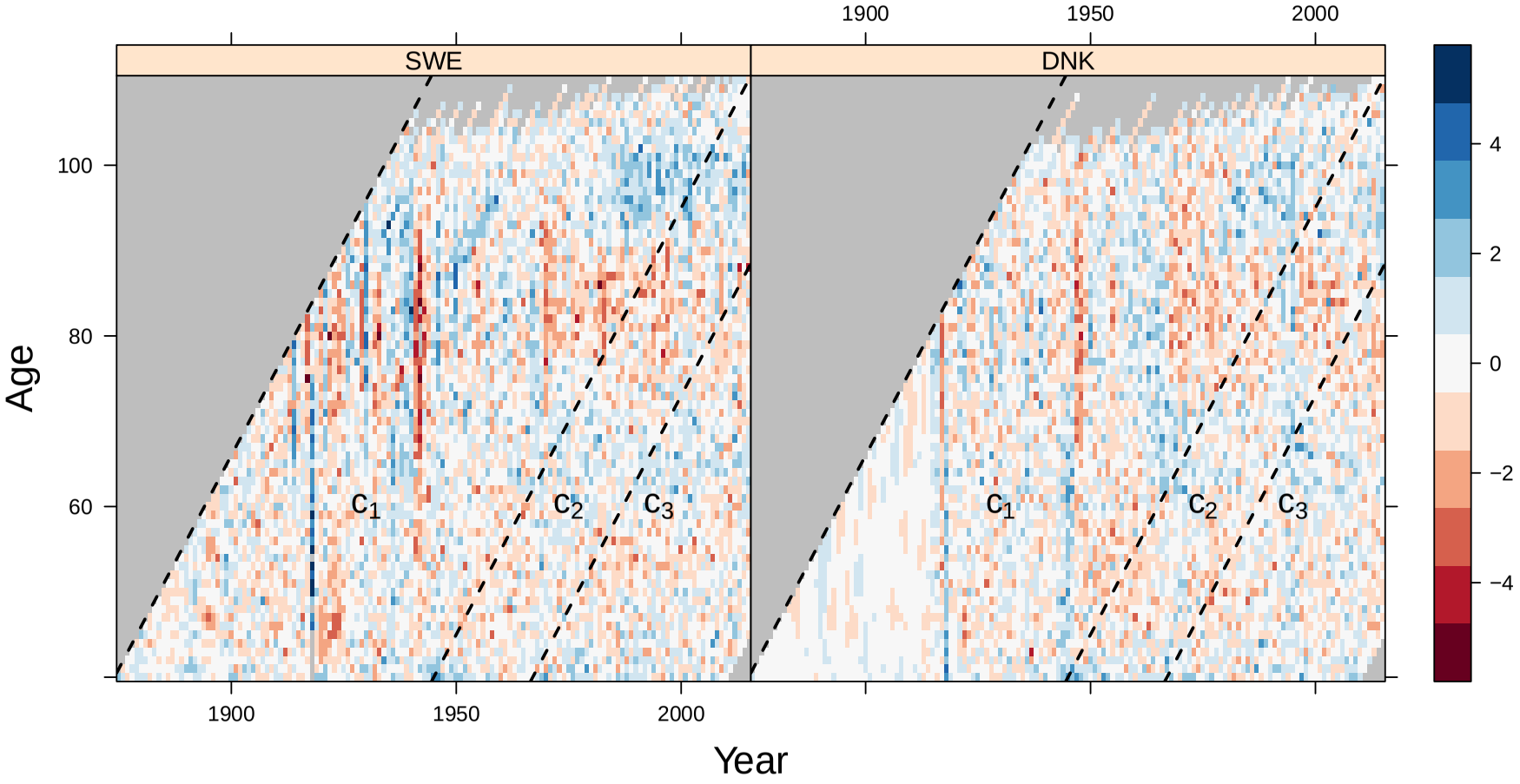

Figure B.3. Poisson deviance residuals of the C-STAD model for adult females in Sweden and Denmark for the cohorts 1835–1970. *Source*: authors' elaborations on data from the Human Mortality Database (2019).