Insurance: Mathematics and Economics 93 (2020) 95–115

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Modelling life tables with advanced ages: An extreme value theory approach

Fei Huang *, Ross Maller, Xu Ning

Research School of Finance, Actuarial Studies and Statistics, College of Business and Economics, Australian National University, Australia

ARTICLE INFO

Article history:
Received September 2019
Received in revised form February 2020
Accepted 14 April 2020
Available online 27 April 2020

JEL classification:
G22

Keywords:
Advanced age mortality
Extreme value theory
Generalised Pareto distribution
Gompertz model
Life table

ABSTRACT

We propose a new model – we call it a smoothed threshold life table (STLT) model – to generate life tables incorporating information on advanced ages. Our method allows a smooth mortality transition from non-extreme to extreme ages, and provides objectively determined highest attained ages with which to close the life table.

We proceed by modifying the threshold life table (TLT) model developed by Li et al. (2008). In the TLT model, extreme value theory (EVT) is used to make optimal use of the relatively small number of observations at high ages, while the traditional Gompertz distribution is assumed for earlier ages. Our novel contribution is to constrain the hazard function of the two-part lifetime distribution to be continuous at the changeover point between the Gompertz and EVT models. This simple but far-reaching modification not only guarantees a smooth transition from non-extreme to extreme ages, but also provides a better and more robust fit than the TLT model when applied to a high quality Netherlands dataset. We show that the STLT model also compares favourably with other existing methods, including the Gompertz–Makeham model, logistic models, Heligman–Pollard model and Coale–Kisker method, and that a further generalisation, a time-dependent dynamic smooth threshold life table (DSTLT) model, generally has superior in-sample fitting as well as better out-of-sample forecasting performance, compared, for example, with the Cairns et al. (2006) model.

1. Introduction

The number of supercentenarians in developed countries has increased considerably in the past few decades (Vaupel and Robine, 2002). For life annuities and defined benefit pension plans, the rise of supercentenarians can lead to low-frequency, high-severity losses (Li et al., 2008). Coupled with an increased awareness of longevity risk (Cairns et al., 2008) as well as the difficulty of reinsuring life annuities (Purcal, 2006), there has been a subsequent increased demand for actuaries to more accurately model extreme age survival probabilities.

However, the relatively less exposure to risk and consequent small numbers of deaths at high ages results in large sampling errors and highly volatile crude observed probabilities of death, which makes modelling probabilities of death at these ages difficult (Wilmoth, 1995). Furthermore, ages at death are often misreported, and substantial effort is needed to ascertain them accurately, see Bernard and Vaupel (1999) and Bourbeau and Desjardins (2002). In light of this, actuaries sometimes close life tables at an arbitrarily chosen age (e.g. Coale and Kisker, 1990). Early closure however results in less accurate estimates of expected shortfall and value-at-risk for life annuity and defined benefit pension plan products, because these measures are by definition determined by the tail of the lifetime distribution (Duffie and Pan, 1997).

The threshold life table (TLT) method of Li et al. (2008) is a recent development that addresses these problems. In this model, from a sample or population of ages at death data, extreme value theory is used to make optimal use of the relatively small number of observations at high ages, while a traditional Gompertz law (Gompertz, 1825) is assumed for earlier ages. The threshold life table has two main attractive features. Firstly, it allows the end point of a life table to be determined in an objective, data-driven way. Secondly, the threshold age – the age at which the modelled lifetime distribution transitions from Gompertz to the distribution suggested by extreme value theory – is also selected objectively. However, it has a couple of drawbacks too. The model is such that the hazard function of the two-part lifetime distribution is not in general continuous at the threshold age, so the mortality transition from non-extreme to extreme ages is not smooth. Furthermore, the model structure is static and does not allow for forecasting of mortality rates into the future.

* Correspondence to: Research School of Finance, Actuarial Studies and Statistics, College of Business and Economics, Australian National University, Canberra, ACT 2601, Australia.
E-mail address: cathy.fei.huang@gmail.com (F. Huang).

https://doi.org/10.1016/j.insmatheco.2020.04.004

In this paper, we propose a smooth threshold life table (STLT) model, as an extension of the TLT model, which overcomes the continuity problem by incorporating a smoothing adjustment at the threshold age. Using a high quality dataset constructed by augmenting Netherlands mortality data from the Human Mortality Database (HMD) with a dataset provided by the Centraal Bureau voor de Statistiek (CBS) of the Netherlands, we demonstrate that the STLT model not only guarantees a smooth mortality transition from non-extreme to extreme ages, but also provides a better and more robust fit to the data than the TLT model. Additionally, we introduce a dynamic adjustment for the STLT model and develop what we call the dynamic smooth threshold life table (DSTLT) model for forecasting. Empirical results using the augmented dataset show that the DSTLT model has both superior in-sample fitting and out-of-sample forecasting performance compared with the Cairns–Blake–Dowd (CBD) model.

The rest of the paper is organised as follows. Section 2 provides a literature review of advanced age mortality models, including those we will use for comparison. Section 3 introduces the data sources used in this paper and describes the construction of the augmented dataset. Section 4 formulates the Li et al. (2008) TLT model as well as the newly proposed STLT model and outlines the estimation methods. Section 5 presents the empirical results and analyses using both the new and previously existing methods. This section begins with an assessment of how much augmenting the HMD with the CBS data improves the fitting of the TLT and STLT models, then goes on to exemplify the methodology on a particular cohort (the 1901 female birth cohort) of the Netherlands dataset. This methodology is subsequently applied to all cohorts (1893–1908 birth cohorts, female and male) with the results reported in detail in Appendices C and D. Section 5 concludes with a comparison of the models summarised in Section 2. Section 6 outlines the construction of the dynamic (DSTLT) model based on results from fitting the STLT to all cohorts in the augmented dataset, and gives a comparison of it with the CBD model, cohort-by-cohort, both in and out-of-sample. Section 7 provides a discussion of the findings, with particular reference to the question of the existence of an upper limit to the length of human life. Some subsidiary results are collected in the Appendices.

2. Advanced age mortality models

There is a large literature on human mortality modelling. Very early on, Gompertz (1825) suggested that mortality increases exponentially with age during the adult years of life. Makeham (1860) extended the Gompertz (1825) model by adding an age-dependent component to better capture younger age mortality. These models are static and one-dimensional, and cannot be used for mortality forecasting. Later, a number of new approaches using stochastic models such as those of Lee and Carter (1992), Renshaw and Haberman (2006) and Currie et al. (2004) were developed for mortality forecasting. A useful literature review of these kinds of mortality models can be seen in Booth and Tickle (2008). However, these models are only suitable for mortality modelling of non-extreme ages, say up to age 90. Due to the sparse, unreliable and volatile nature of advanced age mortality data, there is only a small literature on advanced age mortality modelling and very limited work attempting to model non-extreme and extreme ages together in a coherent framework. We summarise some existing static and dynamic models of advanced age mortality in the following subsections.

The following notation will be used throughout the paper:

- X: age at death of an individual from the population, a continuous random variable
- $f(x)$: probability density function of X
- $F(x)$: cumulative distribution function (cdf) of X
- $S(x) = 1 - F(x)$: survival function of F
- $h(x) = f(x)/(1-F(x))$: force of mortality, or hazard function, corresponding to F
- d_x: the number of deaths between integer ages x and $x+1$
- E_x: the population exposed to the risk of death between integer ages x and $x+1$. E_x is approximated by the midyear population at age x
- l_x: the number of survivors to age x
- $m_x = d_x/E_x$: the central rate of death at age x
- $q_x = d_x/l_x$: the empirical conditional probability of death between ages x and $x+1$ for individuals surviving to age x.

2.1. Static advanced age mortality models

Gompertz–Makeham Law of Mortality

One of the first successful attempts to develop a parametric model of mortality was that of Gompertz (Gompertz, 1825). Gompertz modelled adult mortality with two parameters: a positive scale parameter B that represents the level of mortality, and a positive shape parameter C that measures the rate of increase in mortality with age. The force of mortality in the Gompertz model is

$$h(x) = B\exp(Cx). \tag{1}$$

This *"Gompertz law"* was found to apply fairly well (over appropriate age ranges) in many countries during the last 170 years. Gompertz put forward a possible physiological explanation: that "a man's power to avoid death is gradually exhausted as his age increases". Most modern attempts to explain the law are linked to steady bodily deterioration, perhaps due to the accumulation of molecular and cellular damage, over the age ranges concerned (Thatcher et al., 1998).

It was found by Makeham (1860) that the fit of Gompertz's distribution could often be improved by adding a constant term, so that the hazard is modelled as

$$h(x) = A + B\exp(Cx). \tag{2}$$

This is known as the Makeham, or *Gompertz–Makeham*, law of mortality. The extra parameter A represents mortality resulting from causes, such as accidents or diseases, which are the same for all ages (Saikia and Borah, 2013) and unrelated to either maturation or senescence. If X, the lifetime of a person, has a Gompertz distribution, and Y, the time to a fatal accident, has an exponential distribution, and the random variables X and Y are independent, then the minimum of X and Y has a Makeham distribution. Therefore the Makeham model can be seen as a shock model (Bower et al., 1997).

Logistic Models

A general class of logistic models was formulated by Perks (1932), who found empirically that the values of $h(x)$ in a life table could be fitted by a logistic curve. The hazard function in *Perks's model* is of the form

$$h(x) = A + \frac{B\exp(Cx)}{1 + D\exp(Cx)}. \tag{3}$$

Other old-age mortality models such as the Beard (1971) and Kannisto (1992) models may be written as special cases of the logistic model. Setting $A = 0$ in (3) gives the three-parameter *Beard model*

$$h(x) = \frac{B\exp(Cx)}{1 + D\exp(Cx)}. \tag{4}$$

Kannisto (1992) noticed that modern data for $h(x)$ at high ages can be well-fitted by one of the simplest forms of the logistic

model, in which logit($h(x)$) is a linear function of x. The resulting *Kannisto model* specifies the hazard function as

$$h(x) = \frac{B\exp(Cx)}{1 + B\exp(Cx)}. \tag{5}$$

These logistic-type models have often been shown to have superior fits to the traditional Gompertz or Makeham fits — see for example Rau et al. (2017); Saikia and Borah (2013); Thatcher et al. (1998).[1] Thatcher et al. (1998) mentioned that the logistic model can also be considered as a shock model: if the lifetime X follows a Beard distribution, the time to an accident Y is exponentially distributed and the random variables X and Y are independent, then $\min(X, Y)$ follows a Perks distribution. Beard (1971) showed how (3) can arise as a mixture model of a heterogeneous population. If the members of the population are subject to hazards of the Makeham form (2), but with the parameter B varying from individual to individual according to a gamma distribution, then the average value of $h(x)$ for the survivors who reach age x will have a logistic form (3). The resulting *Gamma–Makeham* model satisfies

$$h(x) = A + \frac{B\exp(Cx)}{1 + BD(\exp(Cx) - 1)/C}. \tag{6}$$

Another, stochastic process, approach to this model was by Le Bras (1976). Yashin et al. (1994) show that Le Bras's formula for $h(x)$ was essentially the same as that given by the Gamma–Makeham model, despite being derived under quite different assumptions. Another version of a logistic model called the *Gamma–Gompertz model* (Vaupel et al., 1979) satisfies

$$h(x) = \frac{B\exp(Cx)}{1 + BD(\exp(Cx) - 1)/C}. \tag{7}$$

For the logistic-type curves (3)–(7), the force of mortality $h(x)$ asymptotes to a constant as $x \to \infty$. There is no such asymptote for the Gompertz and Makeham models, so they cannot capture a late life mortality deceleration effect such as we discuss in Section 7 or other subtle features of the late life mortality curve.

Heligman–Pollard Model

Heligman and Pollard (1980) modelled q_x as follows:

$$\frac{q_x}{1 - q_x} = A^{(x+B)^C} + D\exp(-E[\ln x - \ln F]^2) + GH^x, \quad x = 1, 2, 3, \ldots. \tag{8}$$

Here A, B, C, D, E, F, G, H are constants, and the terms $A^{(x+B)^C}$, $D\exp(-E[\ln x - \ln F]^2)$ and GH^x can be taken to represent early childhood mortality, accidental mortality, and senescent mortality, respectively. GH^x may also be seen as a term capturing the Gompertz law of mortality. Heligman and Pollard (1980) obtained a good fit of their model to Australian mortality rates up to age 85 for three periods, 1946–48, 1960–62, and 1970–72.

Note that $q_x < 1$ for all $x > 0$ in the model (8). This is a drawback for actuaries as life tables need to be closed in order to value insurance products. So in practice, they have to arbitrarily select an age with which to close the life table. There is also strong empirical evidence that advanced-age mortality is not Gompertzian (Olshansky and Carnes, 1997).

Coale–Kisker Method

The Coale–Kisker method (Coale and Kisker, 1990) is another well-known technique for modelling probabilities of death in developed countries (Boleslawski and Tabeau, 2001). The method is recursive, of the form

$$k(x) = k(x - 1) - R, \quad x \geq x_0, \tag{9}$$

where $k(x) = \ln(m_x/m_{x+1})$, R is a constant to be determined, and the extrapolation starts at some integer age x_0. Recall that $m_x = d_x/E_x$ is the central rate of death at age x. Applying the formula up to integer age $x = x_1$, we find

$$R = \frac{(x_1 - x_0)k(x_0) + \ln m_{x_0} - \ln m_{x_1}}{1 + 2 + \cdots + (x_1 - x_0)}. \tag{10}$$

The method requires an assumption of the age x_1 at which the life table is closed as well as the value of the central death rate at this closing age. The age from which to start extrapolating also must be subjectively decided. Coale and Kisker (1990) use $x_0 = 84$, $x_1 = 110$, $m_{x_1} = 1.0$.

Li-Hardy-Tan Threshold Life Table (TLT)

Li et al. (2008) proposed the threshold life table model, which models the age-at-death as a continuous random variable X with cdf F. The model is piecewise: below threshold age N, $F(x)$ is assumed to follow the Gompertz law of mortality; above the threshold age N, $F(x)$ is assumed to follow a Generalised Pareto distribution (GPD). The GPD occurs in the "peaks over threshold" model in extreme value theory (p. 352, Embrechts et al., 2013). The model is completely defined by the Gompertz part,

$$F(x) = 1 - \exp\left(-\frac{B}{\ln C}(C^x - 1)\right), \qquad x \leq N, \tag{11}$$

where $B > 0$ and $C > 1$ are constants, together with the GPD part,

$$F(x) = F_\gamma(x) = \begin{cases} 1 - S(N)(1 + \gamma(x - N)/\theta)^{-1/\gamma}, & \gamma > 0,\ x > N, \\ 1 - S(N)\exp(-(x - N)/\theta), & \gamma = 0,\ x > N, \\ 1 - S(N)(1 - |\gamma|(x - N)/\theta)^{1/|\gamma|}, & \gamma < 0,\ N < x < N + \theta/|\gamma|. \end{cases} \tag{12}$$

Here $S(N) = 1 - F(N)$ and $\theta > 0$ and $\gamma \in \mathbb{R}$ are constants. The cdf $F(x)$ is continuous for $x > 0$, but in general the corresponding density and hazard rate are discontinuous at age $x = N$.

In the Gompertz part of the distribution, (11), B may be interpreted as a "baseline" force of mortality, i.e., the force of mortality of a life at age 0. Parameter C controls for the age effect — the amount by which the force of mortality increases with age.

For the GPD part, (12), θ and γ are scale and shape parameters, and the threshold age N plays the role of a location parameter. Analogous to the Gompertz parameters, θ controls the baseline force of mortality for extreme ages, i.e., the force of mortality of a life at threshold age N.

When $\gamma \geq 0$, $F(x)$ has an infinite right endpoint and indefinitely large lifetime ages are potentially possible. But when $\gamma < 0$, $F(x)$ has a finite right end point at $x = N + \theta/|\gamma|$, and ages greater than this cannot occur. Thus the estimated value, $\hat{\gamma}$ of γ from a dataset is relevant to the question of whether there is a finite upper limit to the lifetime of an individual from the population ($\hat{\gamma} < 0$) or no such upper limit ($\hat{\gamma} \geq 0$).

The threshold life table allows the end point of a life table to be determined in an objective way. But the model structure is static and does not allow for forecasting of mortality rates into the future. We will address those issues by introducing the STLT and DSTLT models.

2.2. Dynamic advanced age mortality models

Combined Models

Models in which the lifetime distribution describes a fixed birth-year or death-year cohort do not lend themselves naturally to prediction. Some existing models forecast future advanced age mortality by combining a static model with an external stochastic

[1] Another model that has been proposed in the literature is the CoDe (de Beer and Janssen, 2016) model; however we do not include it in our analysis.

mortality model for non-extreme ages. The Lee–Carter model (Lee and Carter, 1992) is an important example. In models like these, the stochastic component must first be used to project mortality rates of non-extreme ages into the future. The static advanced age mortality component can then be fitted to the predicted probabilities of death of these birth-year or death-year cohorts. For example, Lee and Carter (1992) combined their model for ages below 85 with the Coale–Kisker method for higher ages. Such a combined model however overlooks the inherent relationship between non-extreme and extreme ages, as well as the temporal dependence of extreme-age mortality rates.

Li et al. in their 2008 paper went on to combine the Lee–Carter model with their threshold life table model in order to estimate probabilities of death over time. In this framework, the mortality of non-extreme ages is modelled twice, first using the Lee–Carter model and then using the Gompertz model. But this redundancy makes the model difficult to interpret, and the temporal dependence of extreme-age mortality rates is ignored in this setup.

Watts–Dupuis–Jones (WDJ) Model

Watts et al. (2006) also took an extreme value approach. They fitted a generalised extreme value (GEV) distribution to Japanese data using the 10 highest ages at death recorded annually between 1980 and 2000. The cdf of the GEV distribution G was parametrised as

$$G(z) = \begin{cases} \exp\left(-\left(1+\xi\left(z-\mu\right)/\sigma\right)^{-1/\xi}\right), & \xi \neq 0, \\ \exp\left(-\exp\left(-\left(z-\mu\right)/\sigma\right)\right), & \xi = 0. \end{cases} \tag{13}$$

To incorporate time-variability, they modelled the parameters μ and σ as functions of standardised calendar year:

$$\mu(i) = \mu_0 + \mu_1 i^* \quad \text{and} \quad \sigma(i) = \exp(\sigma_0 + \sigma_i i^*),$$

where $i^* = (i - 1980)/20$ is the standardised year of death in their data, $1980 \leq i \leq 2000$, $0 \leq i^* \leq 1$. The WDJ model can only be used to forecast the highest attained ages, thus rendering comparison with other models difficult. The dynamic model we propose in Section 6 can be used to forecast both future mortality rates and highest attained ages.

Cairns–Blake–Dowd (CBD) Model

The Cairns et al. (2006) model is designed for modelling old-age mortality and is particularly well suited to forecasting at very high ages (Currie, 2011). In this model,

$$\log h_{x,i} = \kappa_{0,i} + \kappa_{1,i}(x - \bar{x}), \tag{14}$$

where $h_{x,i}$ is the force of mortality in year i at age x, $\bar{x}$ is the mean of the ages used in fitting the model, and $\kappa_{0,i}$, $\kappa_{1,i}$, are latent time-varying variables typically modelled as a bivariate random walk with drift. The model can be rewritten in the form of a Gompertz model with time-varying parameters:

$$\log h_{x,i} = \log B_i + x \log C_i, \tag{15}$$

where $B_i = \exp(\kappa_{0,i} - \kappa_{1,i}\bar{x})$ and $C_i = \exp(\kappa_{1,i})$.

Extensions to the CBD model for mortality modelling of very old ages include adding a quadratic term to improve the fit to the Continuous Mortality Investigation (CMI) assured lives dataset (Currie, 2011). The CBD model is widely used in practice and in Section 6 we use it to benchmark the effectiveness of our proposed DSTLT model.

3. Data

This paper combines Netherlands mortality data from the Human Mortality Database (HMD) and a dataset provided by the Centraal Bureau voor de Statistiek (CBS). For the remainder of this paper, we will refer to data from these sources as HMD data and CBS data, respectively.

The Human Mortality Database is a joint initiative of the Department of Demographics at the University of California, Berkeley and the Max Planck Institute for Demographic Research (Human Mortality Database, 2019). It provides detailed mortality and population data for 40 countries, including the Netherlands. The HMD data for the Netherlands is gathered from various sources, including the Netherlands Interdisciplinary Demographic Institute (NIDI), Eurostat, and the CBS (Jasilionis, 2018). The data contains age and gender-specific death and number-alive counts from 1850–2016. The counts are given annually and thus are interval censored; they are also right-censored at age 100.

The CBS data consists of the ages at death (to the nearest day) of all age $\geq$ 92 Dutch residents whose year of death was between 1986–2015. Calculating the number of persons alive at any time from this data requires two assumptions. One is that every individual in the cohort concerned has died (and thus occurs as an observation) by the end of our observation period. This is because the dataset only records deaths and not numbers alive. The other assumption is that net migration is negligible for high (>92) ages. With these caveats, we can deduce the number of persons alive at each age and year — so approximation techniques for the exposure such as those employed in the HMD data (Jasilionis, 2018) need not be used.

The CBS dataset contains only death counts, so to check whether the HMD and CBS datasets do indeed describe the same population, we computed differences in the number of gender, age and year-specific deaths between the two datasets. The results are shown in Table 11 of Appendix A. For the CBS data, we calculated the age-at-death in years by dividing the age-at-death in days (as given in the data) by 365.25 and rounding down to the nearest integer. The number of deaths in each cell is then found by counting the number of deaths that fall in the corresponding age and year interval. This is essentially the same as defining age as age last birthday, which is the definition used in the HMD.

The method can be seen to be accurate in that the difference matrices for the number of deaths between the two datasets have very small elements. Apart from two discrepancies in two year×age-group cells of 17 and 11 persons, all discrepancies are in single digits.[2] These are small differences given the large number of deaths in each year×age-group cell, which are typically in the thousands or several hundreds (Table 12 of Appendix A). For our analysis, we ignore these minor discrepancies and treat the two datasets as compatible to be combined together.

The combined dataset is constructed as follows. For ages above 92, the death counts in each yearly age interval are computed using the CBS data. The number of persons alive at age x ($x >$ 92) is then calculated as $l_x = \sum_{i=x}^{u} d_i$, where u is the largest observed age at death in the raw data. This assumes all changes in population numbers are from deaths rather than migration. We can then compute the empirically estimated probability of death at age x as $q_x = d_x/l_x$. For ages less than or equal to 92, we extract q_x from the cohort life table of the HMD. The q_x are computed directly from the raw data and no smoothing at older ages is performed (Human Mortality Database, 2019). Combining the computed probabilities of death q_x obtained from the two databases gives an augmented dataset for all birth cohorts from 1893 to 1908. We plot the observed death probabilities, q_x, in Fig. 1. Since there are zero death probabilities for some very high ages, the ages in the combined dataset are right censored at age 108 for females and at age 107 for males for modelling purposes.

[2] These discrepancies necessarily occur at ages less than 100.

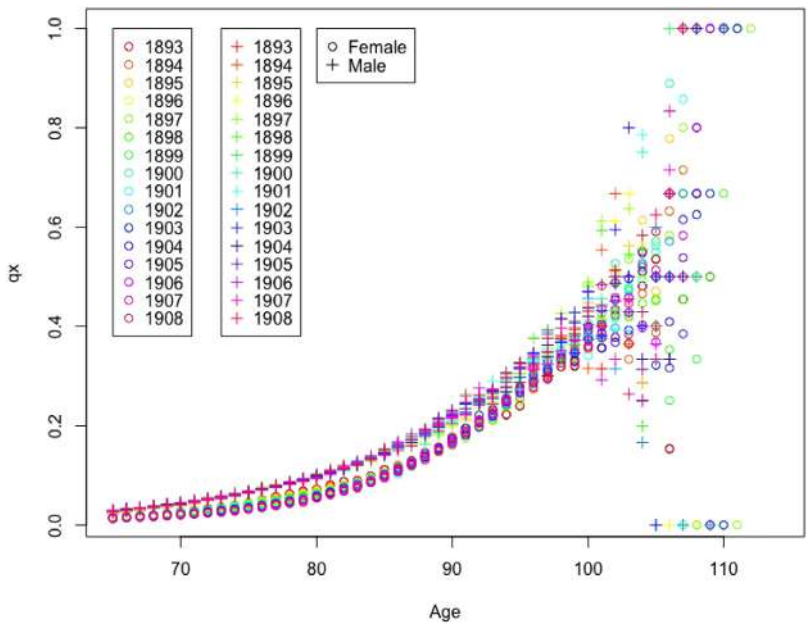

Fig. 1. Observed Netherlands Death Probabilities, 1893–1908 cohorts.

4. Methodology

In this section, we introduce a smoothing constraint to the threshold life table (TLT) model and compare the resulting smooth threshold life table (STLT) with the original Li et al. (2008) TLT model, the Heligman–Pollard model, the Coale–Kisker method, and the logistic models mentioned in Section 2.1. The models are fitted to each female, male and birth-year cohort separately[3] and we examine the effect of fitting with and without the extra CBS data. We use the data of cohorts 1893 to 1901 for in-sample fitting, and cohorts 1902 to 1908 for out-of-sample forecasting. Mortality modelling and forecasting is restricted to ages 65 and above only, as those ages are most relevant for retirement-related insurance products and longevity management.

4.1. TLT And STLT models

The likelihood, for censored data (single-year interval censored from age 65 to age $\tau-1$, and right censored at age τ, where τ is the observed maximum attained age for the cohort), is

$$L(B,C,\gamma;N) = \left[\prod_{x=65}^{N-1}\left(\frac{S(x)-S(x+1)}{S(65)}\right)^{d_x}\prod_{x=N}^{\tau-1}\left(\frac{S(x)-S(x+1)}{S(65)}\right)^{d_x}\right] \times \left(\frac{S(\tau)}{S(65)}\right)^{l_\tau}. \tag{16}$$

Here $S(x)$ is the survival function of the relevant lifetime distribution. For the TLT model as set up in (11) and (12), the log-likelihood can be decomposed into two parts (cf. Eq. (4.4) of Li et al., 2008). The first part is contributed by the Gompertz model, with

$$S(x) = \exp(-B/\ln C(C^x-1)), \quad 65 \le x < N, \tag{17}$$

and

$$l_1(B,C;N) = \sum_{x=65}^{N-1}\ln\left(\frac{S(x)-S(x+1)}{S(65)}\right)^{d_x} + l_N((\ln S(N) - \ln S(65)) = \sum_{x=65}^{N-1} d_x \ln(S(x)-S(x+1)) + l_N \ln(S(N)) - l_{65}\ln(S(65)) \tag{18}$$

(because $\sum_{x=65}^{N-1} d_x + l_N = l_{65}$). The second part is contributed by the GPD model, with

$$S(x) = S(N)(1+\gamma((x-N)/\theta))^{-1/\gamma}, \quad N \le x < \tau, \tag{19}$$

and

$$l_2(\gamma,\theta;N) = \sum_{x=N}^{\tau-1}\ln\left(\frac{S(x)-S(x+1)}{S(N)}\right)^{d_x} + l_\tau((\ln S(\tau) - \ln S(N)) = \sum_{x=N}^{\tau-1} d_x \ln(S(x)-S(x+1)) + l_\tau \ln(S(\tau)) - l_N \ln(S(N)) \tag{20}$$

(because $\sum_{x=65}^{\tau-1} d_x + l_\tau = l_N$).

This shows that the parameters for the Gompertz and GPD parts can be estimated separately for each integer value of $N \ge 85$. The value of N which maximises the profile log-likelihood, $l_p = l_1(\hat{B},\hat{C};N) + l_2(\hat{\gamma},\hat{\theta};N)$, is then chosen as the MLE. We found that the log-likelihood levels off after around age 95 for all cohorts, and selected the minimum age at which this occurs so as to retain the maximum number of observations available for fitting the extreme value part of the model.

For the smooth-constrained (STLT) model, we impose the constraint $h_1(N) = h_2(N)$, where h_1 and h_2 are the hazard functions of the Gompertz and GPD parts of the distribution respectively. This ensures a smooth changeover in hazards between the two parts. The constraint allows us to eliminate one parameter from (17) and (19). A convenient choice is to set

$$\theta = \frac{1}{C^N B} \tag{21}$$

[3] The threshold life table is more appropriately applied to cohorts than to period data because observations are assumed to be independently and identically distributed with constant parameters with respect to time — so the continuous age-at-death random variable should have the same distribution for all individuals to which the method is applied. In consequence of this our analysis is cohort-by-cohort.

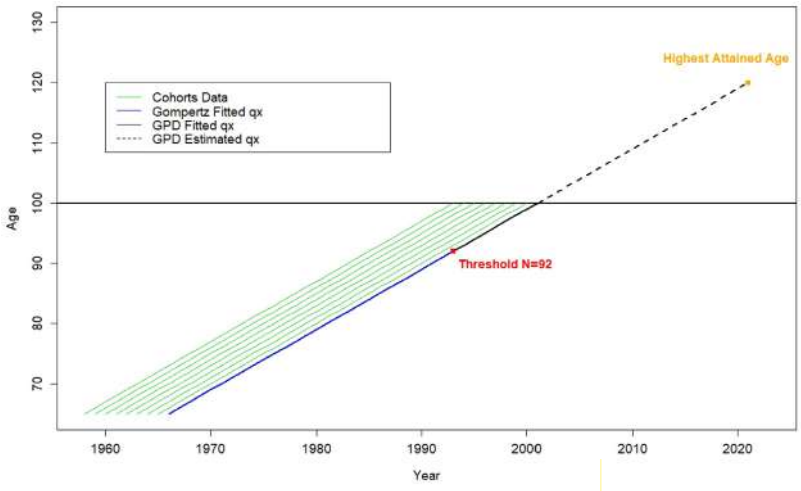

Fig. 2. Schematic for threshold life table fit to the cohort born in 1901, HMD only data, N = 92.

(see Appendix B for derivation) so that, for the STLT, the two component distributions become

$$F(x) = 1 - \exp\left(-\frac{B}{\ln C}(C^x - 1)\right), \qquad x \le N, \tag{22}$$

and

$$F(x) = F_\gamma(x) = \begin{cases} 1 - S(N)(1 + \gamma(C^N B(x-N)))^{-1/\gamma}, \\ \quad \gamma > 0,\ x > N \\ 1 - S(N)\exp(-(C^N B(x-N))), \\ \quad \gamma = 0,\ x > N \\ 1 - S(N)(1 - |\gamma|(C^N B(x-N)))^{1/|\gamma|}, \\ \quad \gamma < 0,\ N < x < N + \theta/|\gamma|. \end{cases} \tag{23}$$

The log-likelihood no longer separates into two disjoint components, and thus we must fit (16) directly to the entire dataset after taking into account (21), (22) and (23).

The TLT and STLT models do not explicitly take migration into account. Consequently, raw death counts cannot be directly used to fit the model. For example, positive net migration would increase the observed number of deaths. A model fitted to the raw death counts would then overestimate death rates, since there are more deaths than would otherwise be observed if there were no migration.

To account for the possibility of migration, we adjust the raw data to create a hypothetical cohort as follows:

1. Set $l'_{65} = l_{65}$;
2. then for $x = 65, 66, \ldots$, with q_x as the empirical conditional probability of death between ages x and $x+1$,
   (a) calculate the number of deaths at age x from $d'_x = l'_x q_x$;
   (b) calculate $l'_{x+1} = l'_x - d'_x$.

The quantities d'_x, l'_x of the hypothetical cohort for $x = 65, 66, \ldots$, are then used in place of d_x, l_x in fitting the model.

For each cohort (slanting green lines in Fig. 2), we fitted a Gompertz model using data before the threshold age N (blue solid line). For ages greater than N (black solid line), we fitted a GPD and then extrapolated to the upper bound of the ages at death (black dashed line). Only data up to age 100 years are available from the HMD (as shown in Fig. 2).

We used the maximum likelihood method for estimation of parameters, with the information matrix to assess precision of estimation. To improve numerical stability, we reparametrised B as e^{α} and C as $e^{e^{\delta}}$.

4.2. Highest attained age

Let $Y = X - N$ denote the exceedance of an individual aged X over threshold age N. In both the TLT and STLT, the distribution of Y conditional on $X > N$ is a GPD of the form:

$$F_Y(y) = \begin{cases} 1 - (1 + \gamma y/\theta)^{-1/\gamma}, & y > 0,\ \gamma \neq 0 \\ 1 - e^{-y/\theta}, & y > 0,\ \gamma = 0. \end{cases} \tag{24}$$

When $\gamma < 0$, the highest possible age at death of individuals in the population, according to the model, is

$$\omega = N - \frac{\theta}{\gamma} = N + \frac{\theta}{|\gamma|}. \tag{25}$$

Consequently, when the estimator $\hat{\gamma}$ of γ is negative, we can take $\hat{\omega} = \hat{N} + \hat{\theta}/|\hat{\gamma}|$ to be the estimated theoretical end point of the life table. The asymptotic variance of $\hat{\omega}$ can be computed from the covariance matrix of $(\hat{N}, \hat{\theta}, \hat{\gamma})$ via the delta method. We calculated these quantities for the TLT and STLT models with and without the CBS data.

5. Empirical results and analysis

We present empirical results and analysis for both the STLT and DSTLT models in this section. To illustrate the methodology, we give detailed results first for a representative cohort — namely, Netherlands females born in 1901. Further calculations confirm that the resulting implications and conclusions remain largely consistent across all cohorts in our dataset (see Appendices C and D).

5.1. Including CBS with HMD data

Our first aim is to investigate the relative merits of including the CBS data with the HMD data for the purpose of parameter estimation. For both the TLT and STLT models, Table 1 shows the estimates for γ and ω with their standard errors, and the estimated threshold age N.

The estimated highest attained ages for the HMD-only regimes are smaller than those for the regimes incorporating the CBS data. For the TLT model using the HMD-only data, the upper bound of the central 95% confidence interval for ω is less than 108. As the highest-aged survivor in this cohort lived to 108 years, this casts

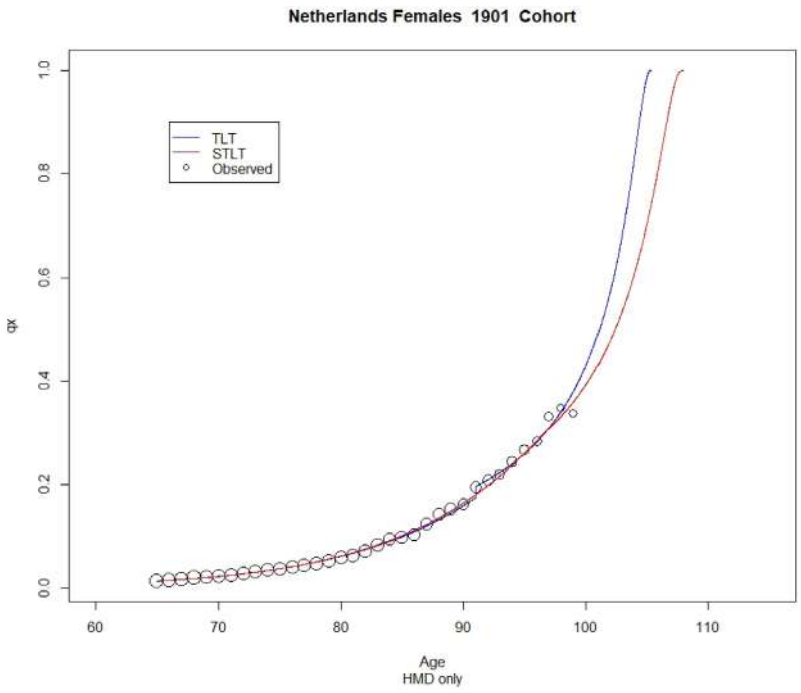

Fig. 3. TLT and STLT fitted curves with HMD data only for female 1901 cohort. Size of point is proportional to log exposure.

Table 1
Estimated γ, highest attained age ω, and threshold age N, female 1901 cohort.

Regime	$\hat{\gamma}$	SE($\hat{\gamma}$)	$\hat{\omega}$	SE($\hat{\omega}$)	N
TLT, HMD only	−0.307	0.0143	106.33	0.526	91
STLT, HMD only	−0.237	0.0451	108.98	2.245	97
TLT, HMD+CBS	−0.240	0.0070	109.59	0.388	91
STLT, HMD+CBS	−0.191	0.0132	111.78	0.976	97

some doubt on the suitability of using the TLT model and HMD data alone. Incorporating the CBS with the HMD data increases the estimated highest attained ages to realistic levels and also significantly reduces the standard errors of $\hat{\gamma}$ and $\hat{\omega}$.

Figs. 3 and 4 show the female 1901 cohort data and the fitted curves for the TLT and STLT models in the HMD and HMD+CBS datasets.[4] The death probabilities for the fitted curves are calculated from the estimated parameters using the definition

$$q_x = P(X < x+1 | X > x) = \frac{F(x+1) - F(x)}{1 - F(x)}. \tag{26}$$

For a TLT model, we use $F(x)$ as defined in (11) with the fitted parameters $\hat{A}, \hat{B}, \hat{C}$ for the Gompertz part, and for the GPD part we use $F(x)$ as defined in (12) with the fitted $\hat{\gamma}, \hat{\theta}, \hat{N}$. We note that there is a discontinuity in the fitted TLT function at the threshold age, due to the piecewise specification of the probability density functions. The smoothed threshold life table removes this discontinuity. The resulting smooth transition is in accordance with our intuition that there should not be a sudden change in mortality experience at any age, including at the threshold.

The fitted TLT and STLT models using both the HMD and the augmented datasets for females and males and all cohorts 1893–1901 are presented in Figs. 12–15 of Appendix C. The smoothing constraint in the STLT constitutes a functional link between the parameters of the Gompertz and GPD distributions, which tends to make the extreme age modelling of STLT more robust than the TLT model. All estimates of γ for the TLT and STLT fits were less than zero.

Overall, strong improvements in interpretability for both the TLT and STLT models are seen when incorporating the CBS data with the HMD.

To compare goodness of fit between models, we use the sum of squared errors (SSE) criterion defined by

$$\text{SSE} = \sum_{x=65}^{\tau} (q_x - \hat{q}_x)^2, \tag{27}$$

where q_x and $\hat{q}_x$ are the observed and estimated death rates at age x. We use SSEs for comparisons rather than a likelihood measure as some of the models we study are not comparable due to differences in the age ranges considered (and the Cole–Kisker method is not fitted by likelihood estimation). The SSEs for the TLT and STLT models are in Table 2.

The STLT provides no bigger SSEs than TLT for all but six of the 18 cohorts, and for those six the difference is small. At this stage, we conclude that adding the CBS to the HMD data significantly improves the fits of the models and adding the smoothing constraint improves on the threshold life table as an overall description of the data.

Table 2
SSEs for the TLT and STLT Models, HMD+CBS data.

Cohort	Females		Males	
	TLT	STLT	TLT	STLT
1893	0.17	0.18	0.45	0.45
1894	0.20	0.16	0.48	0.50
1895	1.00	0.93	0.26	0.26
1896	0.16	0.10	0.61	0.69
1897	0.72	0.80	0.34	0.39
1898	0.10	0.08	0.35	0.28
1899	0.33	0.27	0.12	0.14
1900	0.84	0.74	0.54	0.50
1901	0.08	0.08	0.51	0.51

[4] In Figs. 3 and 4 and similar subsequent figures, the sizes of the plotted points are proportional to the log exposure, so as to give a visual impression of the accuracy.

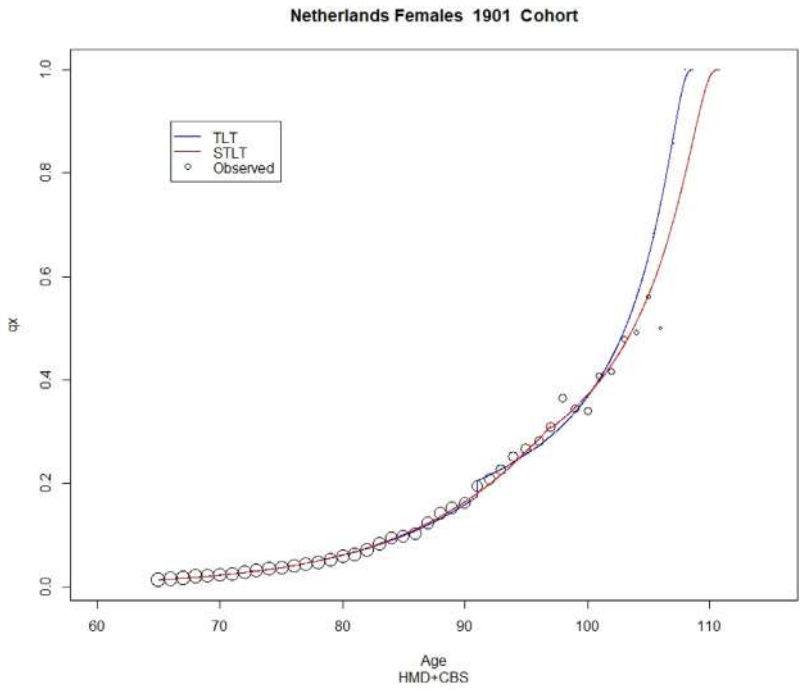

Fig. 4. TLT and STLT fitted curves with HMD and CBS data for female 1901 cohort. Size of point is proportional to log exposure.

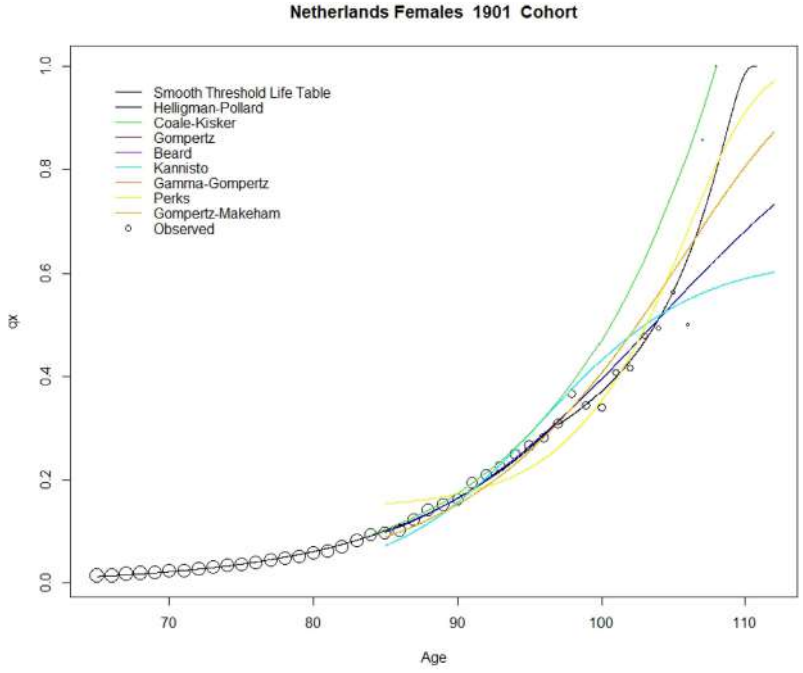

Fig. 5. Fitted lines for STLT, Heligman–Pollard, Coale–Kisker, Gompertz, Beard, Kannisto, Gamma–Gompertz, Perks, and Gompertz–Makeham models, female 1901 cohort. Size of point is proportional to log exposure.

5.2. STLT Vs current methods

The Heligman–Pollard, Coale–Kisker and logistic-type models are used widely in the insurance industry and by governments, and provide good benchmarks for comparison with the STLT model.

We fitted the Heligman–Pollard, Gompertz, Beard, Kannisto, Gamma–Gompertz, Perks and Gompertz–Makeham models to the Netherlands data using probabilities of death from age 85 and above for the 1901 cohort of females. For the Coale–Kisker method we started the extrapolation at age 85 and chose the end point of the extrapolation as age 108 (the highest observed attained age for this cohort). That is, we set $x_0 = 85, x_1 = 108$ and $m_{108} = 1$ in (10). The fits for all models are displayed in Fig. 5.

Here we can make several observations. Firstly, the curve fitted with the Coale–Kisker method consistently overestimates probabilities of death past age 92. Inaccuracy at extreme ages is likely with this method as the fitted curve depends strongly

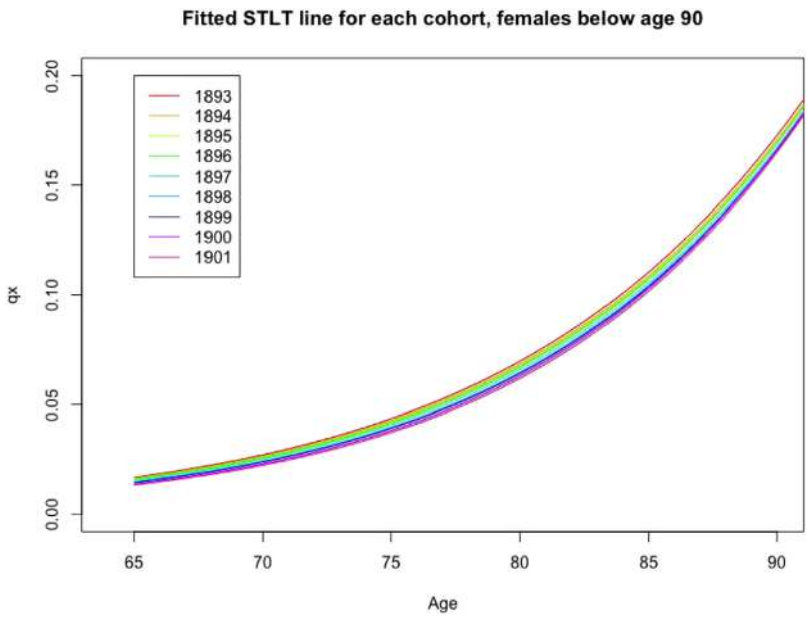

Fig. 6. Fitted STLT lines for each female cohort, shown for ages 90 and below.

on the start and end points of the extrapolation — an obvious weakness of the technique. The Heligman–Pollard model by contrast fits the data reasonably well up until around age 105 but severely underestimates the probability of death at higher ages. The Heligman–Pollard model is similar to a Gompertz (exponential extrapolation) law at high ages, and is clearly inappropriate once extrapolated beyond its range of applicability. For this cohort, the closely related Gompertz, Beard, Gamma–Gompertz and Gompertz–Makeham fits are very similar. The Kannisto model, which imposes that the force of mortality asymptote to 1, is a poor fit. The patterns for this cohort remain consistent over all cohorts analysed, with the logistic-type models (including Gompertz and Gompertz–Makeham) often producing similar fits.

The SSE for the STLT model is significantly smaller than for all other models.[5] Other cohorts of both genders show similar patterns, as seen in Figs. 16 and 17 of Appendix D.

6. The dynamic STLT model

In this section, we make the smooth threshold life table dynamic by modelling the cohort parameters as a function of time. The fitted DSTLT model is then compared with the CBD model using both in- and out-of-sample prediction SSEs as the evaluation criteria.

We start by examining how probabilities of death depend on age for each cohort. We treat the cohorts born in 1893–1901 as the training set, and cohorts born in 1902–1908 as the test set. The results shown in Figs. 6 and 7 are for females born in the years 1893–1901. We see that for lower ages (below 90), earlier cohorts tend to have higher probabilities of death. At higher ages, however, there is no clear pattern. Similar findings have also been observed in recent literature (e.g. Thatcher, 1999).

The estimated parameters for each cohort from the STLT model are in Tables 3 and 4 and the cohort-by-cohort variation of the estimated parameters for all cohorts (1893–1908)[6] is plotted in Figs. 8 and 9. For females, $\hat{B}$ has a noticeable downward trend, while $\hat{C}$ has an upward trend. For males, we observe an increasing trend in $\hat{B}$ for the first three cohorts (1893–1895) and then the trend levels off over the remaining cohorts. The reverse is true for $\hat{C}$. The pattern for males, with only the 1893–1895 cohorts having lower mortality compared to the rest, and the distinct difference from the female experience, may be due to historical and situational factors such as the impact of the two world wars in the Netherlands. Parameters $\hat{\gamma}$, $\hat{\theta}$, $\hat{\omega}$ and N do not have discernible trends for either gender. This finding is consistent with that of Einmahl et al. (2019).

Based on these observations, we model B in the STLT for the first 9 cohorts, 1893–1901, as the following function of time:

$$B_i = \exp(a + bi), \quad i = 1, 2, \ldots, 9, \tag{28}$$

Table 3
Cohort-by-cohort variation of estimated parameters, females.

Cohort	$\hat{B}$	$\hat{C}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\omega}$	N
1893	0.000031	1.1013	2.82	−0.159	114.74	97
1894	0.000029	1.1019	2.58	−0.160	114.13	98
1895	0.000028	1.1022	2.59	−0.205	110.67	98
1896	0.000026	1.1032	2.58	−0.173	112.95	98
1897	0.000024	1.1039	2.87	−0.207	110.88	97
1898	0.000021	1.1054	2.59	−0.154	114.77	98
1899	0.000020	1.1060	2.59	−0.171	113.10	98
1900	0.000017	1.1081	2.55	−0.208	110.24	98
1901	0.000015	1.1093	2.83	−0.191	111.78	97

Table 4
Cohort-by-cohort variation of estimated parameters, males.

Cohort	$\hat{B}$	$\hat{C}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\omega}$	N
1893	0.000071	1.0945	4.55	−0.221	109.55	89
1894	0.000078	1.0934	3.78	−0.195	110.42	91
1895	0.000086	1.0920	4.20	−0.221	109.00	90
1896	0.000103	1.0895	2.18	−0.176	110.38	98
1897	0.000119	1.0877	2.21	−0.179	110.37	98
1898	0.000104	1.0896	2.15	−0.151	112.27	98
1899	0.000112	1.0885	2.84	−0.179	110.82	95
1900	0.000106	1.0893	2.56	−0.134	115.17	96
1901	0.000118	1.0878	3.10	−0.181	111.12	94

[5] SSE(STLT) = 0.0836, SSE (Heligman–Pollard) = 0.216, SSE (Coale–Kisker) = 0.278, SSE (Gompertz) = 0.146, SSE (Beard) = 0.146, SSE (Kannisto) = 0.299, SSE (Gamma–Gompertz) = 0.146, SSE (Perks) = 0.113, SSE (Gompertz–Makeham) = 0.146.

[6] Cohorts 1893–1901 are used for model fitting; and cohorts 1902–1908 are used for out-of-sample validation.

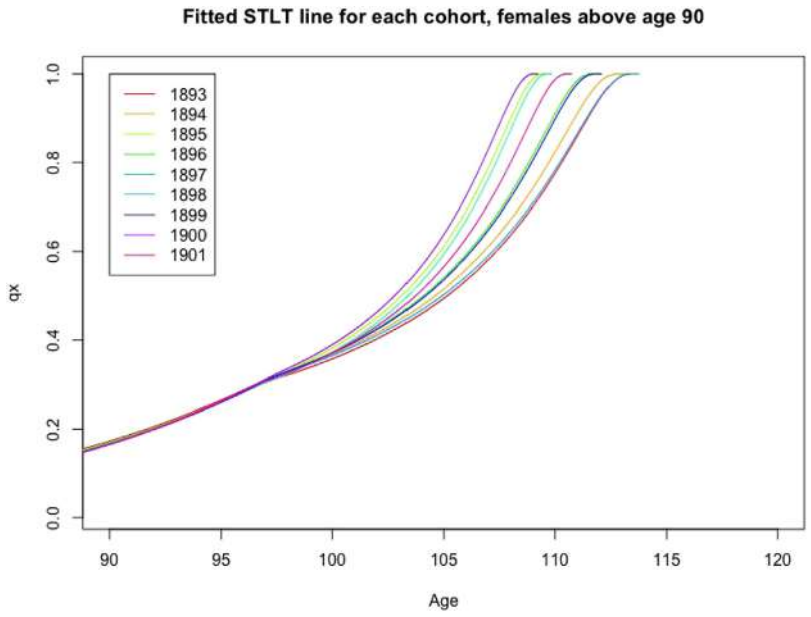

Fig. 7. Fitted STLT lines for each female cohort, shown for ages 90 and above.

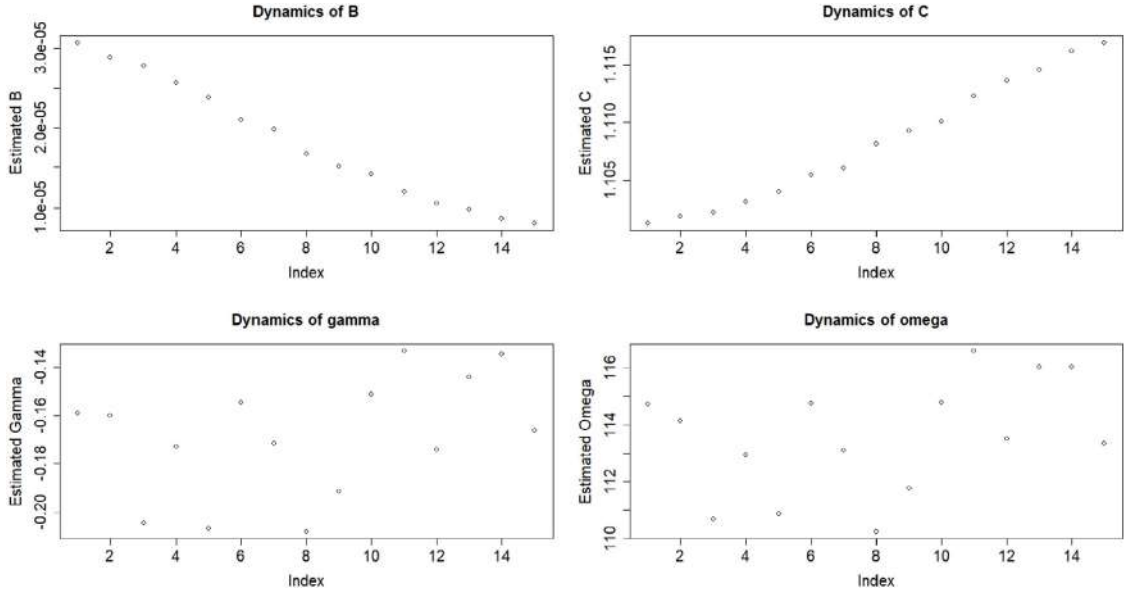

Fig. 8. Time series plots of estimated parameters for the female cohorts. Index 1 = 1893 cohort, index 2 = 1894 cohort, etc.

while keeping γ, N and θ constant. Modelling B in this fashion is consistent with Cairns et al. (2006), who model $\log(B)$ as a random walk with drift. In (28), a and b are constants and $i = 1, 2, \ldots, 9$ is an index representing the year of birth of a given cohort. Thus $i = 1$ corresponds to the cohort born in 1893, $i = 2$ corresponds to the cohort born in 1894, etc.

We call this modified model the dynamic smooth threshold life table (DSTLT). The DSTLT is formulated in the same way as in (11) and (12), with B replaced by $B_i = \exp(a + bi)$ and C replaced by $C_i = (1/(\theta B_i))^{1/N} = \theta^{-1/N}\exp(-(a + bi)/N)$. Thus, (11) is replaced by

$$F_i(x) = 1 - \exp\left(\frac{N\exp(a+bi)}{\ln\theta + a + bi}\theta^{-x/N}\left(\exp\big(-(a+bi)x/N\big) - 1\right)\right) \quad x \le N, \tag{29}$$

and (17) is replaced by

$$S_i(x) = \exp\left(\frac{N\exp(a+bi)}{\ln\theta + a + bi}\theta^{-x/N}\left(\exp\big(-(a+bi)x/N\big) - 1\right)\right) \quad x \le N, \tag{30}$$

but the GPD part remains as in (12). Estimation is performed by finding the values of a, b, θ and γ which maximise the likelihood, which now takes the form

$$L(a, b, \theta, \gamma; N) = \prod_{i=1}^{9}\left(\left(\prod_{x=65}^{N-1}\left(\frac{S_i(x) - S_i(x+1)}{S_i(65)}\right)^{d_{x,i}}\prod_{x=N}^{\tau-1}\left(\frac{S_i(x) - S_i(x+1)}{S_i(65)}\right)^{d_{x,i}}\right) \times \left(\frac{S_i(\tau)}{S_i(65)}\right)^{l_{\tau,i}}\right), \tag{31}$$

for a given N. Here $d_{x,i}$ and $l_{\tau,i}$ are the analogues of d_x and l_τ for cohort i. Then $\hat{N}$ is again chosen as the value of N that results

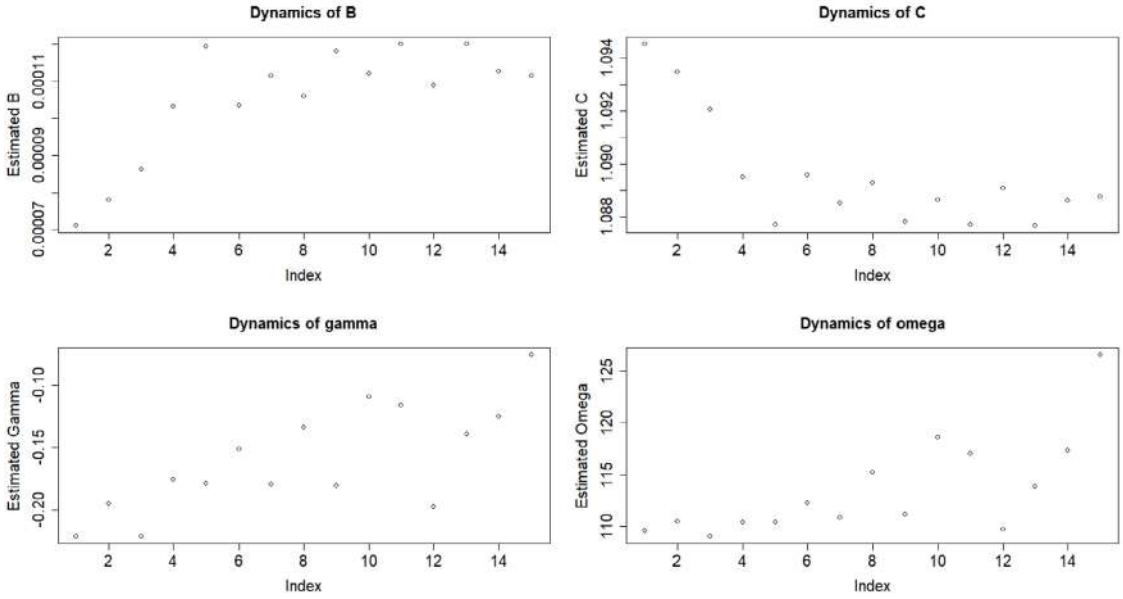

Fig. 9. Time series plots of estimated parameters for the male cohorts. Index 1 = 1893 cohort, index 2 = 1894 cohort, etc.

Table 5
Estimated parameters for the DSTLT model, female cohorts 1893–1901.

Parameter	Estimate	Standard error	95% confidence interval
a	−10.26	0.023	(−10.31, −10.22)
b	−0.085	0.0033	(−0.092, −0.079)
θ	2.58	0.0087	(2.56, 2.60)
γ	−0.174	0.0057	(−0.185, −0.163)

Table 6
Estimated parameters for the DSTLT model, male cohorts 1893–1901.

Parameter	Estimate	Standard error	95% confidence interval
a	−9.29	0.024	(−9.34, −9.24)
b	0.023	0.0033	(0.017, 0.029)
θ	2.57	0.010	(2.55, 2.59)
γ	−0.156	0.0074	(−0.170, −0.141)

in the highest maximised likelihood. This N is assumed to be constant across all cohorts.

Fitting this model, we obtained the parameter estimates in Tables 5 and 6, with the estimated threshold age $N = 98$ for females and $N = 96$ for males.

To determine whether the dynamic modification is a significant improvement over the static model, we test the hypotheses: $H_0 : b = 0$ vs $H_1 : b \neq 0$ using minus twice the difference in maximised log-likelihoods between the two models. Under H_0 this test statistic is asymptotically distributed as χ^2_1. The observed test statistic for females is 657.76, far exceeding the 0.001 value 10.83 of χ^2_1. We thus reject H_0 and conclude that there is a significant time-varying pattern. The same conclusion is obtained for males, though in this case the trend for the parameter B is weakly increasing with cohort.

DSTLT Model vs CBD Model

The suggested method in Currie (2011) for modelling advanced age mortality with the CBD model is to fit the model from age 65 to 89, and extrapolate to higher ages. The predicted q_x for the cohorts born in 1902–1908 under the DSTLT and CBD models are shown in Figs. 10 and 11.

For females, the DSTLT predictions are very close to the CBD predictions for ages below the estimated threshold age 98. From ages 98 to 106, the DSTLT predicts a lower probability of death than the CBD model does, consistent with the data. After age 106 the DSTLT predicts a higher probability of death than the CBD model, again consistent with the data though based on very scant data for these cohorts. For males, the interpretation is similar, but with the threshold age estimated to be 96 years. To make a formal comparison, we summarise the in- and out-of-sample prediction SSEs in Tables 7–10.

For the in-sample fitting, the DSTLT performs worse than the CBD model only for male cohorts of 1895, 1897, and 1899. The discrepancy in these cases invariably occurs at extreme ages. But at extremely high ages, the number of survivors is in single digits and observed probabilities of death may even be 0. (For example, the observed q_{107} for the 1899 male cohort is 0.) The SSE measure then overwhelmingly favours whichever model predicts

Table 7
In-sample fitting errors under the DSTLT and CBD models, females.

Cohort	Below age 98		Above age 98		All ages	
	DSTLT	CBD	DSTLT	CBD	DSTLT	CBD
1893	0.001	0.028	0.262	0.660	0.262	0.688
1894	0.000	0.020	0.125	0.308	0.126	0.328
1895	0.000	0.017	0.628	0.760	0.628	0.777
1896	0.000	0.012	0.093	0.194	0.094	0.206
1897	0.000	0.012	0.578	0.663	0.578	0.674
1898	0.000	0.013	0.170	0.353	0.171	0.366
1899	0.001	0.008	0.287	0.411	0.288	0.420
1900	0.001	0.006	0.604	0.643	0.605	0.649
1901	0.001	0.010	0.121	0.203	0.121	0.213

Table 8
Out-of-sample prediction errors under the DSTLT and CBD models, females.

Cohort	Below age 98		Above age 98		All ages	
	DSTLT	CBD	DSTLT	CBD	DSTLT	CBD
1902	0.001	0.009	0.030	0.178	0.031	0.187
1903	0.001	0.006	0.226	0.445	0.227	0.452
1904	0.001	0.005	0.053	0.214	0.054	0.220
1905	0.001	0.006	0.046	0.210	0.047	0.215
1906	0.001	0.004	0.064	0.200	0.066	0.204
1907	0.002	0.003	0.105	0.171	0.106	0.174
1908	0.002	0.002	0.132	0.148	0.134	0.151

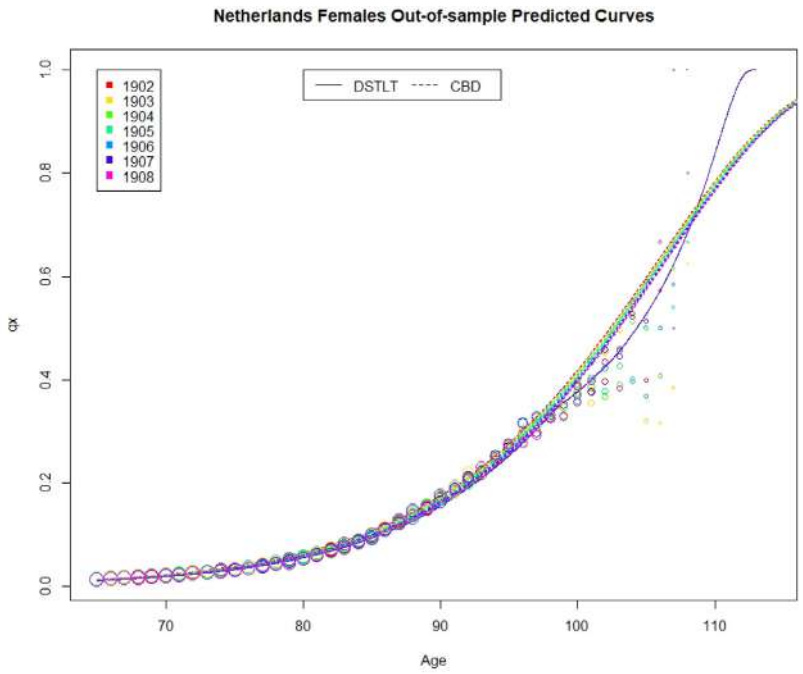

Fig. 10. Out-of-sample predicted DSTLT and CBD curves for the 1902–1908 female cohorts. Size of point is proportional to log exposure.

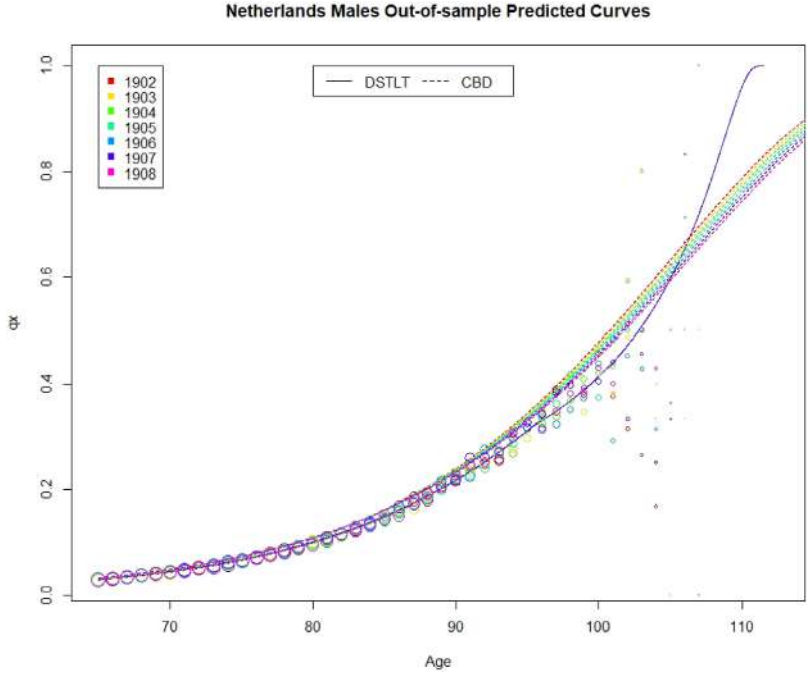

Fig. 11. Out-of-sample predicted DSTLT and CBD curves for the 1902–1908 male cohorts. Size of point is proportional to log exposure.

lower probabilities of death for that age. The DSTLT model outperforms the CBD model across all cohorts in terms of out-of-sample forecasting for both males and females.

For both in-sample fitting and out-of-sample prediction (see Figs. 18–21 of Appendix E), the CBD model and DSTLT perform similarly for ages under 90. From ages 90–105, the CBD model overestimates probabilities of death for almost all cohorts. For the very high ages past 105 years, the CBD model underestimates probabilities of death. The DSTLT also overestimates and underestimates probabilities of death similarly, but not as much as the CBD model.

7. Discussion

We have established for a high quality dataset (including mortality data for old ages beyond 100) that the STLT and DSTLT methods provide better modelling of old age mortality experiences compared with currently existing models, and have useful

Table 9
In-sample fitting errors under the DSTLT and CBD models, males.

Cohort	Below age 96		Above age 96		All ages	
	DSTLT	CBD	DSTLT	CBD	DSTLT	CBD
1893	0.001	0.043	0.262	0.609	0.262	0.652
1894	0.000	0.032	0.125	0.713	0.126	0.745
1895	0.000	0.050	0.628	0.428	0.628	0.478
1896	0.000	0.043	0.093	0.783	0.094	0.825
1897	0.000	0.032	0.578	0.509	0.578	0.541
1898	0.000	0.029	0.170	0.439	0.171	0.468
1899	0.001	0.027	0.287	0.220	0.288	0.247
1900	0.001	0.025	0.604	0.926	0.605	0.950
1901	0.001	0.038	0.121	0.839	0.121	0.877

Table 10
Out-of-sample prediction errors under the DSTLT and CBD models, males.

Cohort	Below age 96		Above age 96		All ages	
	DSTLT	CBD	DSTLT	CBD	DSTLT	CBD
1902	0.001	0.040	0.726	1.227	0.726	1.267
1903	0.000	0.044	0.197	0.573	0.197	0.616
1904	0.001	0.043	0.608	0.900	0.609	0.943
1905	0.000	0.047	0.186	0.431	0.187	0.478
1906	0.001	0.041	0.225	0.596	0.226	0.637
1907	0.001	0.043	0.295	0.625	0.296	0.668
1908	0.001	0.043	0.158	0.424	0.159	0.467

applications to insurance life table construction. The inclusion of the generalised Pareto distribution for very high ages in the TLT and STLT models is very natural in terms of extreme value theory and improves the model fits significantly for this data.

It is true of course that our analysis is very specific to the Netherlands dataset and inference in principle must be restricted to the Netherlands population over the relevant time period. But this high quality dataset serves as test data for the detailed comparisons we make on it with a view to wider application as more reliable data becomes available for other countries and periods.

The significance of the smoothing constraint introduced in this paper is multifaceted. It removes a discontinuity in the force of mortality that may occur at threshold age N in the threshold life table model. It provides a functional link between the Gompertz and GPD distributions, which tends to make the extreme age modelling of STLT more robust than that of the TLT model. Finally it encapsulates also a late life mortality deceleration as well as an advanced life mortality acceleration. We expand on this observation as follows.

In the literature, there have been three distinct philosophies regarding the tail behaviour of the human lifetime distribution. The first proposes that the probability of dying within 1 year may reach 1 at a finite age and thus there is a fixed upper limit to the length of human life (e.g. Vincent, 1951; Gbari et al., 2017; Ferreira and Huang, 2018; Einmahl et al., 2019). The second argues that this probability remains below 1 at finite ages, but nevertheless tends asymptotically to 1 (e.g. Heligman and Pollard, 1980; Gompertz, 1825). This is the "life is unlimited but short" (Rootzén and Zholud, 2017) scenario. The final view suggests that the probability of dying within 1 year may asymptote to a limit less than 1 — this is the case in all the logistic-type models (e.g. Thatcher, 1999; Beard, 1971; Perks, 1932; Kannisto, 1992; Vaupel et al., 1979). In this case, there is no fixed upper limit to the length of human life.

The TLT and STLT models provide an objective way of responding to this question and in fact we find (see Section 5) that all of our empirically analysed cohorts belong to the first case — they estimate a fixed upper limit to the human lifespan. This is consistent with the notion that humans do not live forever, possibly constrained by factors such as telomere attrition (Steenstrup et al., 2017).

Many researchers have observed a *late-life mortality deceleration effect* – that is, that the mortality increase with advanced age becomes slower than exponential (as in the Gompertz law) – and in some cases plateaus to a constant rate (e.g. Greenwood and Irwin, 1939; Gavrilov and Gavrilova, 2011; Barbi et al., 2018). This effect is usually understood to imply a model with an infinite

Table 11
Death count discrepancies for Males (left) and Females (right).

Year/Age	93	94	95	96	97	98	99	Year/Age	93	94	95	96	97	98	99
1986	3	0	4	2	2	1	1	1986	4	2	1	2	0	1	0
1987	2	3	0	0	1	1	0	1987	2	1	2	3	0	3	0
1988	1	2	2	2	1	1	0	1988	1	0	1	0	1	1	1
1989	3	2	1	0	1	0	0	1989	3	5	4	1	3	0	1
1990	1	3	1	3	3	1	0	1990	11	1	1	0	0	2	0
1991	2	1	0	1	0	1	0	1991	2	1	5	5	5	0	1
1992	5	3	0	1	0	1	0	1992	2	1	2	1	0	1	0
1993	1	1	0	1	2	2	0	1993	2	0	1	1	1	2	2
1994	1	2	2	1	1	1	0	1994	1	0	3	2	2	3	2
1995	3	1	1	2	1	0	4	1995	8	2	6	4	2	1	17
1996	0	2	1	0	1	0	1	1996	1	4	1	1	2	3	4
1997	2	2	0	4	2	0	1	1997	1	8	7	1	6	1	1
1998	2	3	1	1	1	0	2	1998	0	5	6	4	4	1	3
1999	3	3	1	2	2	0	0	1999	4	1	1	0	2	1	1
2000	0	0	1	1	0	0	1	2000	0	6	4	2	3	1	2
2001	3	1	2	0	1	0	0	2001	1	0	6	4	1	0	0
2002	2	2	1	2	3	0	0	2002	2	1	4	0	1	1	2
2003	3	0	2	0	0	0	0	2003	3	2	3	1	1	1	1
2004	0	3	3	0	0	1	0	2004	2	1	2	2	1	2	1
2005	6	1	0	0	0	0	0	2005	0	1	1	0	0	2	1
2006	1	0	0	0	0	1	0	2006	0	7	2	5	5	0	0
2007	0	2	0	1	1	0	0	2007	0	1	7	0	1	0	1
2008	0	0	0	1	0	1	0	2008	1	0	3	1	1	0	1
2009	0	0	0	1	0	1	0	2009	8	1	1	0	1	0	1
2010	1	0	1	0	0	1	0	2010	3	4	5	2	1	2	1
2011	3	0	1	0	1	1	0	2011	4	3	8	1	1	2	1
2012	1	0	1	0	2	2	0	2012	3	3	2	1	1	2	1
2013	5	0	0	0	0	0	0	2013	8	0	1	1	3	0	0
2014	1	2	0	0	0	0	0	2014	3	9	1	1	2	2	0
2015	1	4	1	0	0	0	1	2015	2	1	11	4	5	2	2

Table 12
HMD death counts for Males (left) and Females (right).

Year/Age	93	94	95	96	97	98	99	Year/Age	93	94	95	96	97	98	99
1986	438	392	276	175	142	107	66	1986	907	683	557	392	295	204	140
1987	446	378	273	190	129	72	72	1987	978	706	562	454	330	214	136
1988	427	359	284	221	154	96	62	1988	986	825	592	495	356	238	165
1989	466	349	283	227	158	112	64	1989	1140	899	675	544	398	281	185
1990	458	368	290	187	144	116	64	1990	1159	953	751	533	415	269	177
1991	477	399	309	207	172	126	58	1991	1148	964	750	584	431	294	194
1992	487	378	280	210	148	107	78	1992	1219	1013	752	595	456	329	201
1993	514	409	273	215	153	117	71	1993	1404	1120	912	683	505	344	239
1994	503	365	276	242	144	118	76	1994	1356	1120	881	680	482	356	223
1995	473	397	281	217	158	121	64	1995	1397	1159	872	693	518	362	218
1996	517	424	255	202	147	81	67	1996	1459	1219	924	716	530	381	253
1997	491	396	314	226	147	105	59	1997	1536	1278	1017	720	548	374	262
1998	528	358	308	225	152	93	70	1998	1557	1262	1029	824	568	409	265
1999	536	409	310	253	167	112	66	1999	1669	1368	1066	830	643	462	280
2000	478	422	321	218	166	96	82	2000	1645	1386	1104	796	589	430	276
2001	530	411	327	236	153	129	63	2001	1746	1457	1146	965	585	455	294
2002	531	437	309	215	168	110	63	2002	1793	1463	1156	878	654	474	287
2003	596	439	323	228	150	123	77	2003	1800	1488	1201	989	706	479	323
2004	543	420	308	208	168	110	74	2004	1714	1390	1185	878	638	492	311
2005	620	450	336	241	159	107	76	2005	1737	1417	1157	949	671	490	352
2006	642	475	331	230	148	114	64	2006	1864	1545	1233	951	705	538	357
2007	628	480	339	246	163	122	71	2007	1794	1519	1238	894	724	499	356
2008	663	528	373	272	162	133	79	2008	1912	1666	1333	1024	732	606	413
2009	637	496	391	270	163	117	84	2009	1797	1620	1321	1044	749	572	397
2010	654	533	384	313	178	147	86	2010	1906	1567	1346	1121	800	582	396
2011	677	474	389	291	217	138	91	2011	1880	1590	1337	1082	840	604	441
2012	702	607	446	290	231	158	115	2012	1846	1681	1454	1139	942	668	476
2013	920	550	452	307	239	158	108	2013	2269	1631	1477	1158	912	753	510
2014	955	719	414	330	231	151	111	2014	2361	1985	1385	1110	939	717	466
2015	1056	786	620	339	237	182	105	2015	2522	2197	1713	1218	979	706	491

human life span. But, contradictorily, although we do observe a deceleration effect with the STLT model fit to the Netherlands data, it is invariably accompanied with an estimated finite upper limit. In fact, there is no contradiction. The STLT and DSTLT model fits to the Netherlands data are in general agreement with the late-life mortality deceleration effect, as can be seen in Figs. 10 and 11; however, they also imply that there is an *advanced-age mortality acceleration* after the late-life mortality deceleration, which leads to the estimated finite limit to human life span.

To see that this effect is inherent in the STLT and DSTLT model structures, let $h_1(x)$ and $h_2(x)$ denote the forces of mortality at age x for the Gompertz and GPD parts of the distribution respectively. Then

$$\ln h_1(x) = x \ln C + \ln B \quad \text{and} \quad \ln h_2(x) = -\ln(\theta + \gamma(x-N)). \tag{32}$$

In our data, estimates of C are always greater than 1 and estimates of γ are always negative (Tables 3 and 4). So $h_1(x)$ is an increasing function of x for $x > 0$, while $h_2(x)$ decreases for $x < N$ and increases for $x > N$. The first derivatives with respect to x of $h_1(x)$ and $h_2(x)$ are equal when

$$x = x_c = N - \frac{1}{\ln C} - \frac{\theta}{\gamma}. \tag{33}$$

For ages $x > x_c$, the force of mortality under the GPD increases more rapidly than that of the Gompertz. The reverse is true for ages $N < x < x_c$. This advanced age mortality acceleration will be observed subsequent to the threshold age N if and only if $\theta \ln C > |\gamma| = -\gamma$. This is indeed always the case in our data.

In the DSTLT model, the force of mortality of different cohorts is the same at and above the threshold age, that is, $h_i(x) = h(x)$ for all $i = 1, 2, \ldots, 9$ and $x \geq N$. In particular, $h_i(x) = 1/\theta$ at $x = N$ for all i. This is consistent with the *compensation law of mortality*. The law states that for a given species, differences in death rates between different sub-populations decrease with age — because higher initial death rates tend to be accompanied by a lower rate of mortality increase with age, such that mortality rates for different sub-populations (cohorts in our case) become equal after some high age. The age at which this is presumed to happen is termed the "species-specific life span". This concept is captured by the threshold age N in the DSTLT model. The human species-specific life span (i.e. the threshold age N) is estimated as 98 years for females and 96 years for males using our data, which are close to the estimate of around 95 years made by Gavrilov and Gavrilova (1991).

The DSTLT model thus satisfies all three *laws of biodemography*[7]: *the Gompertz–Makeham law of mortality, the compensation law of mortality* and *late-life mortality deceleration* It also allows for a finite limit to human life span. To the best of our knowledge, no other model in the existing literature has all these properties.

R package

An R package to implement the models and reproduce results in this paper can be obtained via https://github.com/u5838836/STLT.

Acknowledgements

We are grateful to Professors Ana Ferreira and Sidney Resnick for helpful discussions in early stages of this work, and to the two reviewers whose comments helped us improve the paper.

Appendix A. Data validation

Table 11 shows the differences in the number of gender, age and year-specific deaths between the HMD and CBS datasets.

Table 12 shows the numbers of deaths in the HMD database in each year×age-group cell.

[7] https://en.wikipedia.org/wiki/Biodemography_of_human_longevity.

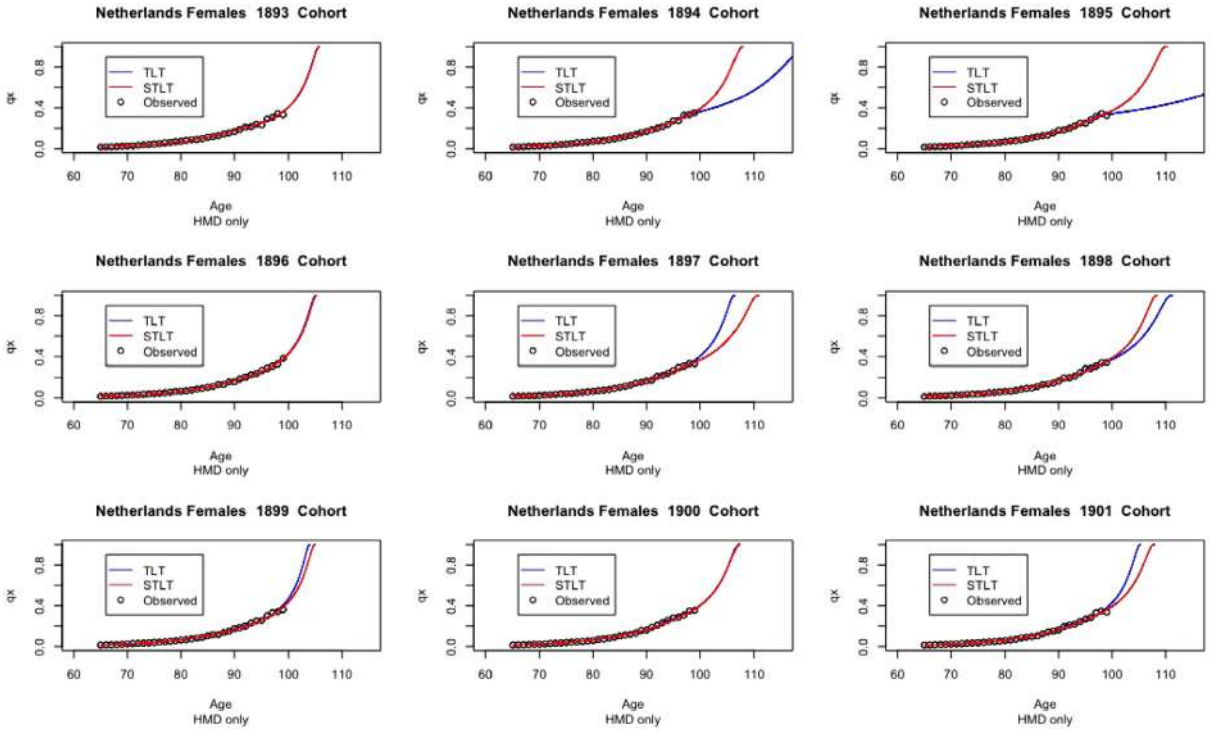

Fig. 12. Fitted TLT and STLT for all cohorts; HMD dataset only, females.

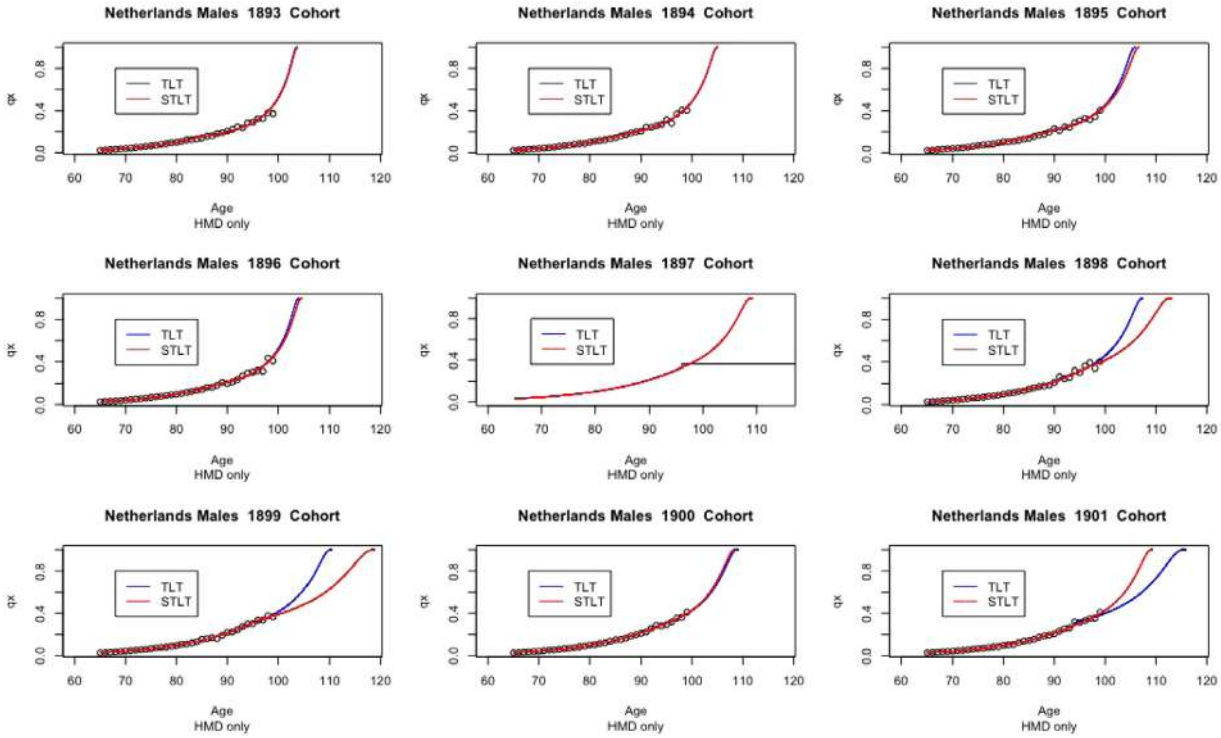

Fig. 13. Fitted TLT and STLT for all cohorts; HMD only, males.

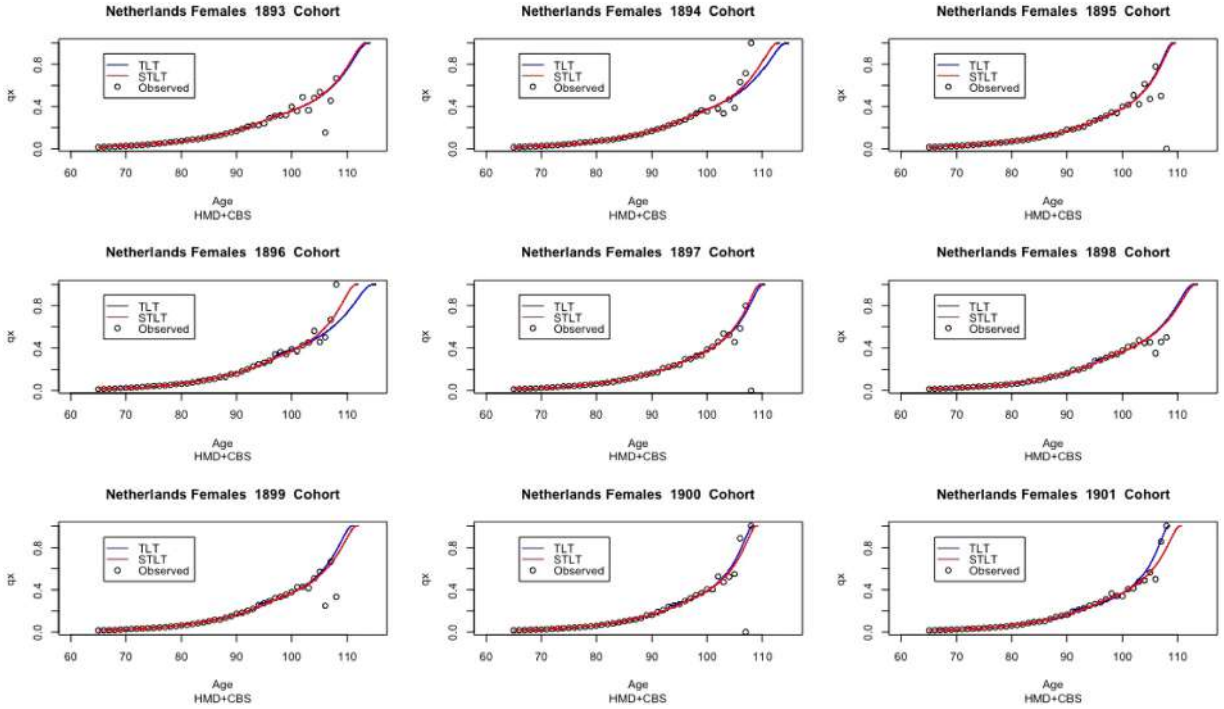

Fig. 14. Fitted TLT and STLT for all cohorts; HMD+CBS dataset, females.

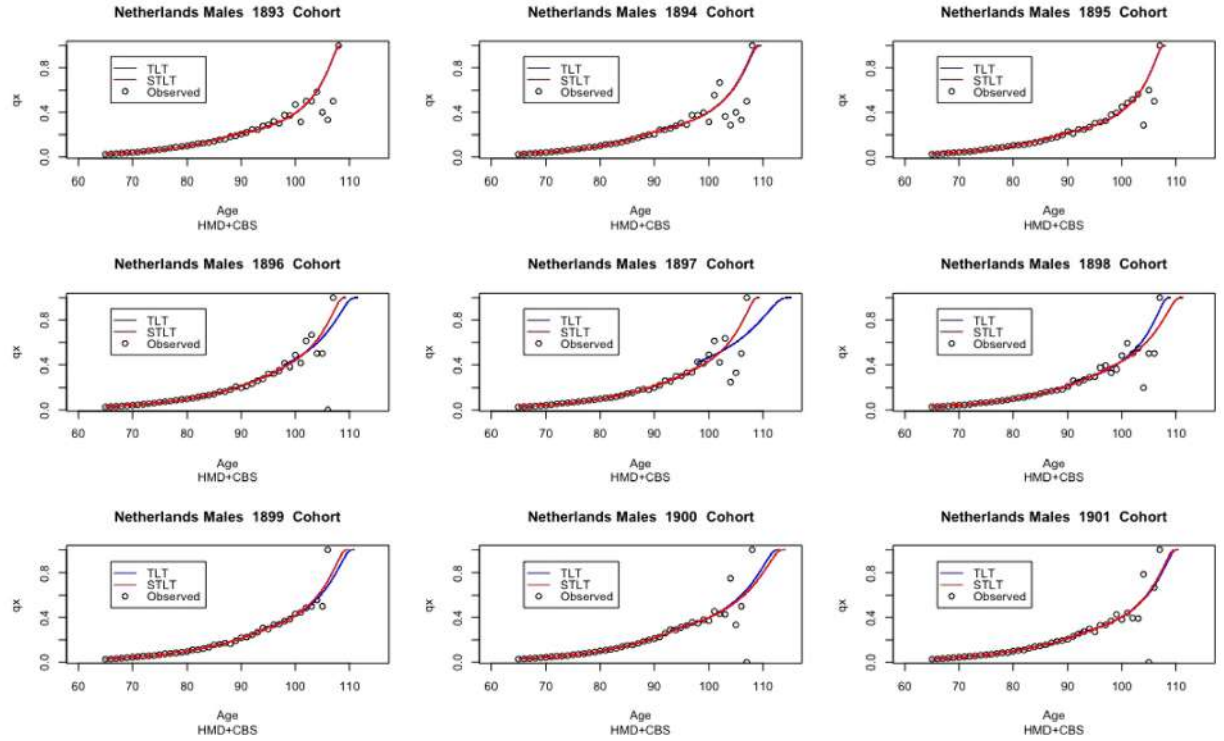

Fig. 15. Fitted TLT and STLT for all cohorts; HMD+CBS dataset, males.

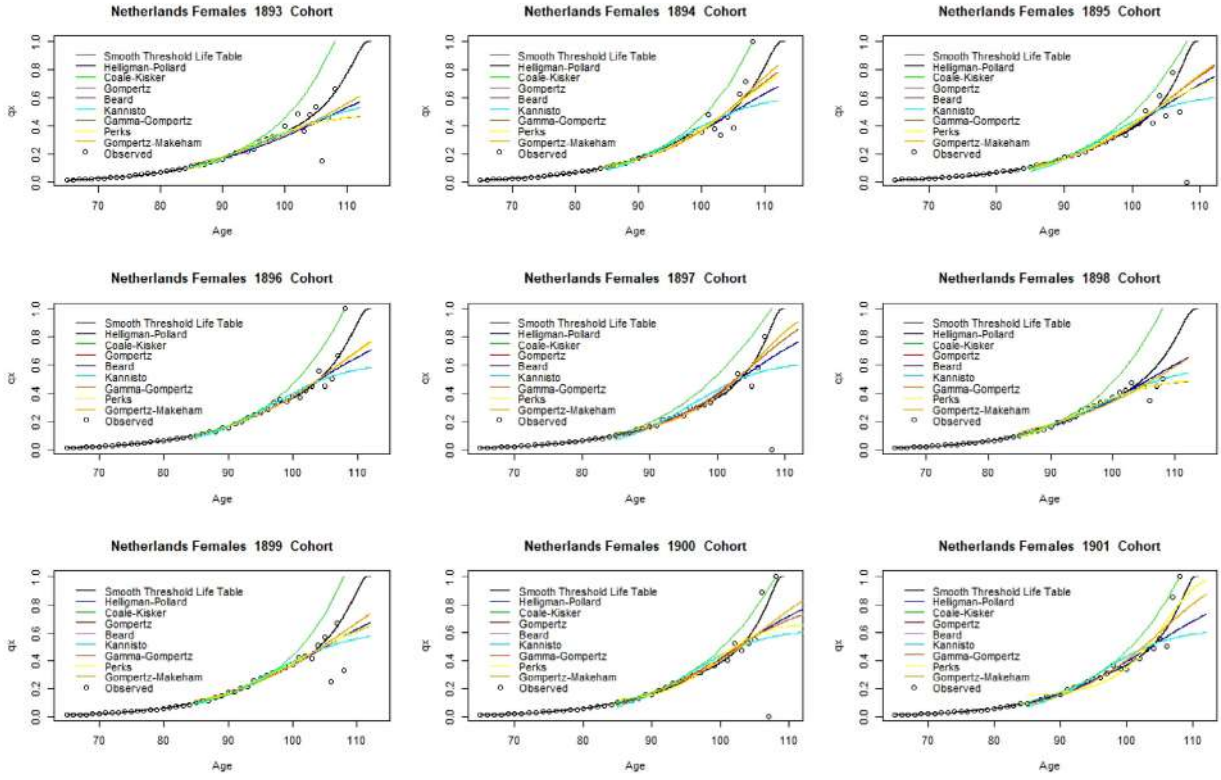

Fig. 16. Fitted lines for STLT, Heligman–Pollard, Coale–Kisker, Gompertz, Beard, Kannisto, Gamma–Gompertz, Perks and Gompertz–Makeham models for all cohorts, females.

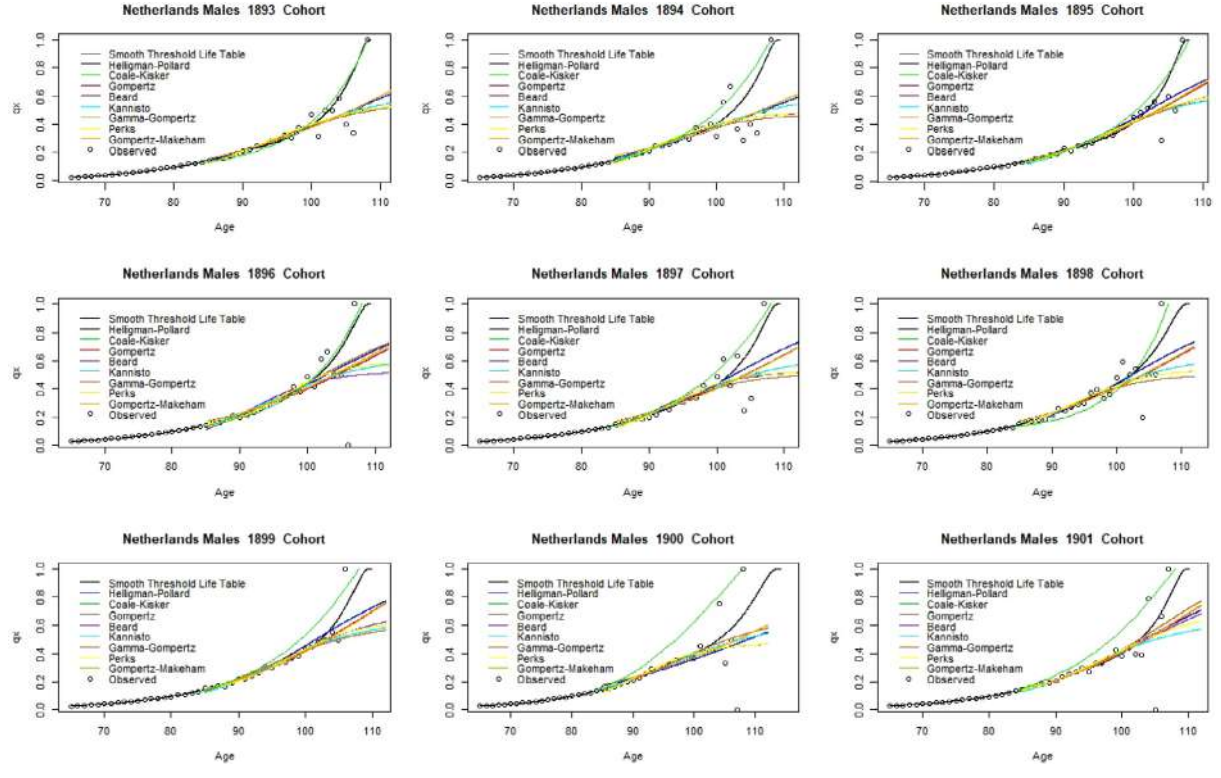

Fig. 17. Fitted lines for STLT, Heligman–Pollard, Coale–Kisker, Gompertz, Beard, Kannisto, Gamma–Gompertz, Perks and Gompertz–Makeham models for all cohorts, males.

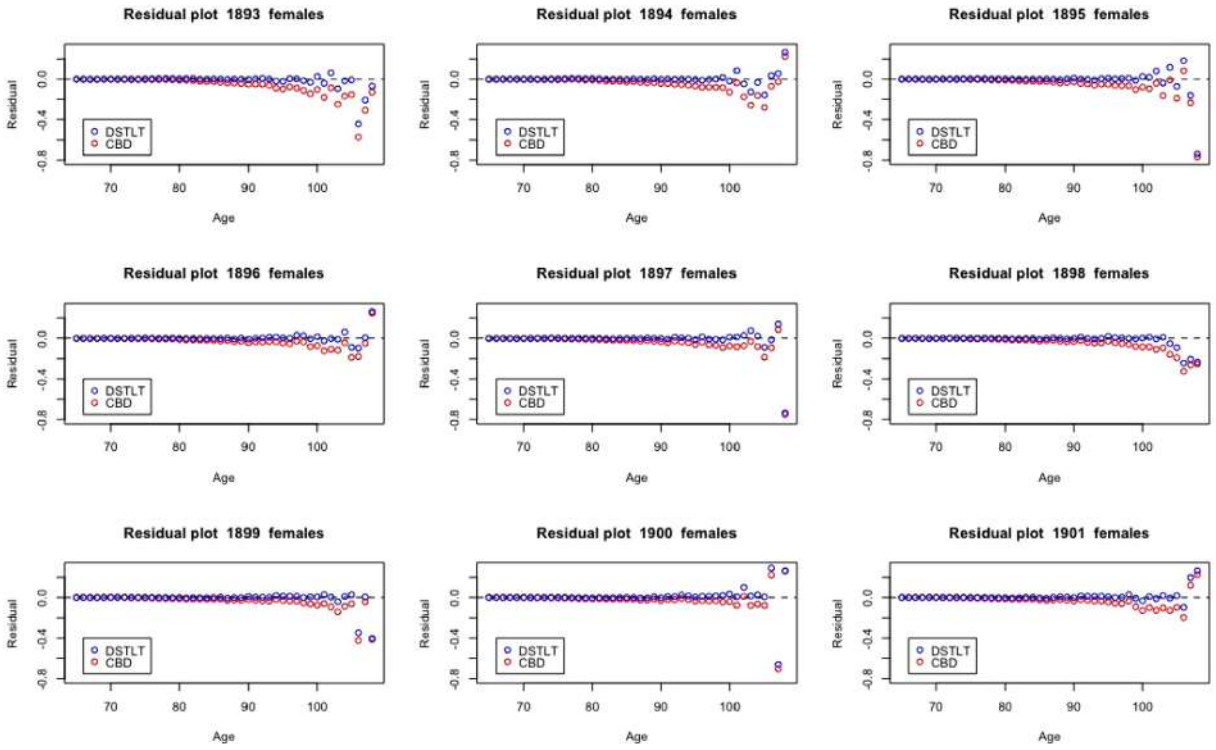

Fig. 18. DSTLT and CBD model residuals, females. Blue points correspond to the DSTLT; red points correspond to the CBD model. (For interpretation of the references to colour in this and other figure descriptions, the reader is referred to the web version of this article).

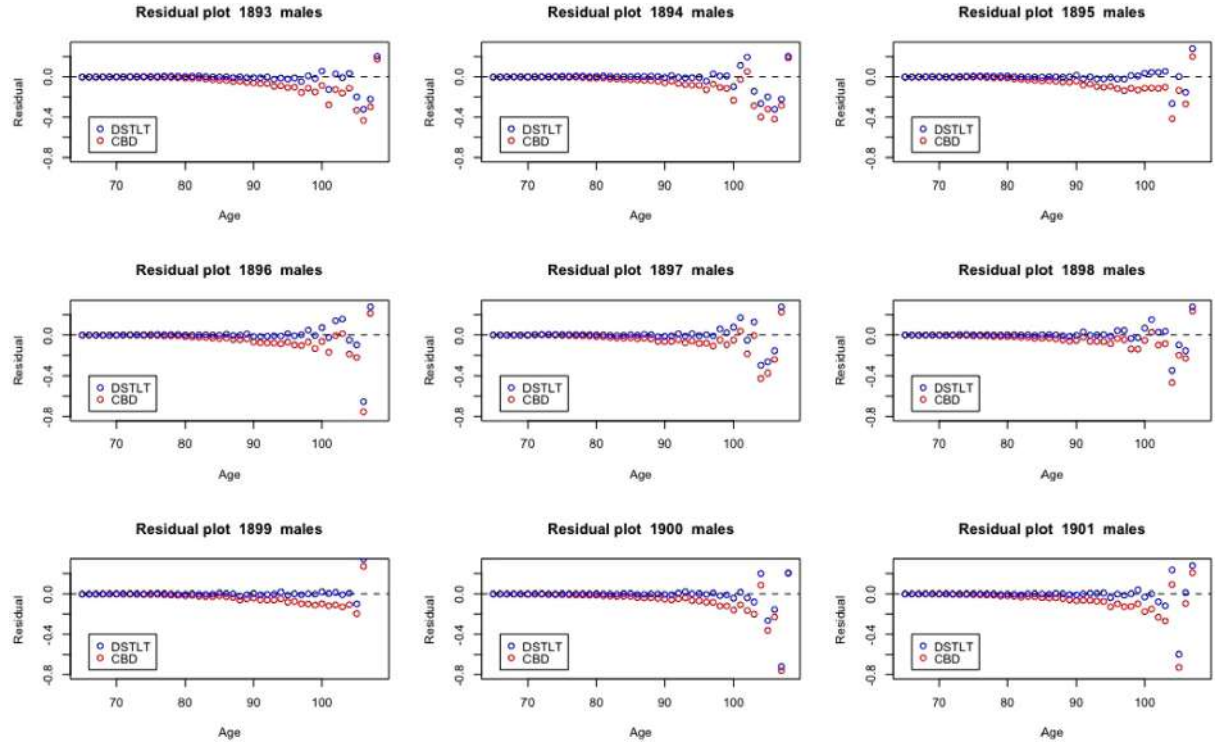

Fig. 19. DSTLT and CBD model residuals, males. Blue points correspond to the STLT; red points correspond to the CBD model.

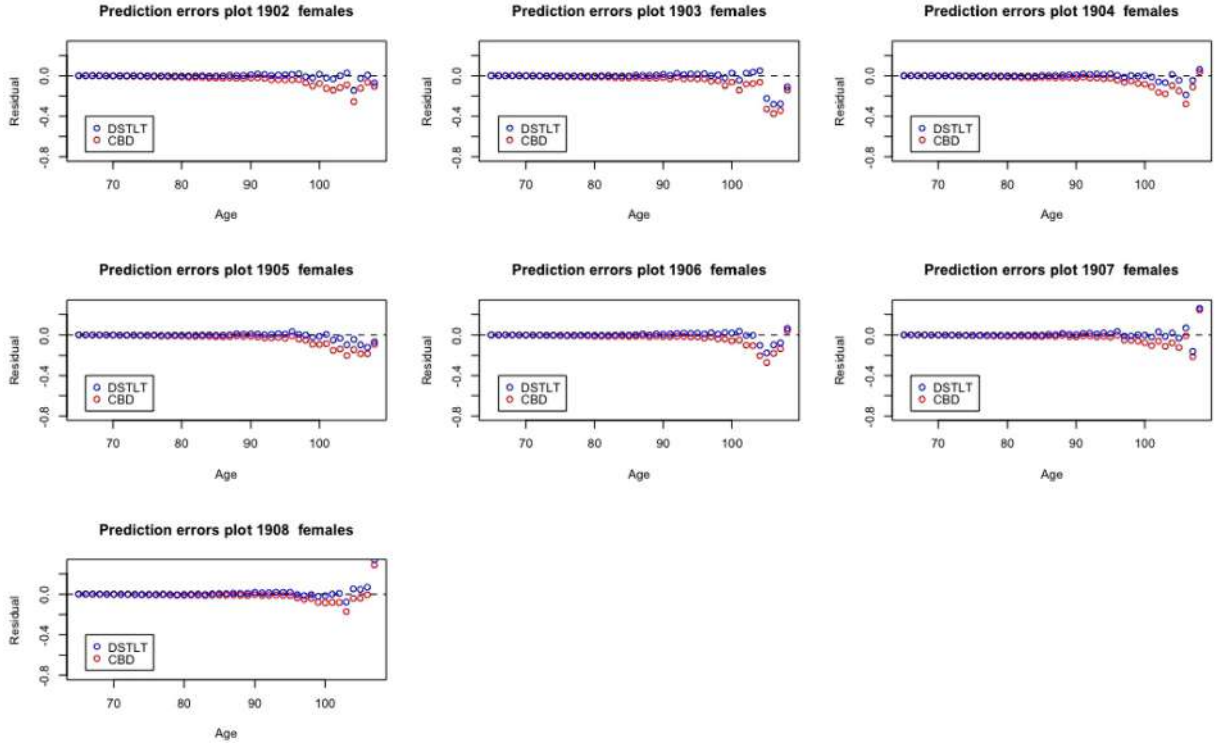

Fig. 20. Prediction errors under the DSTLT and CBD model, females. Blue points correspond to the DSTLT; red points correspond to the CBD model.

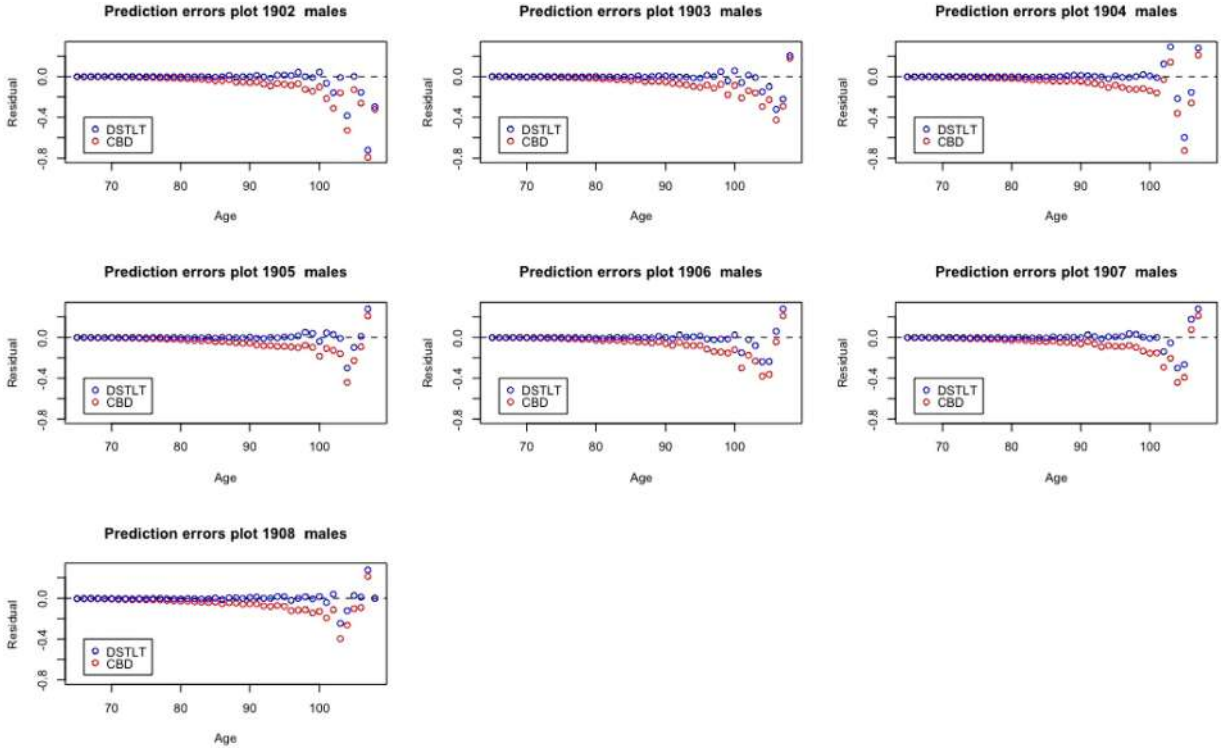

Fig. 21. Prediction errors under the DSTLT and CBD model, males. Blue points correspond to the STLT; red points correspond to the CBD model.

Appendix B. Derivation of smoothness constraint

In the threshold life table model defined by (11) and (12), we set $h_1(N) = h_2(N)$, where h_1 and h_2 are the hazard functions corresponding to the distributions in (11) and (12) respectively.

We can easily check that for the Gompertz part we have

$$f(x) = F'(x) = BC^x \exp\left(-\frac{B}{\ln C}(C^x - 1)\right),$$

so that

$$h_1(x) = \frac{f(x)}{1 - F(x)} = BC^x$$

For the GPD part we have

$$f(x) = F'(x) = \begin{cases} S(N)\,(1 + \gamma(x - N)/\theta)^{-(1+\gamma)/\gamma}\,/\theta, & \gamma \neq 0 \\ S(N)\exp(-(x - N)/\theta)/\theta, & \gamma = 0, \end{cases}$$

so that

$$h_2(x) = \frac{f(x)}{S(x)} = \frac{1}{\theta + \gamma(x - N)}, \quad \gamma \in \mathbb{R}.$$

Setting $h_1(N) = h_2(N)$, we find

$$\theta = \frac{1}{BC^N}.$$

So the distribution to be fitted is as given in (22) and (23).

Appendix C. Fitted TLT and STLT for all cohorts

See Figs. 12–15.

Appendix D. Heligman–pollard, coale–kisker and STLT comparisons for all male and female cohorts

See Figs. 16–17.

Appendix E. Fitting and prediction errors of DSTLT and CBD models

See Figs. 18–21.

References

Barbi, E., Lagona, F., Marsili, M., Vaupel, J.W., Wachter, K.W., 2018. The plateau of human mortality: Demography of longevity pioneers. Science 360 (6396), 1459–1461.

Beard, R.E., 1971. Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In: Biol. Aspects Demography. Taylor & Francis, London, pp. 57–68.

de Beer, J., Janssen, F., 2016. A new parametric model to assess delay and compression of mortality. Popul. Health Metr. 14 (1), 46.

Bernard, J., Vaupel, J., 1999. Validation of Exceptional Longevity. Odense University Press, Odense, Denmark.

Boleslawski, L., Tabeau, E., 2001. Comparing theoretical age patterns of mortality beyond the age of 80. In: Forecasting Mortality in Developed Countries. Springer, pp. 127–155.

Booth, H., Tickle, L., 2008. Mortality modelling and forecasting: A review of methods. Ann. Actuar. Sci. 3 (1–2), 3–43.

Bourbeau, R., Desjardins, B., 2002. Dealing with problems in data quality for the measurement of mortality at advanced ages in Canada. N. Am. Actuar. J. 6 (3), 1–13.

Bower, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., Nesbitt, C., 1997. Actuarial Mathematics, Schaumburg Illinois. Society of Actuaries.

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

Cairns, A., Blake, D., Dowd, K., 2008. Modelling and management of mortality risk: A review. SSRN Electron. J.

Coale, A., Kisker, E.E., 1990. Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. In: Asian and Pacific Population Forum, vol. 4. pp. 1–31.

Currie, I.D., 2011. Modelling and forecasting the mortality of the very old. Astin Bull. 41 (2), 419–427.

Currie, I.D., Durban, M., Eilers, P.H., 2004. Smoothing and forecasting mortality rates. Stat. Model. 4 (4), 279–298.

Duffie, D., Pan, J., 1997. An overview of value at risk. J. Deriv. 4 (3), 7–49.

Einmahl, J.J., Einmahl, J.H.J., de Haan, L., 2019. Limits to human life span through extreme value theory. J. Amer. Statist. Assoc. 114 (527), 1075–1080.

Embrechts, P., Klüppelberg, C., Mikosch, T., 2013. Modelling Extremal Events: For Insurance and Finance, vol. 33. Springer Science & Business Media.

Ferreira, A., Huang, F., 2018. Is human life limited or unlimited? (A discussion of the paper by Holger Rootzén and Dmitrii Zholud). Extremes 21 (3), 373–382.

Gavrilov, L., Gavrilova, N., 1991. The Biology of Life Span: A Quantitative Approach. CRC Press, Chur, Switzerland.

Gavrilov, L., Gavrilova, N., 2011. Mortality measurement at advanced ages: A study of the social security administration death master file. N. Am. Actuar. J. 15 (3), 432–447.

Gbari, S., Poulain, M., Dal, L., Denuit, M., 2017. Extreme value analysis of mortality at the oldest ages: A case study based on individual ages at death. N. Am. Actuar. J. 21 (3), 397–416.

Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality and on the mode of determining the value of life contingencies. Phil. Trans. R. Soc. A 115, 513–585.

Greenwood, M., Irwin, J.O., 1939. The biostatistics of senility. Hum. Biol. 11 (1), 1–23.

Heligman, L., Pollard, J.H., 1980. The age pattern of mortality. J. Inst. Actuar. 107 (1), 437–455.

Human Mortality Database, 2019. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany), www.mortality.org. (Accessed 1 July 2019).

Jasilionis, D., 2018. About Mortality Data for the Netherlands. Tech. Rep., Human Mortality Database, University of California, Berkeley and Max Planck Institute for Demographic Research.

Kannisto, V., 1992. Presentation at a Workshop on Old-Age Mortality. Odense University, Odense, Denmark.

Le Bras, H., 1976. Lois de mortalité et age limite. Population 31 (31), 655–692.

Lee, R.D., Carter, L.R., 1992. Modeling and forecasting U.S. mortality. J. Amer. Statist. Assoc. 87 (419), 659–671.

Li, J.S.-H., Hardy, M.R., Tan, K.S., 2008. Threshold life tables and their applications. N. Am. Actuar. J. 12 (2), 99–115.

Makeham, W.M., 1860. On the law of mortality and construction of annuity tables. J. Inst. Actuar. 8 (6), 301–310.

Olshansky, S., Carnes, B., 1997. Ever since Gompertz. Demography 34 (1), 1–15.

Perks, W., 1932. On some experiments in the graduation of mortality statistics. J. Inst. Actuar. (1886-1994) 63 (1), 12–57.

Purcal, S., 2006. Supply Challenges to the Provision of Annuities. Tech. Rep., University of New South Wales, Working Paper.

Rau, R., Ebeling, M., Peters, F., Missov, C.B.-E.T.I., 2017. Where is the level of the mortality plateau? In: Living to 100 Symposium of the Society of Actuaries, Orlando, FL'.

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

Rootzén, H., Zholud, D., 2017. Human life is unlimited – but short. Extremes 20 (4), 713–728.

Saikia, P., Borah, M., 2013. A study on mathematical modelling for oldest-old mortality rates. Int. J. Math. Arch. (IJMA) 4 (11), 1–9.

Steenstrup, Troels, Kark, Jeremy D., Verhulst, Simon, Thinggaard, Mikael, Hjelmborg, Jacob V.B., Dalgard, Christine, Kyvik, Kirsten Ohm, Christiansen, Lene, Mangino, Massimo, Spector, Timothy D., Petersen, Inge, Kimura, Masayuki, Benetos, Athanase, Labat, Carlos, Sinnreich, Ronit, Hwang, Shih-Jen, Levy, Daniel, Hunt, Steven C., Fitzpatrick, Annette L., Chen, Wei, Berenson, Gerald S., Barbieri, Michelangela, Paolisso, Giuseppe, Gadalla, Shahinaz M., Savage, Sharon A., Christensen, Kaare, Yashin, Anatoliy I., Arbeev, Konstantin G., Aviv, Abraham, 2017. Telomeres and the natural lifespan limit in humans. Aging 9 (4), 1130–1142.

Thatcher, A.R., 1999. The long-term pattern of adult mortality and the highest attained age. J. Roy. Statist. Soc. Ser. A 162 (1), 5–43.

Thatcher, A.R., Kannisto, V., Vaupel, J., 1998. The Force of Mortality At Ages 80–120 Monographs on Population Aging. Odense University Press, Odense.

Vaupel, J.W., Manton, K.G., Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16 (3), 439–454.
Vaupel, J., Robine, J.-M., 2002. Emergence of supercentenarians in low mortality countries. N. Am. Actuar. J. 6 (3), 52–63.
Vincent, P., 1951. La mortalite des vieillards. Population 6 (2), 181–204.
Watts, K.A., Dupuis, D.J., Jones, B.L., 2006. An extreme value analysis of advanced age mortality data. N. Am. Actuar. J. 10 (4), 162–178.
Wilmoth, J., 1995. Are mortality rates falling at extremely high ages: An investigation based on a model proposed by Coale and Kisker. Popul. Stud. 49 (2), 281–295.
Yashin, A.I., Vaupel, J.W., Iachine, I.A., 1994. A duality in aging: the equivalence of mortality models based on radically different concepts. Mech. Ageing Dev, 74 (1–2), 1–14.