Mortality and Longevity

Aging and Retirement

Calibrating Mortality Processes with Trend Changes to Multi-Population Data

January 2020

Calibrating Mortality Process with Trend Changes to Multi-Population Data

AUTHORS

Matthias Börger, Ph.D.
Technische Hochschule Nürnberg Georg
Simon Ohm, Nuremberg

Justin Schoenfeld
Institute of Insurance, Ulm University

Johannes Schupp (corresponding author)
Institute for Finance and Actuarial Sciences,
Ulm & Institute of Insurance, Ulm University

SPONSORS

Society of Actuaries:
Aging and Retirement Research
Committee on Life Insurance Research
Financial Reporting Section
Mortality and Longevity Research
Product Development Section
Reinsurance Section

Calibrating Mortality Processes with Trend Changes to Multi-Population Data

Matthias Börger

Technische Hochschule Nürnberg Georg Simon Ohm, Nuremberg

Keßlerplatz 12, 90489 Nürnberg, Germany

Justin Schoenfeld

Institute of Insurance, Ulm University

Helmholtzstraße 20, 89081 Ulm, Germany

Johannes Schupp (corresponding author)

Institute for Finance and Actuarial Sciences (ifa), Ulm & Institute of Insurance, Ulm University

Lise-Meitner-Straße 14, 89081 Ulm, Germany

Phone: +49 731 20644241

Email: j.schupp@ifa-ulm.de

December 20, 2019

Abstract

The uncertainty in future mortality rates is typically quantified by stochastic mortality models. To this end, the time dependent parameters in these models are projected by stochastic processes. Thus, the choice of these processes and their calibration have a crucial impact on estimates of future uncertainty. Since the commonly applied random walk with drift process has some structural shortcomings (see e.g. Börger et al. (2014)), alternative processes with random changes in the long-term mortality trend have been proposed by several authors. Such trend changes can be observed in the historical data for almost every population. However, data on such trend changes is sparse, and thus, the parameter estimation of these trend processes involves a significant degree of uncertainty.

In this paper, we explain how data on trend changes from several populations can be combined in order to improve the reliability of trend process calibrations for individual populations. We discuss different assumptions on the "similarity" of parameters for different populations and implement those assumptions for the case of the trend change process proposed by Börger and Schupp (2018). In a numerical example we find that the impact on parameter estimates can be substantial. Thus, relying on the sparse data for individual populations only can lead to significant misestimation of future mortality and its uncertainty.

1. Introduction and Motivation

Annuity providers, pension funds, life insurers, and social security systems heavily rely on forecasts of future mortality. For risk management purposes in particular, stochastic mortality models are required in order to quantify the uncertainty in future mortality rates. Therefore, a large number of such models have been developed over the last decades, e.g. the models of Lee and Carter (1992) and Cairns et al. (2006). Most of these models contain one or more time dependent parameters, often referred to as period effects. These parameters are typically projected into the future by stochastic trend processes in order to generate scenarios of future mortality. Obviously, the choice of a specific stochastic trend process has a crucial impact on the forecasts, both in terms of the central, median, or best estimate projection scenario as well as in terms of the uncertainty around this scenario.

Often, a random walk with drift is applied. It is a simple process with a clear parameter interpretation, and it nicely extrapolates the rather linear trends which have been observed in the period effects for many populations over the last decades. However, the random walk with drift also has some structural shortcomings. Most prominently, long-term uncertainty is often underestimated since the drift is fixed and stochasticity is only contained in the annual innovations (see, e.g., Börger et al. (2014) or Börger and Schupp (2018)). This issue is illustrated by Figure 1 which shows the logarithm of probabilities of death for 65-year old males in different countries from all over the world; the data has been obtained from the HMD (2019). We observe the aforementioned rather linear trends in most recent decades, but we also see that the drifts/trends in the log probabilities of death have also changed significantly once in a while in the past. A random walk with fixed drift is not able to generate such patterns, and in particular, it does not adequately allow for the uncertainty which arises from potential trend changes.

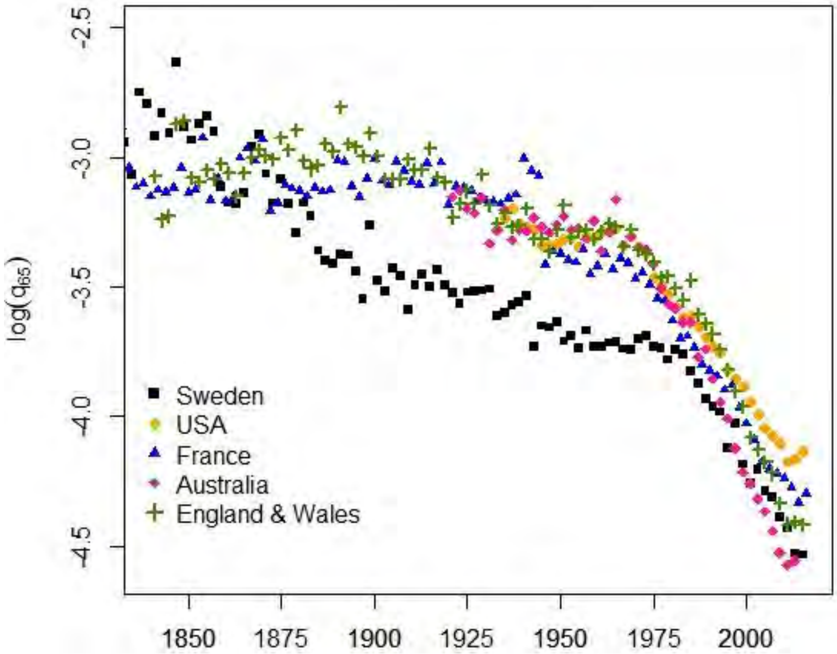

Figure 1: Logarithm of probabilities of deaths for 65-year old males in selected countries

For this reason several authers have proposed trend processes which explicitly allow for trend changes. Hainaut (2012) uses a random walk with drift where the parameters of the random walk (and its drift in particular) can switch between different regimes. Hunt and Blake (2015) allow for a continuous range of future mortality trends by simulating drift changes by a Pareto distribution. Other authors like Sweeting (2011) and Börger and Schupp (2018) have skipped the random walk concept entirely and have instead proposed trend stationary processes with piecewise linear trends where the slope of the linear trend changes randomly over time. However, all these different approaches to stochastic modeling of trend changes have one thing in common: Data on trend changes and their magnitudes is sparse, and therefore, uncertainty in the estimation of the processes' parameters is substantial in general. Even for populations with rather long data histories, typically only a few historical trend changes can be observed. For populations with shorter data histories, e.g. starting after World War II, reliable parameter estimations are often impossible. This clearly limits the applicability of trend process with random trend changes so far. This paper addresses the issue of parameter estimation for these processes and makes propositions how reasonable calibrations can be obtained also for populations with short data histories.

A common concept for reducing parameter uncertainty is to enlarge the data base for parameter estimation by aggregating data from several populations. In this paper, we analyze how this idea can be applied in order to improve the parameter estimation of trend change processes. Exemplarily, we do this for the trend process proposed by Börger and Schupp (2018) and further refined in Schupp (2019). This allows us to illustrate different data aggregation approaches which could easily be applied to other trend change processes, too.

The remainder of this paper is structured as follows: In the following section, we introduce the trend process of Börger and Schupp (2018) and its application within the CBD mortality model of Cairns et al. (2006). We briefly summarize the parameter estimation in a single population setting and discuss the issue of parameter uncertainty. Finally, we provide a concrete example of the trend process for the population of US males. In Section 3, we compare parameter estimates as well as the respective uncertainties for different populations. Possible approaches to improving parameter estimates and reducing the associated uncertainty are then presented in Section 4. The theoretical discussion of these approaches is complemented by a numerical example in Section 5. Finally, Section 6 concludes.

2. Trend Change Mortality Process

2.1 Specification of Trend Change Process

Börger and Schupp (2018) and Schupp (2019) apply their trend change process to the time dependent parameters in the well-known CBD mortality model of Cairns et al. (2006). In the CBD model, annual probabilities of death are described as

$$logit(q_{x,t}) \coloneqq \log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)} \cdot (x - \bar{x}),$$

where $\bar{x}$ is the average age of the age range under consideration. The time dependent parameter $\kappa_t^{(1)}$ determines the general level of mortality, whereas the slope parameter $\kappa_t^{(2)}$ describes the increase of mortality with age.

Building on the observations from Figure 1, Börger and Schupp (2018) and Schupp (2019) propose a trend process which projects piecewise linear trends with random changes in slope. For any future year t, the "observable" processes $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are modeled as the sums of underlying but unobservable true mortality processes and random noise terms:

$$\kappa_t^{(i)} = \hat{\kappa}_t^{(i)} + \varepsilon_t^{(i)},\ i = 1,2.$$

The noise terms $\varepsilon_t^{(i)}$ account for annual fluctuations which are, e.g., due to flu waves, very hot summers, or catastrophes. The vector $\varepsilon_t = \left(\varepsilon_t^{(1)}, \varepsilon_t^{(2)}\right)'$ is assumed to follow a two-dimensional Normal distribution with mean zero and covariance matrix Σ.[1]

The underlying mortality processes $\hat{\kappa}_t^{(i)},\ i = 1,2$ are projected linearly with current slopes $\hat{d}_t^{(i)},\ i = 1,2$:

$$\hat{\kappa}_t^{(i)} = \hat{\kappa}_{t-1}^{(i)} + \hat{d}_t^{(i)},\ i = 1,2.$$

The slopes remain unchanged until the next trend change occurs:

[1] Note that the covariance matrix is assumed constant over time even though heteroscedasticity can usually be observed in the historical data (see Figure 1). However, since the noise does not impact long-term mortality evolutions (in contrast to the innovations in the random walk), this simplification appears appropriate. In the parameter estimation, heteroscedasticity is accounted for though.

$$\hat{d}_t^{(i)} = \begin{cases} \hat{d}_{t-1}^{(i)} & , \; if \; no \; trend \; change \; occurred \; in \; t-1 \\ \hat{d}_{t-1}^{(i)} + \lambda_{t-1}^{(i)} & , \; if \; a \; trend \; change \; by \; \lambda_{t-1}^{(i)} \; occurred \; in \; t-1 \end{cases}.$$

The trend change intensities $\lambda_t^{(i)}$, $i = 1, 2$, are derived as the product of their absolute magnitudes $M_t^{(i)}$ and their signs $S_t^{(i)}$:

$$\lambda_t^{(i)} = S_t^{(i)} \cdot M_t^{(i)}, i = 1, 2.$$

Based on analyses of historical trend changes, Börger and Schupp (2018) propose modeling the magnitudes $M_t^{(i)}$ by Lognormal distributions $LN(\mu^{(i)}, \sigma^{(i)})$. For the signs $S_t^{(i)}$, they use Bernoulli distributions with attainable values -1 and 1 and probability 0.5 each. The probabilities of observing trend changes in $\hat{\kappa}_t^{(i)}$ in any particular year are denoted by $p^{(i)}, i = 1, 2$. Moreover, trend changes in $\hat{d}_t^{(1)}$ and $\hat{d}_t^{(2)}$ are assumed to occur independently as indicated by the occurrences of trend changes in the historical data for a large set of populations (see Börger and Schupp (2018)).

The decomposition of the trend change intensity into absolute magnitude and sign has several convenient implications. First, the distribution of future trend changes is symmetric, i.e. the slope increases and decreases with equal probability and magnitude. Thus, the prevailing trend (even though unobservable)[2] is always the best estimate for the trend at any future point in time. Furthermore, the distribution of $\lambda_t^{(i)}$ has no probability mass at zero and only very little mass around zero. Thus, simulated trend changes can be considered as rather "significant". At the same time, the heavy tail of the Lognormal distribution implies that strong trend changes can occur which is in line with some of the trend changes we observe in Figure 1 (e.g. for Sweden around 1980).

2.2 Parameter Estimation and Uncertainty

In this subsection we explain how the parameters of the trend processes $\kappa_t^{(i)}$ can be estimated from data for an individual population, and we particularly discuss the uncertainty involved. We assume that historical realizations $\kappa_t^{(i)}, t \leq t_0$ are given, where t_0 denotes the final year of the historical data, i.e. the starting point of a simulation. The parameters to be estimated are:

- the probabilities of observing a trend change in a certain year, $p^{(i)}$
- the parameters of the Lognormal distributions for the trend change magnitudes, $\mu^{(i)}$ and $\sigma^{(i)}$,
- the starting values for the underlying but unobservable trend processes $\hat{\kappa}_{t_0}^{(i)}$,
- the prevailing slopes of these trend processes, $\hat{d}_{t_0}^{(i)}$,
- and the covariance matrix Σ of the two-dimensional noise vector ε_t.

[2] See Börger et al. (2018) for a thorough discussion on the observability of mortality data and trends as well as implications for mortality modeling in practical applications.

The parameter estimation is carried out separately for each $\kappa_t^{(i)}$ process, and the covariance matrix Σ is estimated in the very last step. As Schupp (2019) explains, parameter estimation is complex due to the common dependence of realized $\kappa_t^{(i)}, t \leq t_0$ on potential but unknown trend changes in previous years. In particular, a full maximum likelihood estimation of all model parameters seems impossible from a practical point of view. Therefore, we rely on the pseudo maximum likelihood approach proposed by Schupp (2019). An iterative algorithm determines, for any fixed number of trend changes k, (a) the specific realization of the underlying trend process with k trend changes, $\hat{\kappa}_{t,k}^{(i)}, t \leq t_0$, which is closest to the actual data $\kappa_t^{(i)}, t \leq t_0$ in terms of likelihood, and (b) parameter values as (pseudo) maximum likelihood estimates which are consistent with this trend process realization. More precisely, starting with some initial parameter values, the best possible trend process realization $\hat{\kappa}_{t,k}^{(i)}, t \leq t_0$ is determined. Then the parameter values are updated based on this realization and the contained trend changes in particular. Next, these updated parameter values are applied in an update of $\hat{\kappa}_{t,k}^{(i)}, t \leq t_0$. This iterative algorithm typically converges after only very few steps, and we refer to Schupp (2019) for more details.

From the sets of (pseudo) maximum likelihood estimates $p_k^{(i)}$, $\mu_k^{(i)}$, $\sigma_k^{(i)}$, $\hat{\kappa}_{t_0,k}^{(i)}$, $\hat{d}_{t_0,k}^{(i)}$, and $\Sigma_{(i,i),k}$ for different numbers of trend changes k, the final parameter estimates and their respective uncertainties can be determined as follows (a numerical example is provided in the next subsection). The trend change parameters $\theta^{(i)} = \left(p^{(i)}, \mu^{(i)}, \sigma^{(i)}\right)$ are computed as weighted averages of the estimates $\theta_k^{(i)} = \left(p_k^{(i)}, \mu_k^{(i)}, \sigma_k^{(i)}\right)$:

$$\theta^{(i)} = \sum_k w_k^{(i)} \cdot \theta_k^{(i)}, \tag{1}$$

where the weights $w_k^{(i)}$ sum up to one and are based on a relative likelihood measure similar to Bayesian weights. Thus, most weight is applied to the $\theta_k^{(i)}$ for which the respective best possible trend process realization fits the actual data best; for more details on these weights, we refer to Schupp (2019). Even though k can range from zero to the number of data points in theory, in practice only a few values for k need to be considered since most weights $w_k^{(i)}$ are effectively zero.

In estimating the trend change parameters, uncertainty mainly arises from two sources: First, the actual number of historical trend changes cannot be determined exactly as the random noise affects the search for the best possible trend process realizations $\hat{\kappa}_{t,k}^{(i)}, t \leq t_0$. In other words, the trend process realizations which are found for different numbers of trend changes k may fit the actual data similarly well in terms of likelihood. This issue is already taken into account above when deriving the central parameter estimates as weighted averages. Nevertheless, the uncertainty around these central parameter estimates must not be neglected. The second source of uncertainty lies in the estimation of the trend change parameters from a limited (and typically small) number of trend changes k, even if we assume to know the exact number.

In order to quantify the overall parameter uncertainty, we search for a combined (approximate) standard error for both sources of uncertainty. With respect to the second source of uncertainty, the (pseudo) maximum likelihood estimation provides covariance matrices of (approximate) standard errors for each value of k which we denote by $SE_k^{(i)}$. Moreover, we assume that this component of parameter uncertainty in $\theta_k^{(i)}$ can be expressed by some distribution $F_k^{(i)}$. Concerning the unclear number of actual trend changes, the weights $w_k^{(i)}$ provide probabilities for each possible number of trend changes k. Therefore, we assume that the overall parameter uncertainty in $\theta^{(i)}$ is described by the distribution $F^{(i)} = \sum_k w_k^{(i)} \cdot F_k^{(i)}$. Then it can be shown that the covariance matrix of overall standard errors is given by

$$SE^{(i)} = \sum_k w_k^{(i)} \cdot \left(SE_k + \left(\theta_k^{(i)} - \theta^{(i)}\right) \cdot \left(\theta_k^{(i)} - \theta^{(i)}\right)'\right). \qquad (2)$$

To summarize the estimation of the trend change parameters, we have central estimates according to Equation (1) and a covariance matrix of standard errors according to Equation (2). Thus, when projecting the future mortality evolution, parameter uncertainty can be taken into account by randomly drawing, for each simulation path, parameter values from a suitable distribution with according mean vector and covariance matrix.

For the estimation of the starting values of a simulation, $\hat{\kappa}_{t_o}^{(i)}$ and $\hat{d}_{t_0}^{(i)}$, and the associated uncertainty, we follow a different approach. Parameter uncertainty here mainly arises from the uncertainty when the most recent trend change has occurred. When determining the best possible trend process realizations $\hat{\kappa}_{t,k}^{(i)}, t \leq t_0$, a trend change in recent years may be detected for some values of k but not for others. In order to confirm the (non-)occurrence of such a trend change, a couple of additional years of data would be required. As long as this data is not available, however, uncertainty with respect to the starting values may be substantial, depending on the magnitude of the potential trend change.[3] In a simulation of future mortality, obviously both cases, i.e. with and without the potential trend change, should be taken into account. However, since the potential trend change either has occurred or not, we cannot specify central parameter estimates. Instead, we have different estimates for $\hat{\kappa}_{t_o,k}^{(i)}$ and $\hat{d}_{t_0,k}^{(i)}$ with different probabilities/weights $w_k^{(i)}$, and for each simulation path, starting values should be drawn randomly from this empirical distribution. A numerical example is provided in the next subsection.

A central estimate for the covariance matrix Σ of the noise vector ε_t can be derived analogously to Equation (1). If we denote by $\Sigma_{k,m}$ the sample covariance matrix for the case of k trend changes in $\kappa_{t,t\leq t_0}^{(1)}$ and m trend changes in $\kappa_{t,t\leq t_0}^{(2)}$, the final estimate is given as

$$\Sigma = \sum_k \sum_m w_k^{(1)} \cdot w_m^{(2)} \cdot \Sigma_{k,m}.$$

[3] In comparison, when the most recent trend change is assumed to be known, i.e. for a fixed k, the uncertainty in regressing the starting values from the available data appears negligible.

Compared to the uncertainty in the trend change parameters, the uncertainty in Σ is negligible as it is estimated from rather large samples of residuals. Furthermore, the impact of the noise vector ε_t in projections of future mortality is very limited. Therefore, parameter uncertainty can be ignored here.

2.3 Example Calibration

In this subsection we present a full model calibration for US males including specifications of the parameter uncertainties involved. We use the entire data set which is available in the HMD (2019) for ages 60 to 109, i.e. from 1933 to 2016.

Figure 2 shows the historic trend processes $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ and the best possible realizations for the underlying trend processes, $\hat{\kappa}_{t,k}^{(1)}$ and $\hat{\kappa}_{t,k}^{(2)}$, for the relevant numbers of trend changes k, i.e. for those k with $w_k^{(i)} > 0$. Table 1 and Table 2 provide the corresponding parameter estimates and weights/probabilities. We find that the number of actual trend changes is very likely to lie between 2 and 5 or 6, respectively. While some parameter values are very similar for different k, e.g. the $\mu_k^{(i)}$, others vary substantially. Unsurprisingly, this particularly holds for the trend change probabilities $p_k^{(i)}$. This observation clearly underlines why it is important to account for the fact that one cannot clearly observe the exact number of trend changes.

By applying Equation (1), the central parameter estimates can then be derived:

$$\theta^{(1)} = \left(p^{(1)}, \mu^{(1)}, \sigma^{(1)}\right) = (0.0451, -4.4867, 0.2290)$$
$$\theta^{(2)} = \left(p^{(2)}, \mu^{(2)}, \sigma^{(2)}\right) = (0.0480, -7.3466, 0.3932).$$

Furthermore, from Equation (2) we obtain the following covariance matrices of standard errors:

$$SE^{(1)} = \begin{pmatrix} 6.906 \cdot 10^{-4} & -1.245 \cdot 10^{-3} & 2.812 \cdot 10^{-4} \\ -1.245 \cdot 10^{-3} & 2.695 \cdot 10^{-2} & 2.047 \cdot 10^{-3} \\ 2.812 \cdot 10^{-4} & 2.047 \cdot 10^{-3} & 8.691 \cdot 10^{-3} \end{pmatrix},$$

$$SE^{(2)} = \begin{pmatrix} 5.929 \cdot 10^{-4} & -2.127 \cdot 10^{-4} & -1.108 \cdot 10^{-4} \\ -2.127 \cdot 10^{-4} & 4.469 \cdot 10^{-2} & -1.534 \cdot 10^{-3} \\ -1.108 \cdot 10^{-4} & -1.534 \cdot 10^{-3} & 2.396 \cdot 10^{-2} \end{pmatrix}.$$

Comparing the (one-dimensional) standard errors, i.e. the roots of the diagonal entries of the $\mathrm{SE}^{(\mathrm{i})}$, with the central parameter estimates, we find coefficients of variation between 40% and 60% for the $\mathrm{p}^{(\mathrm{i})}$ and $\sigma^{(\mathrm{i})}$. Thus, parameter uncertainty is huge for these parameters. For the $\mu^{(\mathrm{i})}$, on the other hand, the coefficients of variation are only around -3%.

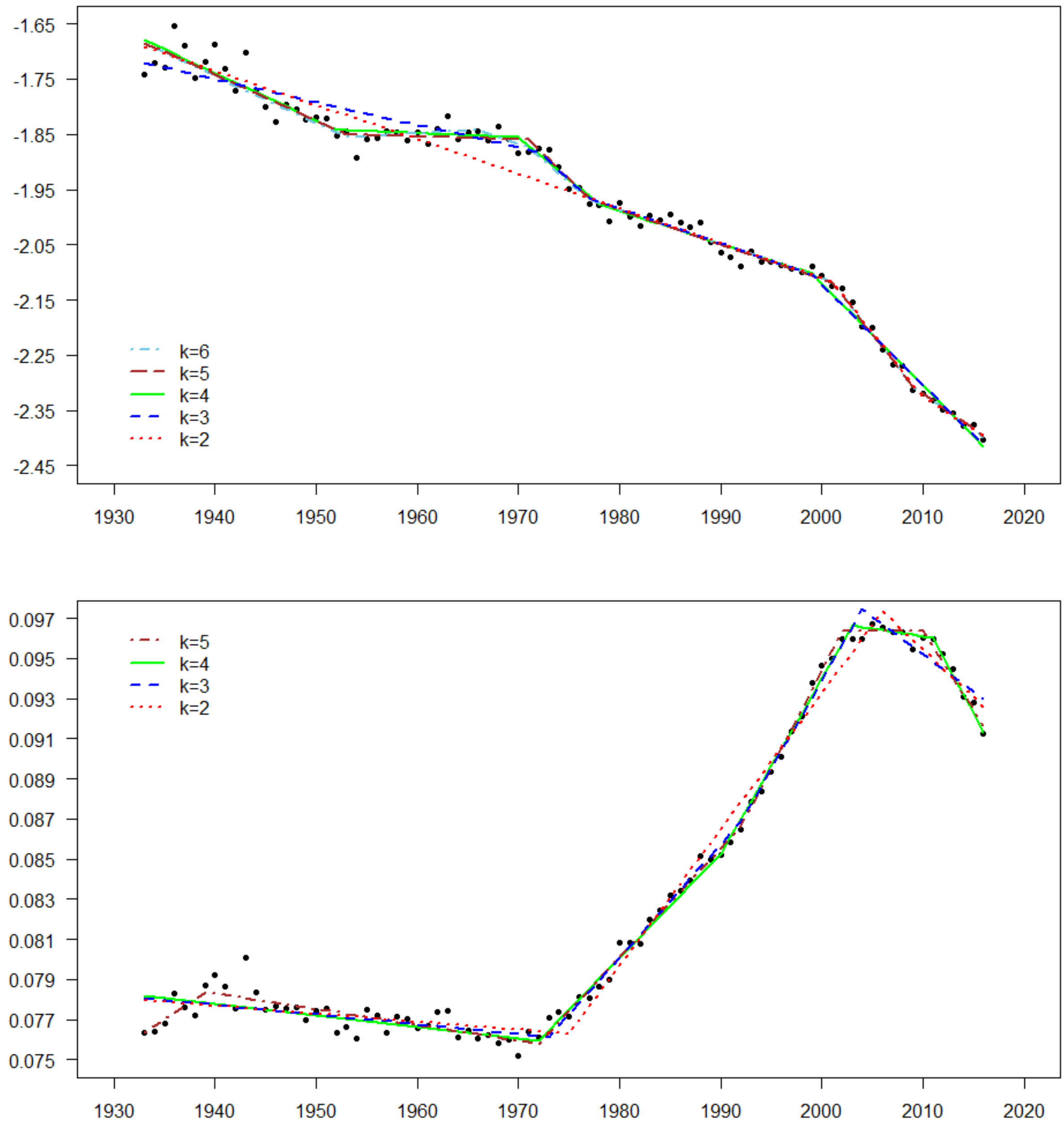

Figure 2: Historical trend processes $\kappa_t^{(1)}$ (top) and $\kappa_t^{(2)}$ (bottom) for US males and best possible realizations for the underlying trend processes, $\hat{\kappa}_{t,k}^{(1)}$ and $\hat{\kappa}_{t,k}^{(2)}$

	$p_k^{(1)}$	$\mu_k^{(1)}$	$\sigma_k^{(1)}$	$\hat{\kappa}_{t_0,k}^{(1)}$	$\hat{d}_{t_0,k}^{(1)}$	$w_k^{(1)}$
k = 2	0.0244	−4.2823	0.1993	−2.3957	−0.0118	0.214
k = 3	0.0366	−4.4642	0.0814	−2.4133	−0.0182	0.034
k = 4	0.0488	−4.5431	0.2414	−2.4148	−0.0184	0.668
k = 5	0.0610	−4.3970	0.3033	−2.3959	−0.0128	0.008
k = 6	0.0732	−4.5850	0.2607	−2.3981	−0.0132	0.076

Table 1: Estimates for trend change parameters $p_k^{(1)}$, $\mu_k^{(1)}$, $\sigma_k^{(1)}$ and starting values $\hat{\kappa}_{t_o,k}^{(1)}$, $\hat{d}_{t_0,k}^{(1)}$ for different numbers of trend changes k and corresponding best possible realizations $\hat{\kappa}_{t,k}^{(1)}$ for the underlying trend process

	$p_k^{(2)}$	$\mu_k^{(2)}$	$\sigma_k^{(2)}$	$\hat{\kappa}_{t_0,k}^{(2)}$	$\hat{d}_{t_0,k}^{(2)}$	$w_k^{(2)}$
$\mathbf{k = 2}$	0.0244	-7.0009	0.2368	$9.259 \cdot 10^{-2}$	$-4.753 \cdot 10^{-4}$	0.024
$\mathbf{k = 3}$	0.0366	-7.3783	0.5627	$9.259 \cdot 10^{-2}$	$-3.739 \cdot 10^{-4}$	0.079
$\mathbf{k = 4}$	0.0488	-7.3497	0.3858	$9.132 \cdot 10^{-2}$	$-9.440 \cdot 10^{-4}$	0.839
$\mathbf{k = 5}$	0.0610	-7.3972	0.3294	$9.166 \cdot 10^{-2}$	$-7.885 \cdot 10^{-4}$	0.057

Table 2: Estimates for trend change parameters $p_k^{(2)}$, $\mu_k^{(2)}$, $\sigma_k^{(2)}$ and starting values $\hat{\kappa}_{t_o,k}^{(2)}$, $\hat{d}_{t_0,k}^{(2)}$ for different numbers of trend changes k and corresponding best possible realizations $\hat{\kappa}_{t,k}^{(2)}$ for the underlying trend process

The figure and tables also show substantial parameter uncertainty for the starting values $\hat{\kappa}_{t_o}^{(i)}$ and $\hat{d}_{t_0}^{(i)}$. For $\kappa_t^{(1)}$ and $k = 3, 4$, the most recent trend change is detected in 1999; the probability for this being the most recent actual trend change is about 70%. However, there is also a 30% chance that the most recent trend change occurred in fact in 2009 or 2010 as detected for $k = 2, 5, 6$.[4] Similarly, we find the most recent trend change in $\kappa_t^{(2)}$ in 2004 or 2006 (for $k = 2, 3$ and with probability of about 10%) or in 2010 or 2011 (for $k = 4, 5$ and with probability of about 90%). In both cases, the $\hat{d}_{t_0,k}^{(i)}$ differ with k in particular, and this uncertainty should be taken into account in projections of future mortality. In contrast to the example at hand, the most recent trend change may be very clear for other populations and/or at other points in time. Thus, parameter uncertainty in the starting values is highly case specific and may even be negligible in some cases.

Also the question whether parameter uncertainty in the starting values can be reduced by combining insights from several populations can only be answered individually for each specific case. Possibly, a potential trend change for one population can be confirmed by detecting similar trend changes for other populations with likely the same reason of occurrence. However, trend changes may also be specific to a single population such that insights from other populations can even be misleading. Due to this need for a case specific analysis, we will not discuss this question further in this paper.

For completeness, the estimate for the covariance matrix of the noise vector is

$$\Sigma = \begin{pmatrix} 1.773 \cdot 10^{-4} & 3.401 \cdot 10^{-6} \\ 3.401 \cdot 10^{-6} & 2.092 \cdot 10^{-7} \end{pmatrix}.$$

In order to illustrate the issue of parameter uncertainty, Figure 3 shows projections of remaining period life expectancies of 65-year old US males with and without parameter uncertainty. Parameter uncertainty in the starting values is taken into account by randomly drawing from the empirical distribution of starting values as given in Table 1 and Table 2 for each simulation

[4] Due to the noise it is obviously difficult to exactly date a trend change. Therefore, we assume that the same trend change may be detected in subsequent years for different k.

path. For the case without parameter uncertainty, the starting values with the largest probability $w_k^{(i)}$ are considered. The path dependent trend change parameters $p^{(i)}$, $\mu^{(i)}$, and $\sigma^{(i)}$ are drawn from correlated (one-dimensional) Beta, Normal, and Gamma distributions, respectively.[5] The Beta and Gamma distributions are chosen in order to ensure that the trend change probabilities always lies between zero and one and that the standard deviations of the trend change magnitudes are always positive. Slight inconsistencies with the standard errors being derived under the assumption of Normality are accepted.

We find that the mean projection changes significantly in the case where parameter uncertainty is taken into account. The reason is that, for both $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$, the most recent trend change could not be clearly determined and that different starting slopes are assigned significant probabilities in particular. The prediction intervals widen only slightly when allowing for parameter uncertainty. This is somewhat surprising at first sight given the substantial standard errors above. However, the parameters $p^{(1)}$ and $\mu^{(1)}$ are negatively correlated for US males which means that in case the likelihood of trend changes is higher (larger $p^{(1)}$), their magnitude is likely to be smaller (smaller $\mu^{(1)}$). For other populations, however, we have found positive correlation between these parameters leading to significantly wider prediction intervals when allowing for parameter uncertainty. This illustrates once again that it is reasonable to combine observations from different populations. The case with (population specific) parameter uncertainty will also be the benchmark in the numerical example later on.

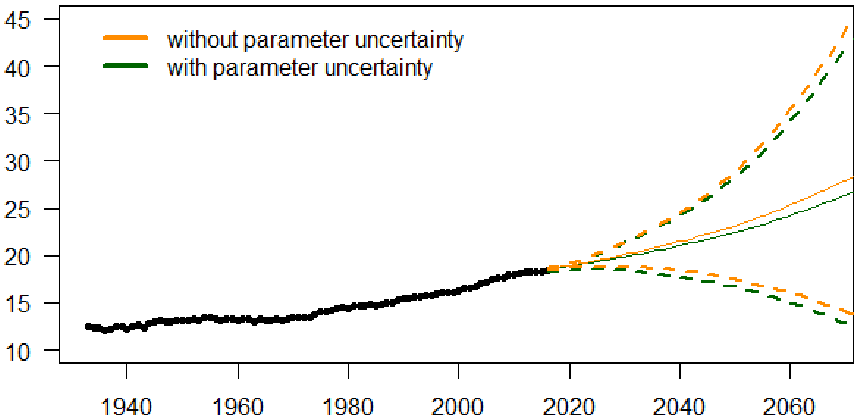

Figure 3: Mean projections (solid lines) and 90% prediction intervals (dashed lines) of remaining period life expectancies for 65-year old US males with and without parameter uncertainty

---

5 More precisely, for each simulation path, a three-dimensional Normal vector with mean equal to the central parameter estimates and covariance matrix $SE^{(i)}$ is generated, and the first and third component of the Normal vector are then transformed to Beta and Gamma distributed values with unchanged mean and variance.

3. Comparison of Population Specific Trend Process Calibrations

We commence our multi-population analysis with a comparison of population specific estimates of the trend change parameters $\theta^{(i)} = \left(p^{(i)}, \mu^{(i)}, \sigma^{(i)}\right),\ i = 1, 2$ and their associated uncertainty. To this end, we calibrate the CBD model and subsequently the trend process from Section 2 to the male and female populations from the following countries: Australia, Austria, Canada, Denmark, England & Wales, Finland, France, Italy, Japan, the Netherlands, New Zealand (non-Maori), Norway, Sweden, Switzerland, the United States, and West Germany.[6] For each population we use the entire HMD data history for ages 60 to 109, as long as it is not explicitly marked as unreliable in the HMD (as e.g. for Sweden before 1860).

Figure 4 shows, for each of the 32 populations, the central estimates and 95% confidence intervals for the trend change parameters. The confidence intervals are derived from Beta, Normal, and Gamma distributions, respectively, which have been calibrated to the population specific central parameter estimates and covariance matrices of standard errors (as explained for US males in Subsection 2.3). We find that uncertainty is substantial for most parameters and populations. This particularly holds for the trend change probabilities. In fact, the central parameter estimates lie within the confidence intervals for most other populations. In comparison, uncertainty in $\mu^{(1)}$ and $\mu^{(2)}$ is smaller which is in line with our findings for US males in Subsection 2.3. Nevertheless, the central parameter estimates are rather similar between the different populations, in particular for $\mu^{(1)}$. The parameter uncertainty in $\sigma^{(1)}$ and $\sigma^{(2)}$ is exceptionally large for a few populations, but again the central parameter estimates lie within the confidence intervals for most other populations.

Given the similarities between parameter estimates for many populations and given the substantial parameter uncertainties, there is reason to believe that parameter calibrations can be improved by aggregating data from different populations. Thus, the question arises which populations should be considered, i.e. which populations can provide insights for the population one is particularly interested in. When modeling multi-population mortality, typically populations with close economic and political links are taken into account, which have thus experienced similar historical mortality evolutions. However, this may not always be a suitable approach when calibrating trend change processes. Populations which are very closely linked may have essentially experienced the same trend changes. Thus, in order to substantially enrich the data base on trend changes, also populations with weaker links and different historical trend process patterns should be considered. In that case, some assumption needs to be made with respect to the similarity (or even equality) of the underlying trend change parameters. We will address this issue in the following section.

[6] Other countries for which data is available in the HMD have been omitted for different reasons: The populations are so small that the noise is too strong to detect trend changes (e.g. Iceland), reliable data is only available for a few decades (e.g. Portugal), or data is missing for some years (e.g. Belgium).

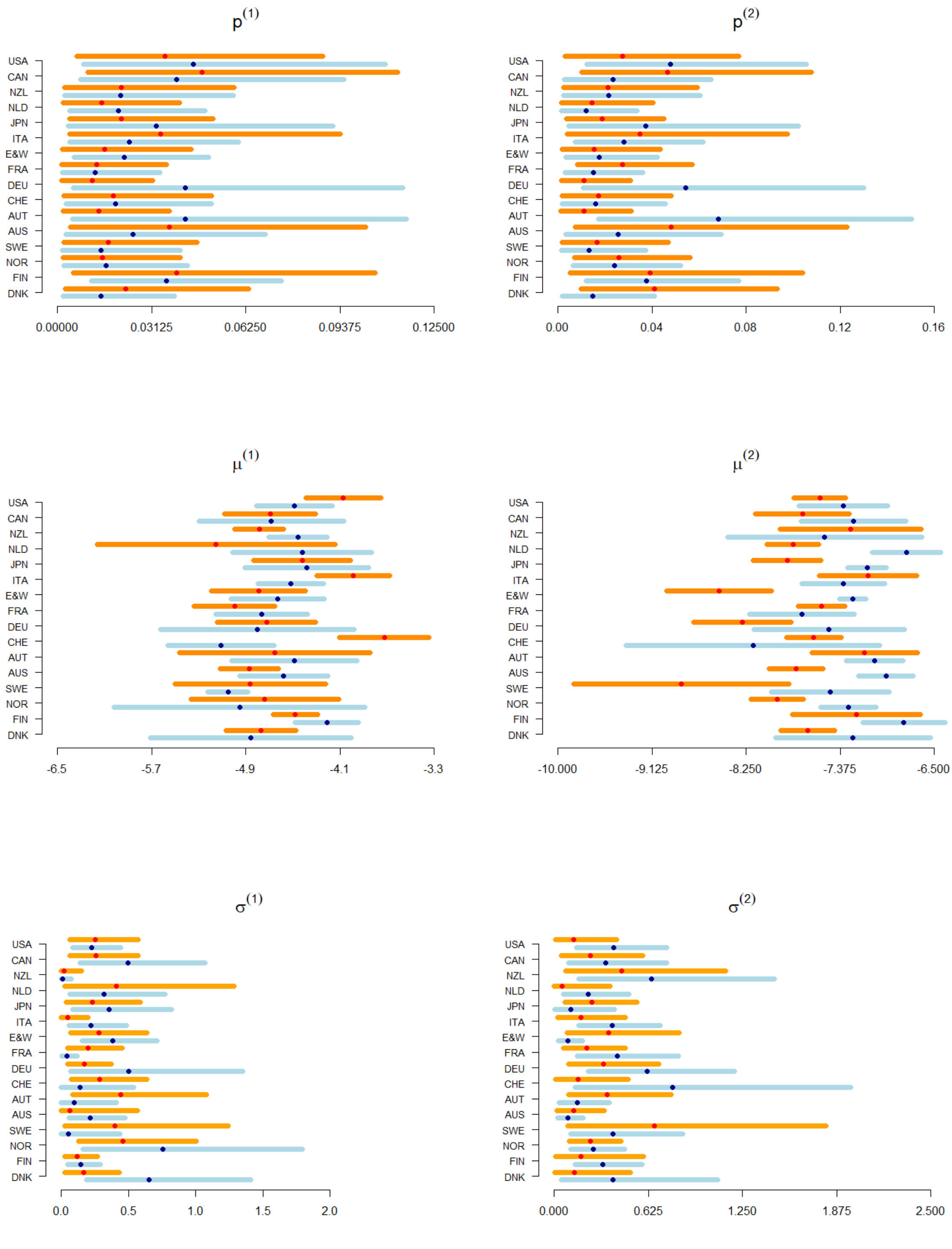

Figure 4: Central estimates and 95% confidence intervals for the parameters $p^{(1)}$, $\mu^{(1)}$, $\sigma^{(1)}$ (left) and $p^{(2)}$, $\mu^{(2)}$, $\sigma^{(2)}$ (right) for males (blue) and females (orange).

4. Trend Process Calibrations to Multi-Population Data

In this section, we will explain different approaches how data from multiple populations can be combined to obtain calibrations for the trend change parameters $\theta^{(i)} = \left(p^{(i)}, \mu^{(i)}, \sigma^{(i)}\right)$. This can be particularly helpful for populations with insufficient data for an individual trend process estimation and also to reduce parameter uncertainty for populations with longer data histories. These data aggregation approaches can be categorized according to whether the trend change parameters for different populations are assumed to be equal or only to come from the same distribution. The assumption of equal parameters for all populations can be motivated by the observations from Figure 4. Parameter uncertainty is huge for most populations and parameters, and it cannot be ruled out that the trend change parameters are equal for all populations. Nevertheless, one may want to relax this assumption, in particular if substantial data is available for the population one is particularly interested in and this data indicates that the population's parameters may be different from those of other populations.

For the remainder of this paper, we assume that data is available for a set of populations P, and the index $\cdot_p$ denotes specific parameter estimates etc. for population $p \in P$. Moreover, let N_p denote the number of data points for population p. Finally, let $p^* \in P$ be the population whose future mortality evolution one wants to project. A numerical example for the proposed data aggregation approaches is provided in the next section.

Parameter estimation from entire data set

Assuming equal parameters for all populations, the most consistent approach certainly is to estimate the parameters from data for all populations simultaneously. This would stabilize the parameter estimation and should reduce parameter uncertainty substantially. However, even though this is a desirable approach from a theoretical perspective, it can be difficult to implement for a large number of populations in practice. For the estimation algorithm proposed in Subsection 2.2, we have found this to be hardly feasible.

Parameter estimation from observed trend changes

Alternatively, the trend change parameters can be estimated from the historical trend changes for all populations. These trend changes, more precisely their occurrences and magnitudes, are determined for each population individually. This obviously implies a slight distributional inconsistency as the historical trend changes are not assumed to be generated by the same set of trend change parameters, but it makes the approach practically feasible. Given the historical trend changes, the trend change parameters are then estimated e.g. via a maximum likelihood approach. Denoting by $\left|\lambda^{(i)}\right|_{p,k}$ the vector of absolute trend change magnitudes for the case of k trend changes for population p, the following likelihood function should be maximized:

$$L(\theta^{(i)}) = \prod_{p\in P}\prod_{k=0}^{N_p}\left(L_{lognormal}\left(\left|\lambda^{(i)}\right|_{p,k};\ \mu^{(i)},\sigma^{(i)}\right)\cdot L_{bernoulli}\left(k;\ p^{(i)}\right)\right)^{w_{p,k}^{(i)}}$$

$$= \prod_{p\in P}\prod_{k=0}^{N_p}\left(\prod_{j=1}^{k} f_{lognormal}\left(\left(\left|\lambda^{(i)}\right|_{p,k}\right)_j;\ \mu^{(i)},\sigma^{(i)}\right)\cdot p^{(i)k}\cdot\left(1-p^{(i)}\right)^{N_p-k}\right)^{w_{p,k}^{(i)}},$$

where $f_{lognormal}$ denotes the probability density function of the lognormal distribution, and the weights $w_{p,k}^{(i)}$ account for the "relevance" of the different trend change realizations. The maximum likelihood estimation also provides a covariance matrix of (approximate) standard errors as a representation of the remaining parameter uncertainty.

Weighted average of population specific parameter estimates

Instead of equal parameters for all populations, we now assume that the parameter values for each population only come from the same distribution of possible parameter values. In this case, the simplest approach to obtain a set of parameter values for population p^* is to take the average of all population specific parameter estimates $\theta_p^{(i)} = \left(p_p^{(i)}, \mu_p^{(i)}, \sigma_p^{(i)}\right)$. This approach is particularly applicable if hardly any information is available on the true trend change parameters for population p^*. A weighted average can be applied in order to account for the relevancy of each population or the credibility of its parameter estimates. A larger weight would then be assigned to a population if, e.g., it is expected to be very informative for population p^* or it has a comparably long data history. Denoting by v_p the weight for each population, the common parameter estimates would be

$$\theta^{(i)} = \sum_{p\in P} v_p\cdot\theta_p^{(i)},$$

with weights e.g. according to the data history, $v_p = N_p/\sum_{q\in P} N_q$. The uncertainty associated with these parameter estimates can be determined as the (weighted) sample covariance matrix of the population specific parameter sets $\theta_p^{(i)}$. For a simulation of future mortality, the same approach as in Subsection 2.3 can then be applied.

Parameter sampling from empirical distribution

Alternatively, a (three-dimensional) empirical distribution can be derived from the population specific parameter estimates $\theta_p^{(i)}$, and parameter values can be drawn randomly from this distribution for each simulation path. In analogy to the weighted average above, the different parameter sets in the empirical distribution can be assigned different probabilities. We denote this empirical distribution by $F_{\theta^{(i)}}$ with $P(\theta_p^{(i)}) = v_p$ for $p \in P$ and zero otherwise. The outcomes of this approach should be very similar to those for the weighted average approach. The means of the randomly drawn parameter values are the same by construction, and also the simulated uncertainty in the parameter values should be comparable if the set of populations is large. In that case, the empirical distribution should be reasonably similar to the (theoretical) distribution parameters are drawn from in the weighted average approach.

Credibility approach

In the previous two approaches we assumed that hardly any information is available on the true trend change parameters for population p^*. However, typically at least some information is available, e.g. in form of the $\theta_p^{(i)}$ for all populations $p \in P$. Hence, when simulating mortality for population p^*, the parameter estimates $\theta_{p^*}^{(i)}$ should be assigned a larger probability than parameter estimates which are significantly different. This can be achieved rather easily in a credibility approach where the probability v_{p^*} is increased to emphasize the population specific information. Thus, this approach represents a compromise between the uncertain estimates for the true population specific parameters and a larger reference group of parameter sets which may differ from the true population specific parameters.

Bayesian approach

Alternatively to the credibility approach, a Bayesian approach can be applied. We again assume that the parameter values for all populations come from the same, but unknown distribution. This is the prior distribution in the Bayesian setting, and we approximate it by the empirical distribution $F_{\theta^{(i)}}$. Without any further knowledge about population p^*, parameter values would be drawn from this prior distribution as in the sampling approach above. The realized $\kappa_{t,p^*}^{(i)}$ processes however provide additional information on likely parameter values $\theta_{p^*}^{(i)}$. Unfortunately, we cannot specify the likelihood $L\left(\kappa_{t,p^*}^{(i)}, t \leq t_0 \middle| \theta\right)$ for the realized $\kappa_{t,p^*}^{(i)}$ process being generated by some parameter set θ. Therefore, in line with the parameter estimation in Subsection 2.2, we instead consider likelihoods for best possible trend process realizations for different numbers of trend changes k which should be approximately proportional to $L\left(\kappa_{t,p^*}^{(i)}, t \leq t_0 \middle| \theta\right)$:

$$L\left(\kappa_{t,p^*}^{(i)}, t \leq t_0 \middle| \theta\right) = \sum_{k=0}^{N_p} L\left(\kappa_{t,p^*}^{(i)}, t \leq t_0 \middle| \theta, k\right) \propto \sum_{k=0}^{N_p} \hat{L}\left(\hat{\kappa}_{t,p^*}^{(i)}(\theta), t \leq t_0 \middle| \theta, k\right).$$

Here $L\left(\kappa_{t,p^*}^{(i)}, t \leq t_0 \middle| \theta, k\right)$ denotes the likelihood under the condition that the data has been generated with k trend changes and $\hat{L}\left(\hat{\kappa}_{t,p^*}^{(i)}(\theta), t \leq t_0 \middle| \theta, k\right)$ is the likelihood as defined in Schupp (2019). In order to avoid the computationally expensive iterative algorithm to obtain the latter likelihood for any parameter set θ, we approximate the best possible trend process realizations $\hat{\kappa}_{t,p^*}^{(i)}(\theta), t \le t_0$ by those from the individual parameter estimation, $\hat{\kappa}_{t,p^*}^{(i)}(\theta_{p^*}^{(i)}), t \le t_0$. Even though the latter have been determined under different parameter estimates, we can assume them to be reasonably similar to those for the parameter set θ since we have observed that the optimal positions of the k trend changes are the same for almost all reasonable parameter sets in general. Thus, we have

$$L\left(\kappa_{t,p^*}^{(i)}, t \le t_0 \middle| \theta\right) \propto \sum_{k=0}^{N_p} \hat{L}\left(\hat{\kappa}_{t,p^*}^{(i)}(\theta), t \le t_0 \middle| \theta, k\right) \approx \sum_{k=0}^{N_p} \hat{L}\left(\hat{\kappa}_{t,p^*}^{(i)}\left(\theta_{p^*}^{(i)}\right), t \le t_0 \middle| \theta, k\right) =: \hat{L}_{p^*}^{(i)}(\theta).$$

Finally, we can derive the posterior distribution $F_{\theta^{(i)}|\kappa_{t,p^*}^{(i)}}$, i.e. the probability for parameter set $\theta_p^{(i)}$ is

$$P\left(\theta_p^{(i)} \middle| \kappa_{t,p^*}^{(i)}, t \le t_0\right) = \frac{1}{c} \cdot v_p \cdot \hat{L}_{p^*}^{(i)}(\theta_p^{(i)}),$$

where c is a constant such that the probabilities sum up to one. This posterior distribution describes the remaining population specific parameter uncertainty, and in a simulation, path dependent parameter values should be drawn from this distribution.

5. Numerical Example

In this section we apply the different data aggregation approaches from Section 4 to the case of US males. Our set of reference populations P consists of the 32 populations which we already considered in Section 3 and the weights/probabilities v_p are derived according to the number of available data points. In the credibility approach we set $v_{p^*} = 0.5$ and reduce all other probabilities proportionally.

Figure 5 shows the central estimates of the trend change parameters and their 95% confidence intervals for the individual calibration for US males and the different data aggregation approaches. We see that the central parameter estimates can vary between the approaches and that parameter uncertainty is not necessarily reduced compared to the individual calibration. We will now explore why this is the case.

Starting with the maximum likelihood estimation based on all historical trend changes, we find that the central estimates of the trend change probabilities $p^{(i)}$ are considerably smaller when data is aggregated. A comparably large number of trend changes has been observed for US males and this is now compensated for. Parameter uncertainty reduces substantially for all three trend change parameters as expected. This means that the same trend change probability and Lognormal distribution may be assumed in order to generate trend changes for different populations. The comparably large central estimates for the $\sigma^{(i)}$ compensate for the fact that the mean of the trend change magnitudes is now derived from all trend changes and not only those of a particular population.

The results for the weighted averages and the sampling from the empirical parameter distribution are very similar as expected. The trend change probabilities $p^{(i)}$ are reduced again compared to the individual calibration, and also parameter uncertainty is smaller. In contrast, parameter uncertainty in the magnitude parameters $\mu^{(i)}$ and $\sigma^{(i)}$ has increased. However, this is primarily uncertainty which arises from the assumption of a distribution for population specific parameter values instead of the assumption of one fixed parameter set for all populations. Thus, it is systematic uncertainty as opposed to rather unsystematic uncertainty arising from parameter estimation from limited data. Only a small portion of the depicted parameter uncertainty can be credited to the randomness in the 32 population specific parameter estimates which are used to approximate the true but unknown distribution of parameter values.

Also the results from the credibility and the Bayesian approaches are rather similar. Both approaches build on the assumption of an unknown distribution for population specific parameters and determine some trade-off between the population specific parameter estimates and those from other populations. In the credibility approach, 50% probability is assigned to the parameter estimates for US males, while it is about 40% for $\theta_{p^*}^{(1)}$ and 11% for $\theta_{p^*}^{(2)}$ in the Bayesian approach. The remaining probability is assigned to other parameter sets, though based on different principles. Again we find reduced trend change probabilities, but less reduced than for the other aggregation approaches due to the substantial weight for the US male parameter set. The same applies to the magnitude parameters which also lie between the individual estimates and the parameter values from the other aggregation approaches. We also observe that, for all parameters, uncertainty remains substantial or even increases compared to the individual calibration. Again, this is primarily systematic uncertainty arising from the assumption of a distribution for population specific parameter values. Given the limited data on trend changes for US males, substantial uncertainty still remains with respect to the credibility of the population specific parameter estimates.

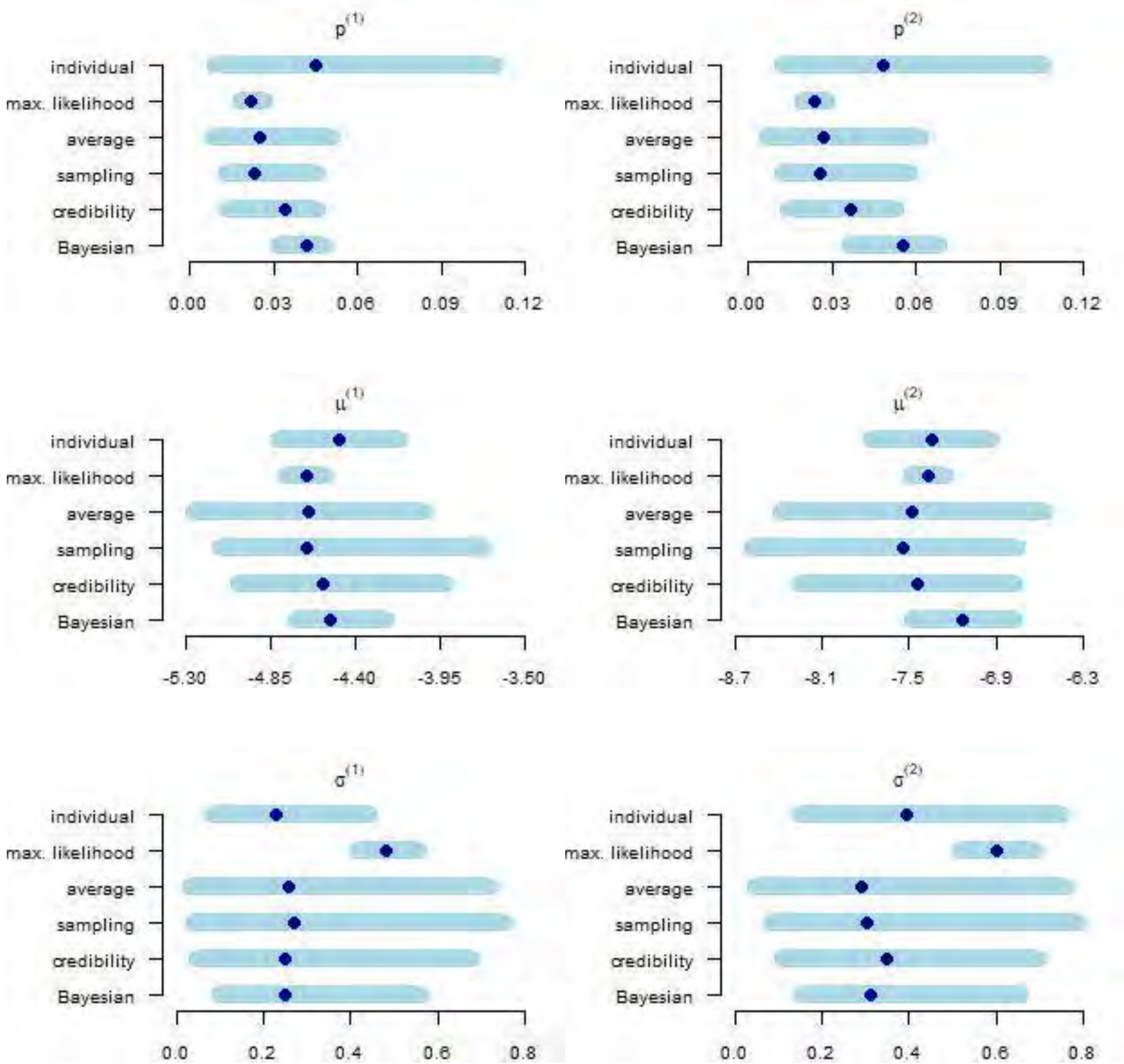

Figure 5: Central estimates and 95% confidence intervals for the parameters $p^{(1)}$, $\mu^{(1)}$, $\sigma^{(1)}$ (left) and $p^{(2)}$, $\mu^{(2)}$, $\sigma^{(2)}$ (right) for US males based on multi-population data.

Finally, we compare the different data aggregation approaches by projecting remaining period life expectancies for 65-year old US males. In each case, parameter uncertainty is accounted for by drawing from Beta, Normal, and Gamma distributions with case specific parametrizations. The starting values are modeled as explained in Subsection 2.3 in any case.

Figure 6 shows the 90% prediction intervals for all approaches considered. The widest prediction intervals can be observed for the population specific calibration. This is in line with the comparably large central parameter estimates for $p^{(i)}$ and $\mu^{(i)}$ which we observe in Figure 4. The maximum likelihood approach yields the most narrow prediction intervals, mainly due to comparably small central parameter estimates for $p^{(i)}$ and $\mu^{(i)}$ and the small parameter uncertainty for all parameters. Prediction intervals for the weighted parameter averaging and the parameter sampling are very similar which is again in line with observations from Figure 4. The widths of the prediction intervals are between those for the aforementioned approaches. Thus, uncertainty is particularly reduced compared to the individual calibration where the trend change probabilities may have been overestimated simply by chance. The credibility and the Bayesian approach yield prediction intervals between the individual calibration and the sampling approach as they define a kind of "mixture" of parameter estimates from these approaches.

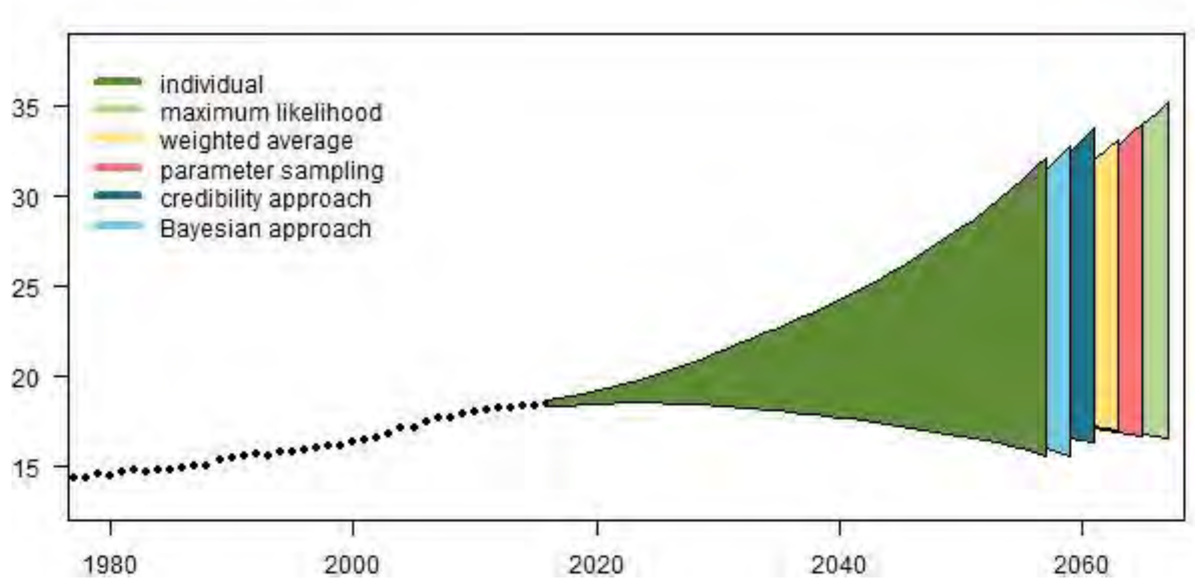

Figure 6: 90% prediction intervals for remaining period life expectancies of 65-year old US males based on different trend process calibrations

6. Conclusion

In this paper, we have discussed the issue of parameter uncertainty in mortality processes with trend changes. Due to the limited number of observed historical trend changes, parameter uncertainty is substantial in general. This particularly holds for those parameters which determine future trend changes as part of stochastic projections. We have identified the main sources of this uncertainty and have explained how it can be quantified for the trend process of Börger and Schupp (2018). A comparison of 32 populations shows that central parameter estimates can vary considerably when trend processes are calibrated for each population individually. However, due to the substantial uncertainty associated with these estimates, it is not clear whether this is mainly due to random effects in the few trend changes which have been observed for each population.

In order to improve the reliability of trend process calibrations, we have then discussed different approaches for aggregating data on trend changes from several populations. This includes approaches which assume the true parameter values to be equal for all populations under consideration as well as approaches where the parameter values for different populations are only assumed to come from the same, but unknown distribution. A maximum likelihood parameter estimation based on the historical trend changes for all populations shows that the assumption of equal trend change probabilities and magnitudes for all populations may be reasonable. Parameter uncertainty can be reduced substantially here.

When allowing for different parameter values from the same underlying distribution for each population, parameter uncertainty reduces only slightly or even increases compared to the population specific calibrations. Depending on how different the population specific parameter estimates are, the uncertainty in the empirical distribution build from these estimates can be larger than the uncertainty in estimating the population specific parameters in the first place. Nevertheless, central parameter estimates can change substantially when aggregating parameter estimates in a common distribution. As we have seen in a numerical example, this may prevent overestimation (underestimation) of the uncertainty in the future mortality evolution in case the population specific parameter estimates may have been rather large (small) simply by chance.

In conclusion, we have found that parameter uncertainty can be much better understood when data from different populations is aggregated. Furthermore, the reliability of trend process calibrations can be improved by reducing random effects in population specific parameter estimates. Whether this leads to more narrow or wider prediction intervals for the quantities of interest like future life expectancies, depends on several factors: most importantly, the population specific parameter estimates, the set of reference populations, and the assumption on how the parameter values for different populations relate to each other.

References

Börger, M., Fleischer, D., & Kuksin, N. (2014). Modeling the Mortality Trend under Modern Solvency Regimes. ASTIN Bulletin 44(1), 1-38.

Börger, M., Ruß, J., & Schupp, J. (2018). It Takes Two: Why Mortality Trend Modeling is more than modeling one Mortality Trend. Working Paper, Ulm University.

Börger, M., & Schupp, J: (2018). Modeling Trend Processes in Parametric Mortality Models. Insurance: Mathematics and Economics 78, 369-380.

Cairns, A., Blake, D., & Dowd, K. (2006). A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. Journal of Risk and Insurance 73(4), 687-718.

Hainaut, D. (2012). Multidimensional Lee-Carter model with Switching Mortality Processes. Insurance: Mathematics and Economics, 50(2): 236-246.

Human Mortality Database (2019) University of California, Berkeley, USA, and Max Planck Institute for Demographic Research, Germany. Available at: www.mortality.org (Data downloaded on 01/02/2019).

Hunt, A. and Blake, D. (2015). Modelling mortality and longevity risk management - chapter 4: Consistent mortality projections allowing for trend changes and cohort effects. Doctoral thesis Cass Business School.

Lee, R., & Carter, L. (1992). Modeling and Forecasting U.S. Mortality. Journal of the American Statistical Association, Volume 87(419), 659-671.

Schupp, J. (2019). On the Modeling of variable Mortality Trend Processes. Working Paper, Ulm University.

Sweeting, P. (2011). A Trend-Change Extension of the Cairns-Blake-Dowd Model. Annals of Actuarial Science, 5(2): 143-162.

About The Society of Actuaries

With roots dating back to 1889, the Society of Actuaries (SOA) is the world's largest actuarial professional organizations with more than 31,000 members. Through research and education, the SOA's mission is to advance actuarial knowledge and to enhance the ability of actuaries to provide expert advice and relevant solutions for financial, business and societal challenges. The SOA's vision is for actuaries to be the leading professionals in the measurement and management of risk.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org