

# MODELING THE MORTALITY TREND UNDER MODERN SOLVENCY REGIMES

BY

MATTHIAS BÖRGER, DANIEL FLEISCHER AND NIKITA KUKSIN

## ABSTRACT

Stochastic modeling of mortality/longevity risks is necessary for internal models of (re)insurers under the new solvency regimes, such as Solvency II and the Swiss Solvency Test. In this paper, we propose a mortality model which fulfills all requirements imposed by these regimes. We show how the model can be calibrated and applied to the simultaneous modeling of both mortality and longevity risk for several populations. The main contribution of this paper is a stochastic trend component which explicitly models changes in the long-term mortality trend assumption over time. This allows to quantify mortality and longevity risk over the one-year time horizon prescribed by the solvency regimes without relying on nested simulations. We illustrate the practical ability of our model by calculating solvency capital requirements for some example portfolios, and we compare these capital requirements with those from the Solvency II standard formula.

## KEYWORDS

Stochastic mortality, mortality trend risk, longevity risk, solvency, one-year view.

## 1. INTRODUCTION

In Europe, new solvency regimes such as Solvency II and the Swiss Solvency Test (SST) imply significant changes in the risk management of insurance companies. Both regimes are conceptually similar and follow the common goal of a comprehensive modeling and assessment of risks insurance companies are exposed to, both in terms of number of risk factors as well as severity of risks.<sup>1</sup>

Under both solvency regimes, capital requirements can either be calculated via a standard formula or an internal model. However, companies are encouraged to develop internal models which reflect company-specific risk profiles as closely as possible. In contrast to the scenario-based standard formulas, internal models are generally stochastic and thus technically more complex. However, in both model types, all relevant risk categories need to be explicitly modeled. For a life insurance company, longevity and/or mortality risks clearly represent such important categories.

In this paper, we focus on mortality trend risk which we define as the risk of unexpected changes in the (long-term) trend underlying the evolution of future mortality rates. This risk is relevant for insurance products which pay out contingent on survival (e.g., annuities) or death (e.g., life insurance) of the insured person. We are therefore considering a common “trend risk component” in business exposed to mortality risk as well as in business exposed to longevity risk. The present analysis excludes catastrophe risk, i.e. the risk arising from natural disasters, pandemics, etc., or random risk, i.e. the risk of random fluctuation arising due to the small size of a portfolio of contracts. The former type of risk requires specific models which differ significantly from trend models, the latter is merely an application of the binomial distribution once underlying mortality rates are given.

For the simulation of mortality rates, a wide range of different models has been proposed in the literature, and for an overview of different types of models, we refer to Cairns *et al.* (2008). However, only very few of the existing models are directly applicable to the calculation of capital requirements under Solvency II or the SST. Specific features of these solvency regimes, such as the quantification of risk over a one-year time horizon, impose strong requirements on the models. Moreover, many models do not meet requirements relevant in practice, e.g. the simultaneous modeling of mortality for several populations.

In this paper, we propose a mortality model which satisfies all these requirements but, at the same time, is not limited to the application for solvency purposes. We discuss all aspects of model specification, calibration and application and provide ideas and practical advice for the implementation of each of these steps. Moreover, we compare capital requirements based on our model to those from the Solvency II standard formula.

With respect to specification, calibration and application, our model addresses the following aspects in particular:

- The model allows for the direct quantification of mortality/longevity risk over pre-defined time horizons as well as in run-off simulations. The pre-defined time horizon is particularly relevant for Solvency II and the SST where risk is measured as the 99.5% Value-at-Risk (VaR) or the 99% expected shortfall, respectively, of the own funds over one year. As Börger (2010) points out, a mortality model therefore does not only have to provide mortality rates for the forthcoming year but also realistic changes in the long-term mortality trend assumption within one year. In fact, these changes in the mortality

trend generally have a much more significant impact on the capital requirement than the one-year realized mortality rates. We account for the one-year view by introducing a stochastic trend component in our model, i.e. we explicitly model changes in the long-term trend assumption. As we will see later, this stochastic trend component also allows for the generation of run-off scenarios which consist of iterative one-year scenarios. In most existing models, in particular in the standard mortality models such as Lee–Carter or Cairns–Blake–Dowd, long-term trend changes are not accounted for explicitly. Here, a full model re-estimation after one year would be required in order to update the long-term trend (nested simulations). Such re-estimations in each simulation path can be very time consuming, and thus, a stochastic trend component is generally more efficient. Moreover, a full model re-estimation implies inconsistency between simulations for different time horizons. In general, it should not matter whether one generates a multi-year mortality path at once or a series of consecutive one-year paths. However, when all model parameters are re-estimated as part of each one-year path, this consistency between multi-year and consecutive one-year simulation paths is violated. This is particularly relevant under Solvency II, where in pillar 1 risk needs to be quantified over a one-year time horizon and in pillar 2 over a 3- to 5-year time horizon. The trend modeling approach proposed in this paper avoids this inconsistency by treating multi-year simulation paths as a series of consecutive one-year paths.

- The model covers a wide age interval, which in this paper we have assumed to range from 20 to 105 years. This is necessary as we want to consider mortality and longevity risk simultaneously in order to exploit diversification effects. Mortality risk is most relevant for younger ages, whereas longevity risk particularly affects older ages. Therefore, for a realistic assessment of diversification effects, a plausible and, in particular, non-trivial dependency structure between the mortality evolutions for different ages is necessary. The five stochastic drivers in our model allow for such a dependency structure, and we pay special attention to this when calibrating the model. The five drivers also lead to large variability in the simulation outcomes, with one driver being dedicated to cohort effects. Given the clear interpretation of all components in our model and given the fact that we cover the full age range, we are convinced that our model is still not overparameterized.
- Since insurance portfolios typically cover several populations, we specify the model in a multi-population setting. These can be populations from different countries or just males and females in the same country. Strong dependencies between those populations make simultaneous modeling indispensable, and we do this by introducing a common mortality trend for all populations and deviations from this common trend for the individual populations.
- As we primarily specify our mortality model for the purpose of risk capital calculation under Solvency II and the SST, we are particularly concerned about tail scenarios. Therefore, we focus on such scenarios in the model calibration and show how some crucial model parameters can be fitted to

extreme mortality events in historical data. The idea behind this approach is that a model for solvency purposes should at least be able to reproduce the most severe historical mortality events with sufficiently high probability. The more common calibration of mortality models to all available data might yield more plausible mortality scenarios close to the best estimate, but it is unclear whether extreme scenarios, e.g. 200-year scenarios for Solvency II, derived from such model calibrations are adequate. For instance, the volatility in mortality rates for most populations is rather low at the moment but might increase again suddenly as it has happened in the past. To account for such possibilities, we analyze where volatility add-ons should be introduced in our model, and we show how they can be calibrated in practice.

- The fitting of mortality models to historical data naturally limits their simulation outcomes to paths which are structurally similar to those observed in the past. In general, events never encountered before, such as the development of a cure for cancer, cannot be properly reflected. We show how such events can be incorporated in risk capital calculations by introducing expert judgment in form of mortality/longevity threat scenarios. Such scenarios can also be used to check the plausibility of a mortality model and/or to adjust the range of simulation outcomes.
- Finally, our model allows for the simulation of mortality rates as well as deviations from best estimate rates. Such deviations are of particular importance as data on insurance portfolios are typically not sufficient for fitting a full mortality model. Using the proposed approach, future mortality rates can be simulated for the general population in the country under consideration, and mortality deviations for the general population can be applied to best estimate mortality rates specific to the particular insured population.

Many of the above issues have been discussed in the literature, but to our knowledge, there is no mortality model yet which accounts for all of them. The model which comes closest seems to be the one proposed by Plat (2011). He models changes in the long-term mortality trend assumption explicitly with a clear focus on mortality and longevity risk over a one-year time horizon. However, there are also considerable differences compared to the model which is proposed in this paper. Plat models mortality reduction factors, while we consider mortality rates. Moreover, his model does not include any margins, and he fits it as a best estimate model to the historical mortality evolution over several decades. As an alternative with a stronger focus on tail scenarios, we propose a calibration of some important model parameters to extreme mortality/longevity events which is more in line with the intended model application. Finally, the models differ in the way several populations are considered simultaneously. While we follow an approach which is structurally similar to cointegration, Plat does not assume a long-term relationship between the mortality evolutions of different populations but only allows for correlation between the year to year innovations.

The topic of multi-population mortality modeling has been discussed by several authors in recent years. Our approach particularly builds on the findings of

Li and Lee (2005) and Jarner and Kryger (2011) who assume a common long-term trend for all populations under consideration and allow for short- to mid-term deviations from this common trend for each population. Li and Hardy (2011) show that such an approach outperforms its alternatives when modeling Canadian and US female mortality, and they apply it in order to measure basis risk in longevity index hedges. An alternative concept has been proposed by Hyndman *et al.* (2011) who do not assume stationary deviations from a common trend but assume stationarity in the ratio of mortality rates for any two populations. Cairns *et al.* (2011) and Dowd *et al.* (2011) introduce models for two populations which are particularly suitable for cases of a larger reference population and a smaller but closely related second population, e.g. a subpopulation of insured. These models do not contain a common trend, but the mortality evolution of the smaller population is linked to the evolution of the reference population.

There is also a considerable literature on the quantification of mortality and/or longevity risk under Solvency II. Richards *et al.* (2012) describe different approaches for quantifying longevity risk over a one-year time horizon via standard mortality models such as Lee–Carter or Cairns–Blake–Dowd. However, the approach which best matches the VaR methodology of Solvency II includes full model re-estimation in each simulation path which we can avoid by relying on the stochastic trend component. Stevens *et al.* (2010) derive closed-form approximations for capital requirements based on the Lee–Carter model in order to avoid model re-estimation. However, both Richards *et al.* (2012) and Stevens *et al.* (2010) work in a single population setting and do not consider mortality risk and its diversification with longevity risk. Hari *et al.* (2008) and Olivieri (2011) use more sophisticated mortality models but focus on longevity risk and the impact of systematic (trend) and non-systematic (random) risk. Moreover, Olivieri (2011) considers a longer time horizon than the one year prescribed under Solvency II and the SST, hence avoiding the modeling of trend changes. The same holds for Olivieri and Pitacco (2008a,b). Börger (2010) does consider longevity risk on a one-year time horizon but focuses on the adequacy of the scenario approach in the Solvency II standard formula.

The remainder of this paper is organized as follows. In the following section, we introduce the structure of our mortality trend model and show how it can be fitted to historical data. Moreover, we discuss the simulation of both future mortality rates and changes in the best estimate mortality trend in a multi-population framework. In Section 3, we explain how volatility add-ons and weighting parameters can be fitted to extreme historical mortality events such that their calibration is most appropriate for solvency capital calculations. We then compare capital requirements calculated with our model to those obtained from the Solvency II standard formula for some representative insurance contracts/portfolios. Subsequently, the sensitivity of the capital requirements to the calibration of the volatility add-ons and the weighting parameters is analyzed. Section 6 contains a discussion on how demographic and epidemiological expert judgment can be used to check a statistical mortality model and to enrich

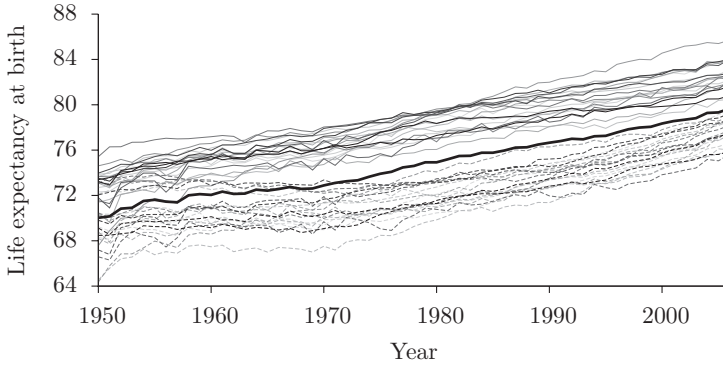


FIGURE 1: Period life expectancies at birth for various populations (solid: females, dashed: males, black: total population): Australia, Austria, Belgium, Canada, Denmark, Finland, France, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA.

its outcomes. This is followed by an overview of all model parameters and short summaries of their meanings and possible calibration approaches in Section 7. Subsequently, we outline the uncertainties in our modeling framework and assess their potential impact. Finally, Section 9 concludes.

## 2. MORTALITY TREND MODEL

As outlined in the Introduction, we want to model mortality in a multi-population setting. In principle, we do this by modeling one common long-term mortality trend for all populations, or the “total population”,<sup>2</sup> and some short- to mid-term deviations from this trend for individual populations. Figure 1 supports this approach. It shows the evolution of life expectancies for a selection of countries across the world, and the existence of a common and rather steady general trend with level deviations for individual countries is clearly visible. However, such deviations seem to be rather short-term, and life expectancies tend to revert back to the overall trend.

In the following sections, we introduce the basic specification of our model and explain how the parameters can be estimated from historical data. This is identical for the total population and all individual populations under consideration. In Section 2.3, we then discuss the simulation of future mortality rates. In particular, we explain the link between the individual populations and the total population.

### 2.1. Basic model structure

We propose to model mortality rates  $q_{x,t}$  as

$$\begin{aligned} \text{logit } q_{x,t} = & \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}(x - x_{\text{center}}) + \kappa_t^{(3)}(x_{\text{young}} - x)^+ \\ & + \kappa_t^{(4)}(x - x_{\text{old}})^+ + \gamma_{t-x}, \end{aligned} \tag{1}$$

where  $\text{logit}(\cdot) = \ln(\cdot/(1 - \cdot))$  is the logit function and  $x^+ = \max\{x, 0\}$ . The term  $\kappa_t^{(1)}$  describes the general, age-independent evolution of mortality. The term  $\kappa_t^{(2)}$  reflects the “slope” of the logit  $q_{x,t}$  or the mortality steepness, and the parameter  $x_{\text{center}}$  should be set somewhere in the middle of the age range under consideration. We set  $x_{\text{center}} = 60$  as we consider ages between  $x_{\text{min}} = 20$  and  $x_{\text{max}} = 105$ . The terms  $\kappa_t^{(3)}(x_{\text{young}} - x)^+$  and  $\kappa_t^{(4)}(x - x_{\text{old}})^+$  account for additional effects and larger volatility which are typically observed in the mortality at young and old ages. Moreover, the terms allow for flexibility regarding the dependency between the mortality evolutions for different ages. An analysis of historical data for different populations suggests setting  $x_{\text{young}} = 55$  and  $x_{\text{old}} = 85$ . The choice of  $x_{\text{young}}$  so close to  $x_{\text{center}}$  might seem counterintuitive at first sight as one would expect significant additional volatility for ages between 20 and 40 in particular. However, since the factor  $(x_{\text{young}} - x)^+$  converges to zero for  $x$  approaching  $x_{\text{young}}$ , the parameter  $x_{\text{young}}$  has to be chosen somewhat larger than 40 in order to generate significant additional volatility for age 40 and below. For a different age range and/or some specific populations, it might be reasonable to use other values for the parameters  $x_{\text{young}}$ ,  $x_{\text{center}}$  and  $x_{\text{old}}$ . However, the concept of separating the age range into young, medium and old ages should be preserved. The process  $\gamma_{t-x}$  represents possible cohort effects, and finally, the parameters  $\alpha_x$  account for age-dependent mortality characteristics which are not covered by the other components of (1), e.g. the “accident hump”. In case of a shorter age range or insignificant cohort effects, the model structure can and should obviously be adjusted accordingly.

Our model structure is rather close to the one proposed by Plat (2009). However, we model the logit of the  $q_{x,t}$  instead of the logarithm of the central mortality rates  $m_{x,t}$ . We choose this approach because, firstly, the  $q_{x,t}$  are the quantities which are actually required for the calculation of insurance reserves and cash flows and, secondly, because it is very convenient for parameter estimation as we will see in the following section.<sup>3</sup> A term of (1) which is not incorporated in the model of Plat (2009) is the term for old ages. We have conducted analyses of various populations<sup>4</sup> and have found this parameter to be significant. Without this term, the volatility at old ages, where very little data are available, would be underestimated.

## 2.2. Parameter estimation

In the following, we describe how all parameters in (1) can be estimated. For the remainder of this paper, let historical mortality rates  $q_{x,t}$  of a given population be available for ages  $x = x_{\text{min}}, \dots, x_{\text{max}}$  and years  $t = t_{\text{min}}, \dots, t_{\text{max}}$ . The full parameter estimation is carried out in four steps:

*Step 1:* First, we estimate the parameters  $\alpha_x$  by setting them to the mean of the logit  $q_{x,t}$  over time, i.e.

$$\alpha_x = \frac{1}{t_{\max} - t_{\min} + 1} \sum_{t=t_{\min}}^{t_{\max}} \text{logit } q_{x,t}.$$

This is mainly done for reasons of parameter uniqueness and computation time. A refinement of these parameters will be carried out in Step 3.

*Step 2:* We then estimate the parameters  $\kappa_t^{(i)}$  for each year  $t = t_{\min}, \dots, t_{\max}$  by a Generalized Linear Model (GLM). We assume a binomial distribution for the annual number of deaths given initial exposures, i.e. the population size at the beginning of each year. In that case, the logit is the canonical link function. The required numbers of deaths and initial exposures for each population are available from the Human Mortality Database. The underlying GLM equation for year  $t$  is

$$\begin{pmatrix} \text{logit}(q_{x_{\min},t}) - \alpha_{x_{\min}} \\ \text{logit}(q_{x_{\min}+1,t}) - \alpha_{x_{\min}+1} \\ \vdots \\ \text{logit}(q_{x_{\max},t}) - \alpha_{x_{\max}} \end{pmatrix} \approx M \begin{pmatrix} \kappa_t^{(1)} \\ \vdots \\ \kappa_t^{(4)} \end{pmatrix}, \tag{2}$$

where  $M$  is the coefficient matrix induced by (1).

*Step 3:* We apply some adjustments in order to facilitate the interpretation of model parameters and to make them comparable across numerous populations. Please note, however, that these adjustments do not affect the response, i.e. the fitted  $\text{logit } q_{x,t}$ . We aim at having parameters  $\alpha_x$  around zero, and at explaining the essential components of the evolutions in mortality rates by the parameters  $\kappa_t^{(i)}$ . To this end, let  $\varphi_1$  be the slope of a regression line to  $\alpha_x$  for  $x \in \{x_{\text{young}}, \dots, x_{\text{old}}\}$ . In the following,  $a \leftarrow b$  denotes that  $a$  gets assigned the value of  $b$ ,<sup>5</sup> and for all  $x$  and  $t$ , we set

$$\begin{aligned} \kappa_t^{(2)} &\leftarrow \kappa_t^{(2)} + \varphi_1, \\ \alpha_x &\leftarrow \alpha_x - \varphi_1(x - x_{\text{center}}), \\ \varphi_2 &= \alpha_{x_{\text{center}}}, \\ \alpha_x &\leftarrow \alpha_x - \varphi_2, \\ \kappa_t^{(1)} &\leftarrow \kappa_t^{(1)} + \varphi_2. \end{aligned}$$

After these adjustments, the parameters have the desired interpretation:  $\kappa_t^{(1)}$  describes the overall level of mortality in year  $t$ ,  $\kappa_t^{(2)}$  is the slope in the  $\text{logit } q_{x,t}$  in that year, and the  $\alpha_x$  only account for age-dependent structures which the  $\kappa_t^{(i)}$  processes cannot describe. With such clear interpretations, the parameters for different populations become comparable. For instance, the overall level of mortality in any two populations can now be compared solely based on  $\kappa_t^{(1)}$ . Without these adjustments, also differences in the  $\alpha_x$  would have to be taken into account.



*Step 4:* Finally, we estimate the cohort parameters  $\gamma_{t-x}$  by fitting them to the residuals from Step 2 using least squares. When doing that, we neglect the 10 oldest and youngest cohorts since there are too little historical data available to reasonably estimate their parameters. At first sight, it seems somewhat counterintuitive not to fit the cohort parameters together with the  $\kappa_t^{(\cdot)}$  parameters in Step 2 as this would improve the overall model fit. However, a simultaneous estimation implies significant additional complexity in a multi-population framework. The cohort parameters can and often do take up some of the mortality trend over time which then has to be projected into the future coherently for all populations under consideration. If, on the other hand, existing trends in the cohort parameters are ignored, forecasts can become biased. We comment on this issue in more detail in Section 2.3.3.

### 2.3. Simulation of future mortality rates

Having estimated the model parameters for all individual populations and the total population, we can now start projecting the mortality rates. Figure 2 shows the fitted  $\kappa_t^{(\cdot)}$  processes for some example populations spread across the world, from 1950 to 2006. For the  $\kappa_t^{(1)}$  processes (top left panel), we observe a common trend similar to the one we have found in the life expectancies (see Figure 1). We expect this common trend to continue in the future and show how this can be achieved by first projecting  $\kappa_t^{(1)}$  for the total population and then adding deviations for the individual populations. For the other  $\kappa_t^{(\cdot)}$  processes, we do not observe such clear common trends. Therefore, we simply project these processes for each individual population (considering correlations) and not in relation to the total population. The cohort parameters  $\gamma_{t-x}$  lie around zero for all populations simply by the way they are estimated. Thus, they do not possess any trend and can be projected for each population individually as well.

**2.3.1. Projection of  $\kappa_t^{(1)}$ .** The process  $\kappa_t^{(1)}$  is the most important process in our model with respect to mortality trend risk. This is because the long-term mortality trend assumption is essentially determined by the assumed trend in  $\kappa_t^{(1)}$ . Thus, changes in  $\kappa_t^{(1)}$  and its trend have a significant impact on reserves/liabilities and hence also capital requirements. Therefore, adequate projection of  $\kappa_t^{(1)}$  is crucial for a realistic assessment of mortality and longevity trend risk.

*A new stochastic process for  $\kappa_t^{(1)}$ .* We commence by deriving an appropriate structure for the stochastic process to describe the future evolution of  $\kappa_t^{(1)}$  for any given population. In Figure 2, we observe rather linear trends for most populations with some occasional breaks where the slope of the linear trend changes. Therefore, it is obvious to model the future evolution of  $\kappa_t^{(1)}$  as a linear trend with stochastic drift (and intercept) and some additional noise around this stochastic trend. The assumption of a linear time trend in mortality is very

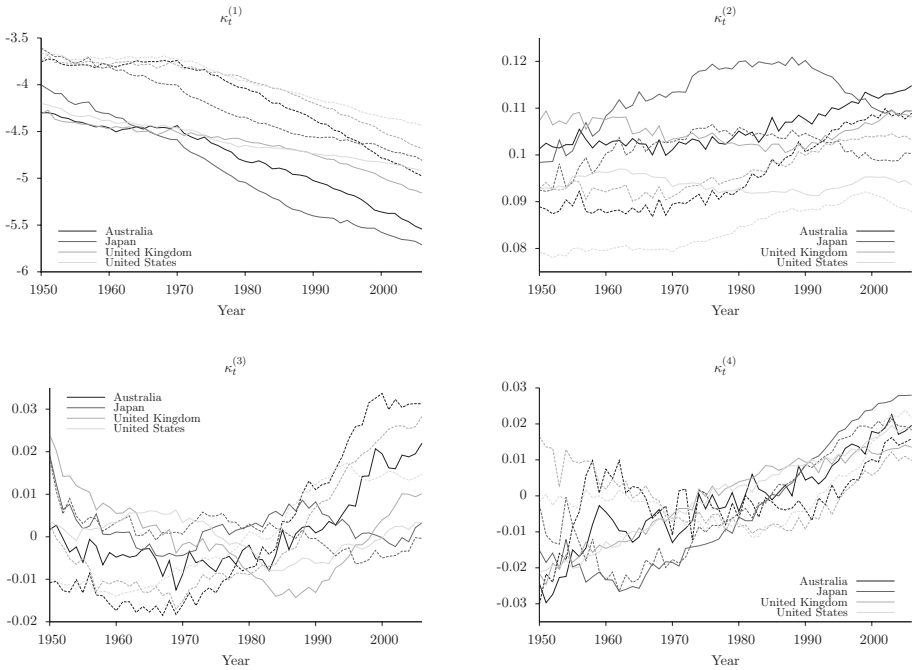


FIGURE 2: Parameters  $\kappa^{(1)}$ ,  $\kappa^{(2)}$ ,  $\kappa^{(3)}$  and  $\kappa^{(4)}$  for some example populations (solid: females; dashed: males).

common; for instance, it holds whenever a random walk with drift is applied in order to project time trends in parametric mortality models. However, it is also possible to consider other trend structures in our setting. For instance, a quadratic trend would take into account curvature in the historical  $\kappa_t^{(1)}$  process as can be observed for the United Kingdom (UK) (cf. Figure 2). However, despite the better fit, we refrain from using a quadratic trend since its adequacy for forecasting is questionable (see also Currie *et al.*, 2004). For the UK, it would presume continuously increasing mortality improvements for the future.

In order to introduce stochasticity in the slope of the linear trend, we take up the idea of model re-estimation which we mentioned in the Introduction. We apply this idea to  $\kappa_t^{(1)}$  by re-estimating the linear trend each year based on a data set which includes the additional realization of  $\kappa_t^{(1)}$  for the then current year. This approach matches with actuarial practice in deriving mortality projections. One always considers the most recent data, determines the trend in these data and projects it into the future.<sup>6</sup>

The projection starts with fitting a regression line  $\ell_{t_{\max}}$  to the historical  $\kappa_t^{(1)}$ ,  $t = t_{\min}, \dots, t_{\max}$ , by weighted least squares. This line constitutes the current best estimate trend. As weights, we apply  $w_t = (1 + 1/h)^{t-t_{\max}}$  which fade out going backward in time. The parameter  $h$  determines how much weight is assigned to most recent data points compared to older data points.<sup>7</sup> The smaller

is  $h$ , the greater is the impact of an additional data point and thus the stronger the changes in the best estimate trend within one year. In order to obtain stochastic forecasts for  $\kappa_{t_{\max}+1}^{(1)}$ , we then add a (normally distributed) random variable on top of the regression line at  $t_{\max} + 1$ . Then, we recalculate the regression line based on the  $\kappa_t^{(1)}$  values up to  $t_{\max} + 1$  and repeat the procedure to obtain forecasts for years  $t_{\max} + 2$ , etc.

More precisely, a stochastic forecast for  $\kappa_t^{(1)}$ ,  $t > t_{\max}$  and given data up to time  $t - 1$ , is obtained as

$$\kappa_t^{(1)} = \ell_{t-1}(t) + \varepsilon_t^{(1)}(\bar{\sigma}^{(1)} + \sigma^{(1)}),$$

where  $\ell_{t-1}(t)$  is the weighted regression line on the  $\kappa_t^{(1)}$  up to year  $t - 1$ , evaluated for year  $t$ , and  $\varepsilon_t^{(1)} \sim \mathcal{N}(0, 1)$  identically and independently. The volatility  $\sigma^{(1)}$  is the weighted sample standard deviation of the empirical errors  $\kappa_t^{(1)} - \ell_{t-1}(t)$  for  $t = t_{\min} + 2, \dots, t_{\max}$ , where the weights are as above but with weighting parameter  $h_\sigma$ .<sup>8</sup> The term  $\bar{\sigma}^{(1)}$  is an optional volatility add-on to ensure a reasonably conservative volatility for solvency purposes. When the model is applied in other settings, in particular as a best estimate model, this add-on may of course be omitted.<sup>9</sup> The calibration of the weighting parameters and the add-on obviously depends on the population under consideration, and we refer to Section 3 for practical examples.

It is important to note that, in contrast to the volatility, the mean of the innovations  $\varepsilon_t^{(1)}$  should not be derived from the historical errors  $\kappa_t^{(1)} - \ell_{t-1}(t)$ ,  $t = t_{\min} + 2, \dots, t_{\max}$ , but should be set equal to zero. If the innovations had a non-zero mean, the fundamental assumption that the regression line fitted to most recent data constitutes the best estimate trend at any point in time would be violated. Thus, in our setting, we “know” that the mean of the innovations is zero. If they had a negative (positive) mean, one would already know today that future  $\kappa_t^{(1)}$  and thus also future mortality rates would be likely to be smaller (larger) than anticipated according to today’s regression line. Therefore, the regression line would not constitute a best estimate trend. In fact, determination of the best estimate for future mortality rates would become fairly complex as it would have to take into account the systematic year to year changes in the regression lines which stem from the  $\kappa_t^{(1)}$  being mostly below (above) the last regression line.

However, a mean in the historical errors which is significantly different from zero can be accounted for indirectly in our setting. Such a mean would indicate that there is more curvature in the historical  $\kappa_t^{(1)}$  than the chosen trend allows for. If one changed to a trend of higher order, the model fit would be improved and the mean of the historical errors would be closer to zero.

Figure 3 shows, for UK males, 99% confidence intervals for  $\kappa_t^{(1)}$  based on some example values for the weighting parameter  $h$ .<sup>10</sup> In the left panel, we see confidence bounds for the one-year view, i.e. percentiles of the realized  $\kappa_t^{(1)}$  for the first year and percentiles for the updated best estimate  $\kappa_t^{(1)}$  thereafter. We

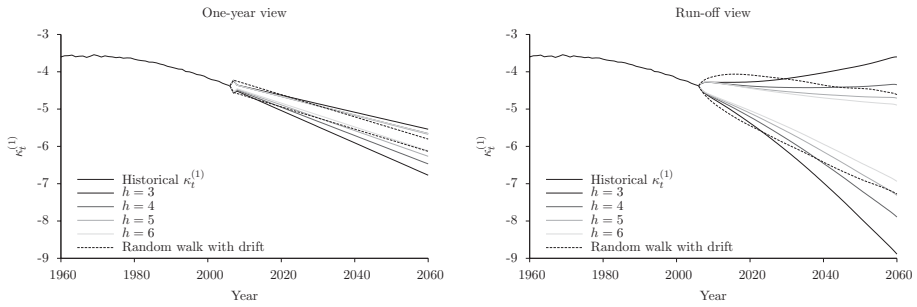


FIGURE 3: 99% confidence intervals based on the stochastic linear trend approach with different parameter values  $h$  and the commonly used random walk with drift approach.

observe that our trend approach allows for significant uncertainty in the realized  $\kappa_t^{(1)}$  but that the long-term trend changes only slightly within one year. In the right panel, the one-year view is applied iteratively which results in the run-off view. In both cases, we see that the smaller  $h$ , the larger the (trend) uncertainty.

Note that the above approach of projecting  $\kappa_t^{(1)}$  can also be formulated in terms of an autoregressive time series process. Let  $n = t_{\max} - t_{\min} + 1$  be the number of years of historical data. Then, the re-estimation of the linear trend can be written as an AR( $n$ ) process<sup>11</sup>

$$\kappa_t^{(1)} = a_1 \kappa_{t-1}^{(1)} + \dots + a_n \kappa_{t-n}^{(1)} + \varepsilon_t(\bar{\sigma}^{(1)} + \sigma^{(1)}),$$

where  $a_j = (1 + 1/h)^j (c_1 + c_2 j)$  for values  $c_1, c_2$  which can be expressed in terms of  $n$  and  $h$ . The most significant advantage of this formulation is that it provides a calculation method which is faster than repeated least squares regression. This formulation also shows the similarities between our regression approach and the concept of (double) exponential smoothing which has been applied, amongst others, by Plat (2011).<sup>12</sup> In both approaches, future time series values are determined via a linear function of the historical observations.

**Comparison with the random walk with drift.** In the literature, the process  $\kappa_t^{(1)}$  is usually projected as a random walk with drift, and in the following, we compare the stochastic trend process from the previous section to the random walk. Both processes share the normally distributed innovations, i.e. the  $\varepsilon_t^{(1)}$ , which describe the short-term uncertainty in the mortality evolution. The most important conceptual difference between the processes lies in the modeling of the long-term trend. In case of the random walk, this trend is fixed by the drift parameter once the random walk is estimated as part of the model fitting. In contrast to that, the annual re-estimation of the trend  $\ell_{t-1}(t)$  allows for random changes in this trend as part of the simulation. Therefore, we refer to the trend as being stochastic in our  $\kappa_t^{(1)}$  process.

The stochastic trend approach offers several advantages compared to the random walk with drift. First, it can be directly applied to quantify mortality/

longevity risk over limited time horizons, e.g. one year, as well as in a portfolio run-off (see Figure 3). The random walk with drift is primarily intended for run-off simulations. This does not mean that it cannot be applied for limited time horizons as well. However, that would require model re-estimation at the end of the simulation horizon which generally implies a significant computational effort. Without model re-estimation, the fixed drift would result in implausible time constant confidence intervals as shown in the left panel of Figure 3 for the one-year case.<sup>13</sup> Moreover, model re-estimation can be critical since it implies inconsistencies between simulations over different time horizons (see also the Introduction). When all model parameters can change in the re-estimation after one year, it makes a difference whether one simulates for two years at once or twice for one year with a re-estimation in between. The stochastic trend approach overcomes this issue as each path in a multi-year simulation is a combination of consecutive one-year simulation steps.

Also for the run-off view, the stochastic trend approach seems more adequate. The random walk with drift has come under some criticism for yielding, e.g. in the Lee–Carter model, rather wide confidence bounds in the short run and implausibly narrow confidence bounds in the long run (cf., e.g., Lee and Miller, 2001). In the right panel of Figure 3, we see that the stochastic trend approach overcomes this issue. The depicted 99% confidence intervals are narrower in the short run but can be significantly wider in the long run depending on the calibration of the weighting parameter  $h$ . Furthermore, the long-term uncertainty in the random walk with drift case is fully determined by the one-year random fluctuations in  $\kappa_t^{(1)}$  which is not plausible in general. If random fluctuations for a rather small population are, say, twice as high as for a large population, this should not automatically imply that long-term uncertainty is also twice as high. In fact, it may well be the case that the long-term uncertainty is equal for both populations, in particular if they are closely related in terms of social, political and economic factors. In our stochastic trend approach, we are able to disentangle the one-year random fluctuations and the long-term uncertainty by appropriate calibrations of the parameters  $(\bar{\sigma}^{(1)} + \sigma^{(1)})$  and  $h$ .

**Projection of  $\kappa_t^{(1)}$  for the total population.** The straightforward approach to projecting  $\kappa_t^{(1)}$  in our multi-population setting would be to apply the stochastic trend process from above for each population under consideration. The innovations  $\varepsilon_t^{(1)}$  for different populations could be modeled according to their historical correlation to obtain some common movement. Unfortunately, this approach is not feasible. The mortality evolution of each individual population would be modeled adequately, but correlations in the one-year innovations are too low in general to imply reasonable dependencies between the long-term mortality trends of different populations. Even between males and females in the same country, we have only found correlations of around 0.5, with a maximum of 0.7. If we applied such correlations, there would be a decent chance that best estimate trends for closely related populations would diverge after only a few

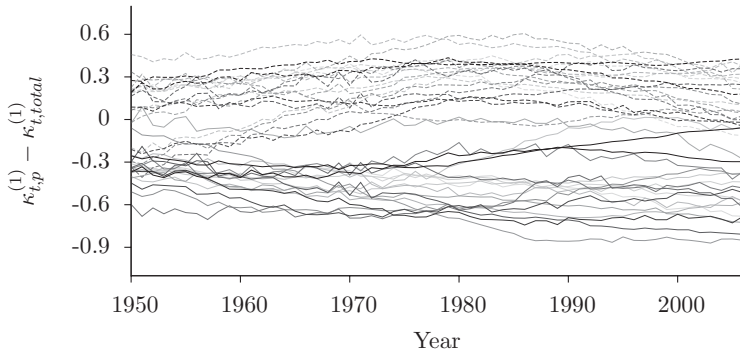


FIGURE 4: Differences  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$  for various populations  $p$  (solid: females, dashed: males): Australia, Austria, Belgium, Canada, Denmark, Finland, France, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA.

years of simulation. This would contradict the strong co-movements of mortality evolutions observed in historical data (see Figures 1 and 2).

A similar phenomenon has been observed by Coughlan *et al.* (2011) and Dowd *et al.* (2011). They find correlations in the same range as ours between the one-year mortality improvements of closely related populations. For aggregate mortality improvements over longer time periods, however, the correlations increase to more than 0.9. Coughlan *et al.* (2011) argue that short-term correlations are probably so low due to noise in the mortality data and thus misleading for a reasonable assessment of long-term dependencies between populations.

Therefore, we use a different approach to coherently model the mortality trends of several populations. Our method builds on the findings of Li and Lee (2005) and Jarner and Kryger (2011) in that we also consider a common trend for all populations under consideration. More precisely, we apply the stochastic trend process described above to the total population only, denoting the resulting process by  $\kappa_{t,total}^{(1)}$ .<sup>14</sup> For the individual populations, we modify this process as described in the following paragraph to allow for short- to mid-term deviations from the common trend.

**Projection of  $\kappa_t^{(1)}$  for the individual populations.** Figure 4 shows the differences between the historical  $\kappa_{t,p}^{(1)}$  and  $\kappa_{t,total}^{(1)}$ , where  $p$  denotes the individual population under consideration. We observe mostly horizontal patterns but substantial variation in the levels of the  $\kappa_{t,p}^{(1)}$ . These observations agree with what was seen in Figure 1. Moreover, the differences  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$  are not constant over time but mostly seem to increase or decrease for some years/decades before starting to revert back. However, the overall range of the differences is rather constant over time, despite the fact that the populations with the largest/smallest differences change several times. These findings suggest modeling the differences as mean reverting processes with population-specific means which describe the long-term relation between the respective population and the total population.

Also from a demographic perspective, mean reversion to some long-term difference appears to be the most plausible assumption. It is hard to imagine that, given the strong political, economic and social links worldwide, mortality trends for populations which we consider here will diverge in the long run.

Using the Bayesian Information Criterion as a goodness-of-fit measure, an AR(1) process is the optimal process to describe the differences  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$  for about two-thirds of the populations under consideration. For the sake of consistency, we therefore model short- to mid-term deviations from the total population’s mortality trend as AR(1) processes for all populations:

$$\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)} = a_p + b_p(\kappa_{t-1,p}^{(1)} - \kappa_{t-1,total}^{(1)}) + \varepsilon_{t,p}^{(1)}, \tag{3}$$

where we apply the constraint  $1 > b_p \geq 0$  to guarantee mean reversion and to avoid oscillations. Note that this approach of modeling a common trend and mean reverting deviations is similar to the concept of cointegration introduced by Engle and Granger (1987). In fact, both approaches are equal in structure. They differ in that we do not estimate the cointegration vector from the data but explicitly use the population sizes for each individual population as well as the (negative of the) total population size. Furthermore, all coefficients corresponding to populations’ year to year differences are zero.

The innovations  $\varepsilon_{t,p}^{(1)}$  are  $\mathcal{N}(0, \sigma_p^{(1)})$  distributed and serially independent. The volatility  $\sigma_p^{(1)}$  is derived from the historical data, weighted according to  $h_\sigma$ . Across populations, we allow for dependencies as defined by the sample correlation matrix of the  $\varepsilon_{t,p}^{(1)}$ . The mean reversion speed  $b_p$  can be obtained from the historical data in general but should be adjusted in the very unlikely case its estimated value is not within the prescribed interval. Given  $b_p$ , the parameter  $a_p$  determines the assumed long-term difference between  $\kappa_{t,p}^{(1)}$  and  $\kappa_{t,total}^{(1)}$ , i.e.  $a_p/(1 - b_p)$ . Obviously,  $a_p$  can be obtained from the historical data as well. However, alternative calibrations based on estimates for the long-term difference  $a_p/(1 - b_p)$  may be more appropriate in some cases. Here, expert judgment is required.<sup>15</sup>

Table 1 shows estimated long-term differences for some populations based on four different methods using historical data from 1950 to 2006:

1. With  $a_p$  derived from the historical data;
2. Unweighted average of the historical differences;
3.  $w$ -weighted average of the historical differences with  $h = 6$ ;
4. Extrapolation of the current linear trend ( $w$ -weighted linear regression, again  $h = 6$ ) of the differences five years into the future.

For UK females, for instance, all four methods yield very similar values for the long-term difference as  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$  has been rather constant over time. For Japanese females, however, the values differ significantly as there has been a long-lasting downward trend in  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$ . We do not know whether and when this trend will break, but it would be bold to assume a long-term difference

TABLE 1  
 EXPECTED LONG-TERM DIFFERENCES  $\kappa_{t,p}^{(1)} - \kappa_{t,total}^{(1)}$  BASED ON FOUR DIFFERENT METHODS FOR SOME  
 EXAMPLE POPULATIONS.

Population	1.	2.	3.	4.
US males	0.34	0.39	0.35	0.34
US females	-0.31	-0.25	-0.14	-0.04
Japan males	-0.07	-0.04	-0.06	-0.05
Japan females	-1.30	-0.73	-0.94	-1.01
Australia males	0.63	0.47	0.22	0.05
Australia females	-0.49	-0.47	-0.55	-0.62
Dutch males	0.09	0.06	0.09	0.05
Dutch females	0.26	-0.54	-0.33	-0.27
Swedish males	-0.20	-0.21	-0.24	-0.30
Swedish females	-0.76	-0.81	-0.72	-0.68
UK males	0.65	0.49	0.35	0.24
UK females	-0.30	-0.30	-0.30	-0.33

$a_p/(1 - b_p)$  according to Method 2. This method would imply an instant and sharp reversion of the historical trend. Method 4, somewhere in between Methods 1 and 3, seems to provide a more reasonable assumption in this case.

2.3.2. *Projection of  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  and  $\kappa_t^{(4)}$ .* As indicated above, Figure 2 suggests that there is no common trend in the  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  or  $\kappa_t^{(4)}$  for different populations. Hence, we do not project them for the total population but only for the individual populations. More precisely, we forecast  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  and  $\kappa_t^{(4)}$  as a three-dimensional random walk, with volatilities and correlations estimated from the historical errors using the same weighting parameter  $h_\sigma$  as for the  $\kappa_{t,total}^{(1)}$  process.

We choose this very simple projection for  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  and  $\kappa_t^{(4)}$  to keep the model complexity as low as possible. Furthermore, the application of more sophisticated time series models would increase computation time in simulations significantly as the three processes have to be simulated for each population under consideration. The random walk also offers the advantage that the best estimates of future  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  and  $\kappa_t^{(4)}$  are equal to the current values of these processes. Thus, no re-estimation is required, in contrast to  $\kappa_{t,total}^{(1)}$ . Moreover, we ensure consistency between multi-year and one-year view since, in case of a random walk, a multi-year simulation path is a combination of consecutive one-year simulation steps. For the process  $\kappa_t^{(2)}$ , a random walk also seems to be a conservative choice. In Figure 2, we observe that the slope in the logit mortality rates has been increasing for most countries. By applying a random walk, however, the best estimate slope for the future is constant at the current level. This means that we tend to overestimate mortality rates for younger ages which are particularly exposed to mortality risk and that we tend to underestimate mortality rates for older ages where longevity risk is more relevant.



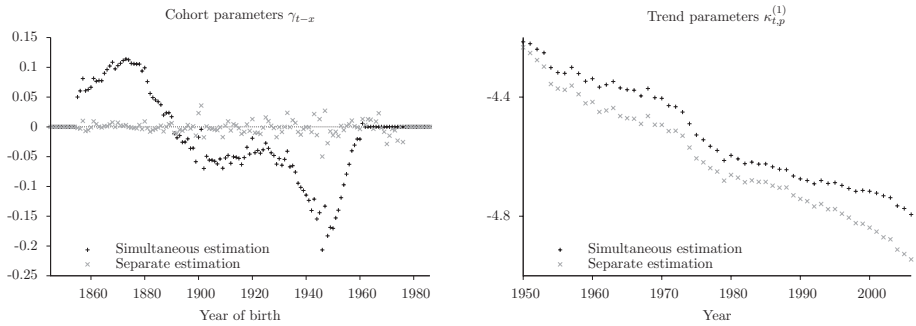


FIGURE 5: Parameters  $\gamma_{t-x}$  and  $\kappa_{t,p}^{(1)}$  for US females and different approaches of fitting the cohort parameters.

For a similar reason, we introduce an optional volatility add-on  $\bar{\sigma}^{(2)}$  for  $\kappa^{(2)}$ , i.e.

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \varepsilon_t^{(2)}(\bar{\sigma}^{(2)} + \sigma^{(2)}).$$

A larger volatility in  $\kappa_t^{(2)}$  reduces dependencies between the mortality evolutions for young and old ages, and thus, this add-on limits diversification effects between mortality and longevity risk. We discuss the calibration of  $\bar{\sigma}^{(2)}$  in Section 3.

An analysis of the historical data of various populations reveals mostly significant and similar correlations between the  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$  for each population. Therefore, we carry over these correlations into the projection. The correlations between  $\varepsilon_{t,p}^{(1)}$ , on the one hand, and  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$ , on the other hand, are mostly insignificant and not persistent. For instance, between  $\varepsilon_{t,p}^{(1)}$  and  $\varepsilon_t^{(2)}$ , the average correlation is  $-0.07$  for the countries cited in Figure 4. We thus interpret them as not persistent and assume independence between  $\kappa_{t,p}^{(1)}$  and the other  $\kappa_t^{(c)}$  processes for the projection.

Across populations, we find significant correlations between the  $\varepsilon_t^{(2)}$  and assume these correlations to persist in the future. Note that the correlations in the  $\varepsilon_t^{(2)}$  for different populations and the correlations between  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$  for each individual population also imply some correlation between the  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$  for different populations.

**2.3.3. Projection of  $\gamma_{t-x}$ .** In Section 2.2, we indicated that the fitting and projection of the cohort parameters  $\gamma_{t-x}$  requires special attention. To explain this in more detail, we have fitted the  $\gamma_{t-x}$  in two ways: as in Estimation Step 4 (separate estimation) and together with the  $\kappa_t^{(c)}$  in Estimation Step 3 (simultaneous estimation).<sup>16</sup> Figure 5 shows the resulting  $\gamma_{t-x}$  and  $\kappa_{t,p}^{(1)}$  for the case of US females from 1950 to 2006.

As expected, the cohort parameters from the separate estimation lie closely around zero without any obvious trend. In the case of simultaneous estimation, on the other hand, the  $\gamma_{t-x}$  follow a more pronounced pattern, and they tend to

decrease with increasing year of birth. Thus, they explain some part of the generally decreasing trend in mortality over time. This is confirmed by the graphs for the  $\kappa_{t,p}^{(1)}$  as the trend is less steep for the simultaneous estimation case. Similar observations can be made for most other populations.

If one wanted to forecast future mortality in the simultaneous estimation case, these observations would have several critical implications. Changes in the general level of mortality over time would not be described by changes in the  $\kappa_{t,p}^{(1)}$  alone. Thus, the  $\kappa_{t,p}^{(1)}$  would lose some of their interpretation, and for an assessment of mortality trends,  $\kappa_{t,p}^{(1)}$  and  $\gamma_{t-x}$  would have to be considered in combination. This also implies that trends in the cohort parameters would have to be projected in the first place. However, mean reverting AR(1) processes are often applied, and thus, trends are ignored and forecasts can get biased. In our example of US females, future mortality would probably be overestimated since the overall trend in mortality is likely to be steeper than the trend in the  $\kappa_{t,p}^{(1)}$  alone. In a multi-population setting, things would become even more complicated since trends in the  $\kappa_{t,p}^{(1)}$  and the  $\gamma_{t-x}$  would have to be projected coherently for several populations. This could even mean that the  $\kappa_{t,p}^{(1)}$  and the  $\gamma_{t-x}$  diverge for two populations as long as the “combined mortality trends” from both processes do not diverge.

In order to avoid these issues, we estimate the  $\kappa_{t,p}^{(c)}$  and the  $\gamma_{t-x}$  separately and accept the slightly worse model fit. This approach is in line with the findings of Hunt and Blake (2013) who state that cohort parameters should only be estimated after all age- and period-dependent effects have been extracted from the data. In the separate estimation case, there is no trend in the cohort parameters for any population, and thus, trends only need to be projected for the  $\kappa_{t,p}^{(1)}$ . In such a setting, we have a clear idea what we regard as coherent between populations (see the previous sections). Moreover, the projection of the  $\gamma_{t-x}$  is fairly straightforward. As we have found only very little, sometimes positive and sometimes negative, serial correlation in the  $\gamma_{t-x}$  for all populations we have looked at, we simply project the cohort parameters as random noise. In contrast to the  $\kappa_{t,p}^{(c)}$  processes, we derive the volatility  $\sigma_\gamma$  without weighting. The latest cohort parameters only depend on a few data points, and therefore, we think it is more reasonable to estimate the volatility from the entire set of historical cohort parameters. Note also that, in case of simulations over just one or a few years, the effect of the cohort component is negligible. For each year of simulation, only one additional cohort enters the model. Moreover, when modeling deviations from the best estimate only, e.g. in order to determine risk capital charges, the historical cohort parameters are basically irrelevant.

### 3. VOLATILITY ADD-ONS AND WEIGHTING PARAMETERS

In the previous section, we introduced four parameters which still need to be calibrated: the volatility add-ons  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$  and the weighting parameters  $h$

and  $h_\sigma$ . The latter define the weights in the fitting of the regression line  $\ell_t(\cdot)$  and the estimation of the volatilities for the  $\kappa_t^{(\cdot)}$  processes, respectively. The add-ons can be used to artificially increase the estimated volatilities in order to have a more conservative model.

More precisely, the add-on  $\bar{\sigma}^{(1)}$  determines in particular how strong sudden drops or increases in realized mortality can be. Note though that we only consider rather persistent mortality changes here. Our model is not meant to cover transitory shocks like the Spanish Flu. The add-on  $\bar{\sigma}^{(2)}$  impacts the dependency between mortality rates for young and old ages. A large add-on implies smaller correlation and thus less diversification between mortality and longevity risk in general. The weighting parameter  $h$  determines how strong one-year changes in the long-term mortality trend for the total population can be. Moreover, it defines the width of the long-term confidence bounds in run-off simulations (see Figure 3). It is important to note that, by  $\bar{\sigma}^{(1)}$  and  $h$ , we are able to separately model the severity of sudden short-term mortality events (by  $\bar{\sigma}^{(1)}$ ) and changes in the long-term mortality evolution (by  $h$ ). By  $h_\sigma$ , we can control how much the volatilities of the  $\kappa_t^{(\cdot)}$  innovations depend on most recent observations. The smaller  $h_\sigma$ , the more sensitive the model is to changes in the volatilities over time.

In the following sections, we explain how these four parameters can be calibrated. We first discuss the calibration rather generally and then illustrate our approach by concrete examples.

### 3.1. Calibration of volatility add-ons and weighting parameters

For the calibration of  $h_\sigma$ , we find it rather difficult to provide a precise criterion. In general, we suggest using a rather large value, e.g. around 30. By a small value, one would implicitly assume that mortality trend risk is time dependent in the sense that there might be large differences in trend risk within only a few years in which the volatility changes. Given the long time horizons in which trend risk manifests itself, this is counterintuitive. Thus, a large value for  $h_\sigma$ , together with a sufficiently large value for  $\bar{\sigma}^{(1)}$ , prevents the typical cyclic model behavior that, throughout periods of low volatility, the measured/perceived risk is continuously decreasing until a regime change “surprisingly” blows up the risk.

The volatility add-on  $\bar{\sigma}^{(2)}$  should be chosen rather large for a conservative mortality model since a large add-on limits the diversification between mortality and longevity risk. On the other hand, a very large add-on can imply negative correlation between mortality rates for young and old ages which is not a reasonable assumption from a demographic point of view. Therefore, we propose fixing  $\bar{\sigma}^{(2)}$  such that the correlation between mortality rates for ages at the boundaries of the age range, i.e. 20 and 105 in our case, is close to zero. This seems to be the most conservative of all demographically reasonable calibrations.

Regarding the add-on  $\bar{\sigma}^{(1)}$  and the weighting parameter  $h$ , we suggest fitting them to severe mortality events/evolutions in the historical data. The idea behind this approach is that a mortality model for solvency purposes should

at the very least be able to reproduce all mortality events we have already observed with reasonably high probability. If that is not the case, it is questionable whether a model will be able to generate sufficiently severe events for the future.

Usually, mortality models are fitted to series of historical data which do not include extreme events. In fact, the available data set for model calibration is often shortened in order to exclude events like World War II. Thus, the model parameters are calibrated such that the model performs well in generating mortality scenarios around the best estimate. However, the adequacy of extreme scenarios, e.g. 200-year scenarios, is unclear. Therefore, we believe that a parameter calibration with a focus on extreme mortality events is more suitable for solvency models.

In such a calibration, the most difficult part certainly is to determine how rare an observed historical mortality event is or, in other words, at which percentile of a mortality simulation it should be observed. This always implies a considerable degree of subjectivity, and one has to rely on expert judgment to some extent. Obviously, this is not a desirable feature of our model, but it is not critical either from our point of view. In fact, similar questions arise in any model that is to be used for solvency purposes. For any model, one has to decide whether volatility estimates are conservative enough or whether the volatilities estimated from the historical data should be increased by some add-ons. Furthermore, the calibration of  $h$  is closely related to the question of how many years of data should be used when re-estimating a mortality model in each simulation path; the more data points are considered, the smaller are the one-year trend changes. The specification of a suitable data period is basically the analog to the calibration of  $h$  in our setting. Therefore, we do not bring up new issues here, but we re-formulate existing ones and provide possible solutions.

In principle, any extreme historical mortality event/evolution could be used to determine  $\bar{\sigma}^{(1)}$  and  $h$ . However, it is wise to choose events which the parameters are particularly related to. Since  $\bar{\sigma}^{(1)}$  defines the severity of sudden mortality drops or increases in our model, at least one of the historical scenarios should represent such a rapid change in the mortality level. The weighting parameter  $h$ , on the other hand, particularly determines the magnitude of trend changes within one year and thus also the width of long-term confidence intervals. Therefore, a scenario with a strong trend change appears to be most suitable for the calibration of  $h$ .

However, when calibrating the parameters  $\bar{\sigma}^{(1)}$ ,  $\bar{\sigma}^{(2)}$ , and  $h$ , one needs to be aware of possible interactions between the parameters (here and in the following, we assume  $h_\sigma$  to be fixed as outlined above). For instance, the larger  $\bar{\sigma}^{(1)}$ , the larger one-year trend changes in  $\kappa_{t, total}^{(1)}$  become if  $h$  is not adjusted accordingly. Thus, the impact of one parameter depends on the calibration of the others. Therefore, the parameters  $\bar{\sigma}^{(1)}$ ,  $\bar{\sigma}^{(2)}$ , and  $h$  need to be calibrated simultaneously. This also means that  $\bar{\sigma}^{(1)}$  and  $h$  cannot be fitted to different historical scenarios independently. Instead, they must be calibrated such that the combination of

both parameters is most appropriate for the re-production of the entire set of historical scenarios. At first sight, one might anticipate an identifiability problem here as different combinations of  $\bar{\sigma}^{(1)}$  and  $h$  might be equally appropriate to re-produce the historical scenarios. However, this is generally not the case due to the significant difference in the parameters' meanings (short-term volatility vs. long-term trend uncertainty). The weighting parameter  $h$  only slightly impacts the severity of sudden drops or increases in the level of mortality, and therefore, this severity is basically determined by  $\bar{\sigma}^{(1)}$ . Based on  $\bar{\sigma}^{(1)}$ ,  $h$  can then be fixed such that long-term trend uncertainty is as desired.

This relationship also indicates a reasonable calibration procedure. The parameters  $\bar{\sigma}^{(1)}$ ,  $\bar{\sigma}^{(2)}$ , and  $h$  can be fitted iteratively, with each parameter updated based on the current calibration of the others. An efficient order for the parameter adjustments in each iteration step is as follows:

1. Update of  $\bar{\sigma}^{(1)}$  such that the short-term mortality event is observed at the desired percentile;
2. Update of  $h$  such that the trend change event is observed at the desired percentile;
3. Update of  $\bar{\sigma}^{(2)}$  such that the correlation between ages at the boundaries of the age range is close to zero.

The add-on  $\bar{\sigma}^{(1)}$  should be updated before  $h$  because of the weak dependency of the short-term event on  $h$ . The add-on  $\bar{\sigma}^{(2)}$  should be updated last since the correlation between mortality rates for young and old ages can only be properly assessed once the volatility in the  $\kappa_{t,p}^{(1)}$  is fixed. The iterative calibration of the parameters can stop as soon as changes in the parameter values from one iteration step to the next become insignificant.

In the following section, we illustrate our calibration approach for the four parameters  $\bar{\sigma}^{(1)}$ ,  $\bar{\sigma}^{(2)}$ ,  $h$  and  $h_\sigma$ . In particular, we provide examples for extreme mortality events in the historical data to which  $\bar{\sigma}^{(1)}$  and  $h$  can be fitted. For a more thorough understanding of the impact of each of the four parameters, we refer to the sensitivity analyses in Section 5. A possible refinement of the parameter values based on hypothetical extreme scenarios is addressed in Section 6.

### 3.2. Illustrative parameter calibration

We illustrate the calibration of  $\bar{\sigma}^{(1)}$ ,  $\bar{\sigma}^{(2)}$ ,  $h$  and  $h_\sigma$  for two different versions of the model: a single population model and a multi-population model. The only difference between these versions arises from the simulation of  $\kappa_{t,p}^{(1)}$ . In the single population model, the stochastic trend approach is directly applied for the population under consideration, while in the multi-population case, it is applied for the total population as explained in the previous section.

For  $h_\sigma$  and  $\bar{\sigma}^{(2)}$ , we follow the approaches which we proposed above. We set  $h_\sigma = 30$  and choose  $\bar{\sigma}^{(2)}$  such that the correlation between mortality rates for ages 20 and 105 is (close to) zero. In the multi-population case, a

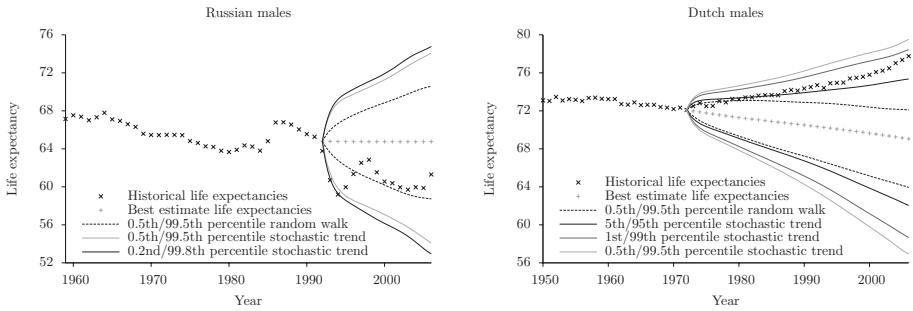


FIGURE 6: Historical life expectancies and relevant percentiles of simulated life expectancies for Russian and Dutch males.

correlation of exactly zero is not achievable for all populations under consideration since  $\bar{\sigma}^{(2)}$  is assumed to be equal for all populations for simplicity. In fact, for the final calibration of  $\bar{\sigma}^{(2)}$ , we found correlations in the range of  $\pm 15\%$  for all populations which are reasonably close to zero from our point of view.

For the calibration of  $\bar{\sigma}^{(1)}$  and  $h$ , we use the most extreme mortality events which we could find in the historical data with respect to sudden drops/increases in the level of mortality and long-term trend changes. As outlined above, it is difficult to state exactly along which percentiles these events should be observed, and we can only provide indications here. To this end, we assume that  $n$  observations are sufficient to define an underlying distribution up to the  $(1 - 1/n)$ th percentile. This obviously is a simplification, and we intend to consider more sophisticated methods in future work (cf. also Section 7).

**3.2.1. Short-term mortality event.** We first look for severe short-term mortality events in the  $\kappa_{t,p}^{(1)}$  for all individual populations for which data are available in the Human Mortality Database. Such an event would allow us to update  $\bar{\sigma}^{(1)}$  as part of the iterative calibration procedure. The most extreme event which we can find is the drop in the life expectancies of Russian males shortly after the fall of communism at the beginning of the 1990s. We see this drop in the crosses in the left panel of Figure 6. The rationale behind our calibration approach is now to fit our (single population) model to the available data up to year 1992 (beginning in year 1959), to start a simulation from there, and to adjust the add-on  $\bar{\sigma}^{(1)}$  such that the drop in life expectancies from 1993 onward can be observed at a reasonable percentile — given preliminary estimates for  $h$  and  $\bar{\sigma}^{(2)}$ .

A suitable percentile could be the 99.8th percentile for the following reason: Data are available for at least 50 years for basically all 37 countries in the Human Mortality Database. However, as the innovations in the  $\kappa_{t,p}^{(1)}$  processes are significantly correlated between those populations — for instance, a similar drop in life expectancy can be observed for Ukrainian males — the number of

actually “independent” observations for each year is much smaller than 37. For a rather conservative, i.e. large, estimate of 0.7 for the correlations between the innovations of different populations, the likelihood of a rare scenario for one of the 37 populations is equal to the likelihood of the same event for one of 10 independent populations. Thus, assuming 50 serially independent innovations, the drop in Russian life expectancy can be interpreted as a 500-year scenario.<sup>17</sup>

This scenario can be applied to determine a reasonable value for  $\bar{\sigma}^{(1)}$  not only for Russian males but for any single population. By carrying the parameter values for Russian males over to other populations, the model is able to generate similarly extreme scenarios also for those populations. Note that this does not necessarily mean that confidence bounds for other populations will be as wide as for Russian males. The width of the confidence bounds also depends on the currently observed volatilities, in particular  $\sigma^{(1)}$ , which may be significantly different from those for Russian males. However, for a population with much smaller volatilities, a drop in life expectancies smaller than the one in Russia can already be regarded as similarly rare from our point of view.

Having found an extreme short-term event for a single population, we now look for a multi-population counterpart. When we consider all populations for which data are available in the Human Mortality Database at least from 1959 onward, that is 64 of the total 74 populations,<sup>18</sup> we can observe a severe short-term event in  $\kappa_{t,total}^{(1)}$ . As indicated above, the drop in life expectancy at the beginning of the 1990s can also be observed for other Eastern European populations, e.g. Ukrainian males, and therefore, it significantly affects the total population, too. Given the roughly 50 years of data for the total population (1959–2006), one might want to see this life expectancy drop at about the 98th percentile. In analogy to the single population case, we fit our multi-population model to the available data up to 1992 (beginning in 1959) and start a simulation for the years from 1993 onward. The add-on  $\bar{\sigma}^{(1)}$  for the multi-population case is again determined such that the desired percentile is met.

**3.2.2. Trend change event.** Next, we look for extreme trend changes in the historical data which can be used to update the weighting parameter  $h$  as part of the iterative calibration procedure. Figure 7 shows the historical  $\kappa_{t,p}^{(1)}$  for various populations and  $\kappa_{t,total}^{(1)}$  for the corresponding total population. Even though we think that extreme short-term mortality events in Russia can be representative also for extreme events in other countries, we have not included populations from former Soviet countries in this figure. In the past, mortality trends in these countries have been significantly different from the common trend in most other industrialized countries around the world. Therefore, trend estimates might get blurred if populations from former Soviet countries were included. However, most of the former Soviet countries seem to be catching up at the moment, and they are likely to follow the same trend as the populations given in Figure 7 in the future (see, e.g., Li and Lee, 2005).

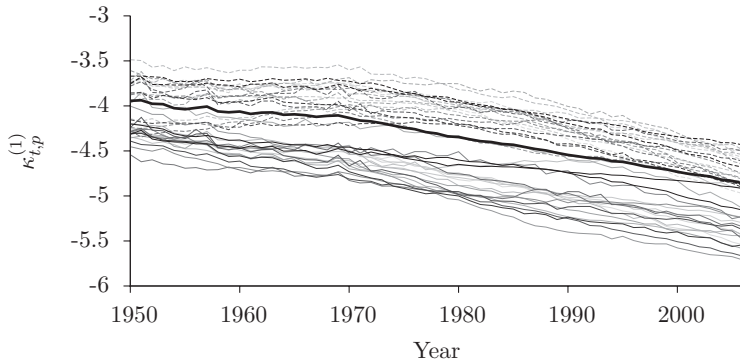


FIGURE 7: Parameters  $\kappa_{t,p}^{(1)}$  for various populations (solid: females, dashed: males): Australia, Austria, Belgium, Canada, Denmark, Finland, France, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA; parameters  $\kappa_{t,total}^{(1)}$  for the corresponding total population in black.

We observe that there has only been one significant trend change in  $\kappa_{t,total}^{(1)}$ , i.e. in 1971. Therefore, we fit our multi-population model to the available data up to 1970 (beginning in 1950) and simulate life expectancies from 1971 onward. The crucial question is now at which percentile of a run-off simulation we roughly want to observe this trend change in the realized  $\kappa_{t,total}^{(1)}$ .

Unfortunately, we cannot proceed as for  $\bar{\sigma}^{(1)}$  and state that the trend change in 1971 is likely to be a one in about 50 years event. A sudden drop or increase in mortality can basically occur every year, and thus, we have as many (approximately independent) observations as years of data. For trend changes, the situation is different as it always takes some years until a trend change becomes observable or, in other words, until we perceive changes in mortality over time as a change in the long-term trend. An exact number of years necessary cannot be specified, and therefore, the historical mortality evolution can only provide some indication for the calibration of  $h$ . Reasonable choices for the percentile at which one might want to observe the trend change in 1971 might lie between the 90th and the 95th percentile.

Finally, we search for an extreme trend change event for a single population to which  $h$  can be calibrated in the single population model variant. As we see in Figure 7, the trend change in 1971 occurred for most populations, and for a single population model,  $h$  should be chosen such that even the strongest trend change amongst all populations is seen at a reasonable percentile. This strongest trend change can be observed for Dutch males, with the corresponding life expectancy evolution shown in the right panel of Figure 6. As in the multi-population case, it is difficult to provide one exact percentile at which we would want to observe the Dutch mortality trend in a run-off simulation starting in 1971. However, the strongest of all country-specific trend changes should obviously be seen at a larger percentile than the trend change for the total population. Therefore, percentiles between the 95th and the 99th percentile might be reasonable choices.



3.2.3. *Results from iterative calibration.* In the multi-population case, we started the iterative calibration of  $\bar{\sigma}^{(1)}$ ,  $h$  and  $\bar{\sigma}^{(2)}$  with initial values  $\bar{\sigma}^{(1)} = \bar{\sigma}^{(2)} = 0$  and  $h = 4$ . However, these values can be chosen rather arbitrarily from our point of view as we have experienced that the iteration converges very quickly. After only two iteration steps, we obtained the final parameter values of  $\bar{\sigma}^{(1)} = 0.02$  (compared to  $\sigma^{(1)} \approx 0.015$  for the total population at the beginning of the 1990s),  $\bar{\sigma}^{(2)} = 0.0001$ , and  $h = 5$ , with the trend change for the total population observed at about the 95th percentile. For  $h = 4$  and only slightly different volatility add-ons, the trend change can be observed at about the 90th percentile. The fast convergence of the iteration is particularly due to the fact that the severity of short-term events is almost exclusively determined by  $\bar{\sigma}^{(1)}$ . Thus,  $\bar{\sigma}^{(1)}$  is already very close to its final value after the first iteration step, and  $h$  and  $\bar{\sigma}^{(2)}$  have to be re-adjusted to significant changes in  $\bar{\sigma}^{(1)}$  only once.

In the single population case, we started the iteration with the same initial values and also required two iteration steps. Here, the optimal parameter values are  $\bar{\sigma}^{(1)} = 0.05$  (compared to  $\sigma^{(1)} \approx 0.06$  in Russia),  $\bar{\sigma}^{(2)} = 0.0005$ , and  $h = 5$ , with the trend change for Dutch males observed at about the 99th percentile. For  $h = 4$ , the trend change can be observed slightly above the 95th percentile. Note though that the values for the add-ons cannot be directly compared to those in the multi-population setting as the modeling approaches for  $\kappa_{t,p}^{(1)}$  and thus also the observed volatilities  $\sigma^{(1)}$  differ.

In addition to the historical life expectancy evolutions for Russian and Dutch males, Figure 6 shows confidence intervals for simulated life expectancies based on the mentioned volatility add-ons and a weighting parameter  $h = 4$ . In both cases, the single population model variant was fitted to the available historical data up to the extreme mortality event, i.e. 1959 to 1992 for Russian males and 1950 to 1970 for Dutch males. Note that the choice of the starting year of the data set only has a small impact on the parameter estimates in general due to the weighting which is applied in the estimation of volatilities and trends.<sup>19</sup>

For comparison, the plots also contain confidence intervals in case a random walk with drift is applied to simulate  $\kappa_{t,p}^{(1)}$ .<sup>20</sup> However, the resulting confidence intervals are much too narrow to include the realized life expectancy evolutions, and the random walks' volatilities would have to be blown up enormously for that to happen. In the right panel, also the shape of the confidence bounds looks dubious. The parabolic shape does not allow for substantial changes in the slope of linear mortality/life expectancy trends. This again questions the adequacy of the random walk with drift in describing the general level of mortality (see also Section 2.3.1). The stochastic trend approach, on the other hand, yields rather triangular confidence bounds which fit the historical data much better.

After specifying the volatility add-ons and the weighting parameters, we have got a fully calibrated multi-population mortality trend model at hand. In order to check the model's validity and adequacy, we have run extensive back tests and

have benchmarked our models to other mortality models. However, we refrain from giving more results on these tests here than those provided in Figure 6. Instead, we want to illustrate the applicability of our model by some numerical examples in the following section.

#### 4. NUMERICAL COMPARISON WITH THE SOLVENCY II STANDARD FORMULA

We now apply our model for some illustrative calculations of the Solvency Capital Requirement (SCR) under Solvency II. In principle, the SCR is defined as the 99.5% VaR of the Basic Own Funds over one year, where the Basic Own Funds are equal to the difference between the market value of assets and the Best Estimate Liabilities (BEL). Loosely speaking, one could also say it is the capital required to cover losses which may occur over the following year with at least 99.5% probability. For simplicity, we assume that assets are fixed in the following. Then, we can determine the SCR as the 99.5th percentile of (a) the actual payments to policy holders during the one year plus (b) the difference between year-end BEL and its expected value at the beginning of the year, all discounted back to the beginning of the year.

In the Solvency II standard formula, the 99.5% VaR is approximated by deterministic mortality/longevity shocks. For mortality prone contracts, the underlying mortality rates for all ages are increased by 15%, and the resulting change in Basic Own Funds (or BEL in case assets are fixed) constitutes the SCR. In the case of longevity, the mortality rates are reduced by 20%, with the change in Basic Own Funds again being the SCR. The SCRs for mortality- and longevity-linked contracts/portfolios are then aggregated by a variance-covariance approach assuming a correlation of  $-0.25$ . The shocks as well as the variance-covariance approach have come under some criticism for yielding implausible or even unrealistic results (see, e.g., Doff, 2008; Devineau and Loisel, 2009; Börger, 2010). Nevertheless, the standard formula is the benchmark for any internal solvency model. Therefore, we illustrate in the following how our model performs compared to the Solvency II standard formula, both in terms of capital required and diversification between mortality and longevity risk.

We consider both a multi-population and a single population model variant for UK males with weighting parameters and volatility add-ons as calibrated in the previous section, i.e.  $h = 5$ ,  $h_\sigma = 30$ ,  $\bar{\sigma}^{(1)} = 0.02$  and  $\bar{\sigma}^{(2)} = 0.0001$  for the multi-population model and  $h = 5$ ,  $h_\sigma = 30$ ,  $\bar{\sigma}^{(1)} = 0.05$  and  $\bar{\sigma}^{(2)} = 0.0005$  for the single population model.<sup>21</sup> We apply both model variants to quantify the mortality/longevity risk in two types of standard insurance contracts: a term insurance with terminal age 65 and a life annuity. Figure 8 shows the SCRs for such contracts with different initial ages, relative to the BEL (a constant interest rate of 2% is assumed for discounting).<sup>22</sup>

We observe that, in relation to the BEL, the multi-population model yields significantly smaller SCRs than the standard formula for both products and all initial ages. Part of this difference in SCRs may be credited to an implicit

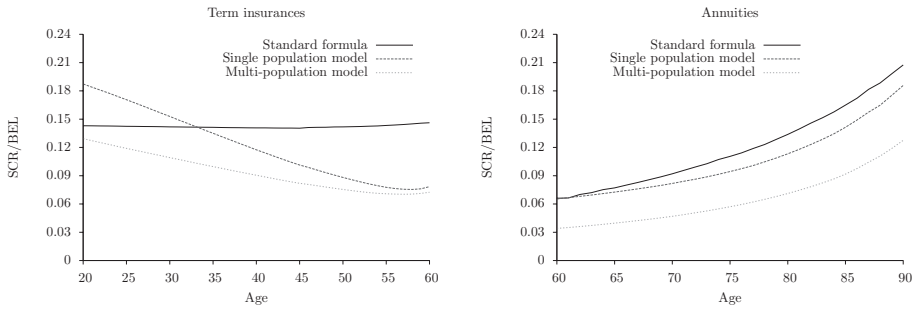


FIGURE 8: SCRs relative to BEL for UK males based on the values for  $h$ ,  $h_{\sigma}$ ,  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$  derived in Section 3, excluding contributions from random risk, parameter risk and threat scenarios.

margin for parameter uncertainty in the standard formula but nevertheless, the standard formula seems to overstate mortality/longevity risk.<sup>23</sup> When comparing the multi-population model’s results to those from the single population model, we find a positive effect from pooling data from several populations; overall trend uncertainty is reduced. Thus, running a multi-population model can obviously be beneficial even if one is interested in the mortality evolution of a single population only.

A multi-population model does not necessarily yield smaller risk capital charges in any case. This depends on the specific combination of populations in particular. However, in general, the pooling of mortality data and thus mortality information from several populations has a risk mitigating effect. The question which populations should be considered is difficult to answer and certainly depends on the application in view. We agree with Jarner and Kryger (2011) that there must be large similarities between the populations considered, in particular with respect to their presumed future mortality evolutions. If we added a population which exhibits significantly different mortality patterns, we would not be surprised to observe an increase in overall uncertainty due to the implied inhomogeneity in our set of populations. Moreover, data for all considered populations have to be credible.

Another interesting observation can be made from Figure 8 with respect to the evolution of the SCR/BEL ratios with initial age. While the shapes of all three curves look very similar for the life annuity, this is not the case for the term insurance. According to the standard formula, the SCR/BEL ratios are rather constant, whereas for the single and the multi-population model, the SCR/BEL ratios significantly depend on the initial age. For ages up to 35, the single population model even yields larger SCRs than the standard formula. The reason for the observed variation with age is the following: The large SCR/BEL ratios for young ages are due to the rather long-term risk as we always consider a terminal age of 65. Here, a change in the long-term mortality trend can have a substantial effect. This is a feature the standard formula with its constant one-off increase in mortality rates for all ages and maturities does not account for. With increasing initial age, the SCR/BEL ratios then decrease because of shrinking maturities.

However, when we get to ages above 50, the BEL start to decrease even stronger than the SCRs, resulting in increasing SCR/BEL ratios.

Finally, we want to illustrate the diversification properties of our model by comparing the SCRs for some insurance portfolios to those based on the variance–covariance approach in the standard formula. Our sample portfolios consist of term insurance contracts for initial ages 30, 40, 50 and 60 (with a death benefit of £100,000 for each contract), of annuity contracts for initial ages 60, 70, 80 and 90 (with a yearly annuity amount of £2,000 in each case), and a combination of both. The benefit amounts are chosen such that the SCRs for the term insurances and the annuities are roughly the same. The diversification benefit, i.e. the reduction of the combined SCR relative to the sum of term insurance SCR and annuity SCR, is 39% for the standard formula, while it increases to 48% for the multi-population model. Note that our model is specified such that diversification between mortality and longevity risk is already limited, in particular due to the volatility add-on  $\bar{\sigma}^{(2)}$ . Therefore, the correlation assumption between mortality and longevity SCRs in the standard formula seems very prudent. A reduction relative to the current correlation of  $-0.25$  seems to be appropriate, allowing for stronger diversification between mortality and longevity risks. In fact, a correlation of  $-0.45$  would be required in the standard formula to obtain a capital relief which is equal to the 48% in the multi-population model.

## 5. PARAMETER SENSITIVITIES

The results of the SCR comparison in the previous section significantly depend on the calibration of the parameters  $h$ ,  $h_\sigma$ ,  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$ . Since a certain degree of subjectivity is involved in their calibration, in this section, we analyze how sensitive the SCRs are with respect to each of the four parameters in the multi-population model.

More precisely, we present two sets of sensitivity analyses. In the first set, we vary each parameter under the assumption that the other parameters remain unchanged. This shows the stand-alone impact of each parameter on the SCR. In the second set, we vary the percentiles at which we want to observe the extreme mortality events from Section 3. The specification of these percentiles significantly relies on expert judgment, and therefore, it is important to analyze how sensitive the SCRs are with respect to concrete percentile choices. Here, we need to take into account the interactions between  $\bar{\sigma}^{(1)}$ ,  $h$  and  $\bar{\sigma}^{(2)}$ . Altering one of the percentiles can and typically does change the calibration of all three parameters.

The standard parameter values are  $h = 5$ ,  $h_\sigma = 30$ ,  $\bar{\sigma}^{(1)} = 0.02$  and  $\bar{\sigma}^{(2)} = 0.0001$  (cf. Sections 3 and 4), and the alternative scenarios for the first set of sensitivity analyses are as follows:<sup>24</sup>

TABLE 2  
SENSITIVITIES OF SCRS WITH RESPECT TO THE DESCRIBED PARAMETER CHANGES.

	Mortality	Longevity	Mixed
$h = 4$	+14.6%	+14.5%	+0.1%
$h = 6$	-8.9%	-8.7%	-0.0%
$h_\sigma = 10$	-8.9%	-7.1%	-0.0%
$h_\sigma = \infty$	+6.2%	+5.1%	+0.0%
$\bar{\sigma}^{(1)} = 0.016$	-7.7%	-6.2%	-0.0%
$\bar{\sigma}^{(1)} = 0.031$	+21.7%	+18.1%	+0.0%
$\bar{\sigma}^{(2)} = 0.00005$	-0.1%	-2.1%	-2.8%
$\bar{\sigma}^{(2)} = 0.00015$	+0.1%	+2.1%	+2.8%

- $h = 4$  and  $h = 6$ ;
- $h_\sigma = 10$  (more weight on most recent data points) and  $h_\sigma = \infty$  (no weighting);
- $\bar{\sigma}^{(1)} = 0.016$  and  $\bar{\sigma}^{(1)} = 0.031$ ;
- $\bar{\sigma}^{(2)} = 0.00005$  and  $\bar{\sigma}^{(2)} = 0.00015$ , i.e. a 50% decrease and increase of the add-on.

Table 2 shows changes in the SCRs for the different example portfolios which stem from changes in one of the parameters. Some of the SCR changes are fairly large; however, we regard them as moderate given the fairly substantial parameter shifts which are applied. We observe that  $h$ ,  $h_\sigma$  and  $\bar{\sigma}^{(1)}$  only have a significant impact for the “single risk portfolios”, while  $\bar{\sigma}^{(2)}$  mostly affects the SCR in the mixed portfolio case. The reason is that the first three parameters control the general level/trend of mortality independent of age. Changes in these parameters affect the severity of scenarios which are critical with respect to either mortality or longevity risk but not the severity of combined mortality/longevity scenarios.<sup>25</sup> The latter scenarios are relevant for the well-balanced mixed portfolio, and their severity is particularly controlled by the add-on  $\bar{\sigma}^{(2)}$ . For instance, an increase in  $\bar{\sigma}^{(2)}$  implies less diversification and, as we see, a larger SCR for the mixed portfolio. However, the SCR changes induced by the variation of  $\bar{\sigma}^{(2)}$  appear rather small.

When assessing the impact of the parameters  $h$ ,  $h_\sigma$  and  $\bar{\sigma}^{(1)}$ , it is important to keep in mind that the parameter variations are not necessarily comparable. Thus, it is difficult to state, at this stage, which parameter is the most influential one. Moreover, the impact of  $h_\sigma$  is particularly large for UK males. In the UK, the volatility in the  $\kappa_t^{(1)}$  has shrunk considerably over the last decades, and therefore, volatility estimates vary significantly for different values of  $h_\sigma$ . For other populations, we find a significantly smaller impact of  $h_\sigma$  on the volatility estimates and thus also on SCRs. Therefore,  $h$  and  $\bar{\sigma}^{(1)}$  seem to have the biggest impact on the SCRs.

Next, we vary the percentiles along which we want to observe the extreme historical mortality events. Here, we take interactions between  $\bar{\sigma}^{(1)}$ ,  $h$  and  $\bar{\sigma}^{(2)}$  into account. Altering the percentile for the short-term event, i.e. the life

TABLE 3

PARAMETER CALIBRATIONS AND SENSITIVITIES OF SCRS IN CASE OF VARIATION OF PERCENTILES FOR EXTREME MORTALITY EVENTS.

	$h$	$\bar{\sigma}^{(1)}$	$\bar{\sigma}^{(2)}$	Mort.	Long.	Mixed
Short-term event at 95th percentile	6.5	0.031	0.0001	+2.7%	+3.7%	-0.0%
Short-term event at 99th percentile	4.6	0.016	0.0001	-3.6%	-2.1%	-0.0%
Trend change event at 90th percentile	4.0	0.018	0.0001	+12.8%	+13.2%	+0.1%
Trend change event at 97th percentile	6.0	0.023	0.0001	-6.9%	-7.5%	+0.0%

expectancy drop at the beginning of the 1990s, obviously impacts the calibration of  $\bar{\sigma}^{(1)}$ . However, as part of the iterative calibration, also  $h$  and  $\bar{\sigma}^{(2)}$  need to be adjusted such that the trend change event is still observed at the desired percentile and the correlation at the boundaries of the age range is still close to zero. Similarly, altering the percentile for the trend change event does not only affect the calibration of  $h$  but also those of  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$ . The alternative percentiles are:

- Short-term event: 95th and 99th percentile instead of the 98th percentile;
- Trend change event: 90th and 97th percentile instead of the 95th percentile.

Table 3 shows the parameter calibrations for the altered percentiles as well as changes in the SCRs for the different portfolios. Both parameters  $\bar{\sigma}^{(1)}$  and  $h$  vary significantly for each of the four percentile variations, and we clearly see their interaction. For instance, an increase in short-term volatility (increase in  $\bar{\sigma}^{(1)}$ ) implies that less weight is to be assigned to the new data point in the trend re-calibration (increase in  $h$ ). The add-on  $\bar{\sigma}^{(2)}$ , on the other hand, is equal for all four percentile variations (after rounding). Apparently, changes in  $\bar{\sigma}^{(1)}$  and  $h$  hardly affect the correlation between ages at the boundaries of the age range, and thus,  $\bar{\sigma}^{(2)}$  needs to be adjusted only very slightly.

The SCR changes in Table 3 are rather small in case the percentile for the short-term event is altered. This holds in particular when comparing the SCR changes with those in Table 2 where the same values for  $\bar{\sigma}^{(1)}$  are considered. The effect of a change in  $\bar{\sigma}^{(1)}$  is, to a large extent, counterbalanced by the re-calibration of the weighting parameter  $h$ . Thus, the concrete percentile for the short-term event only seems to have a small impact on the SCRs. Over a one-year time horizon, mortality/longevity risk seems to be mostly driven by long-term trend uncertainty. This observation is confirmed by the SCR changes in case the percentile for the trend change event is altered. The resulting SCR changes are only slightly smaller than those in Table 2 (in absolute value). Therefore, the weighting parameter  $h$  clearly is the most crucial parameter in our model.

## 6. MORTALITY/LONGEVITY THREAT SCENARIOS

For the calculation of risk measures such as the 99.5% VaR or the 99% expected shortfall, we are particularly interested in tail mortality/longevity scenarios.

However, such scenarios are naturally quite rare in the historical mortality data, and thus, a fair amount of uncertainty remains whether tail scenarios generated by a mortality model are adequate and sufficiently severe. Epidemiological and demographic expert judgment could be used to check the appropriateness of extreme model outcomes and — if necessary — to modify the model. Such expert judgment can be implemented as unlikely but possible severe scenarios which we refer to as mortality/longevity threat scenarios.

Examples for such threat scenarios are the mortality/longevity stress scenarios in the Solvency II standard formula which we introduced in Section 4 — even though their adequacy has been questioned (see, e.g., Börger, 2010). Potentially more suitable threat scenarios could, e.g., be derived by shocking an existing mortality projection, i.e. by assuming 0.5% or 1% stronger or weaker annual mortality improvements than in the original projection. Alternatively, a longevity threat scenario could reflect the likely effects on the future mortality evolution of finding a cure for certain diseases like cancer. Similarly, a mortality threat scenario could describe possible effects of a significant increase in obesity and/or diabetes. Moreover, threat scenarios could be used to allow for very unlikely scenarios which the statistical model cannot generate. Given the historical mortality evolution, it seems reasonable to assume a global long-term mortality trend. However, although fairly unlikely from today's point of view, there might be diverging trends for certain populations in the future. Such diverging trends could also be covered by a threat scenario.

The application of mortality/longevity threat scenarios can be twofold. First, they can be used to check and possibly refine the model calibration. If the scenarios generated by the statistical model should cover a threat scenario in terms of severity but do not, some model parameters may have to be adjusted, e.g. the volatility add-on  $\bar{\sigma}^{(1)}$  or the weighting parameter  $h$ . Secondly, threat scenarios can be used to enrich the outcomes of the statistical model. Capital requirements can be calculated as a weighted average of the requirements based on threat scenarios and those stemming from the statistical model.

## 7. SUMMARY OF PARAMETER CALIBRATION

The model proposed in this paper contains a significant number of different parameters. This is due to the numerous requirements it has to fulfill: multi-population modeling, risk quantification over pre-defined time horizons, coverage of the full age range, conservative calibration, etc. In this section, we provide an overview of all parameters and summarize their meaning and estimation.

The parameters which need to be determined first are  $x_{\text{young}}$ ,  $x_{\text{center}}$  and  $x_{\text{old}}$ . They divide the age range roughly into young, middle and old ages and thus define the meaning of the  $\kappa_t^{(i)}$  processes. We set these parameters to 55, 60 and 85, respectively, which seems to be a reasonable choice for the populations under consideration in this paper. In general, they could also be estimated from data in

combination with the  $\kappa_t^{(\cdot)}$ . However, this is only a feasible approach in a single population model. In a multi-population framework, it is important that the parameters are equal for all populations. Otherwise, the  $\kappa_t^{(\cdot)}$  processes would not be comparable between populations.

With  $x_{\text{young}}$ ,  $x_{\text{center}}$  and  $x_{\text{old}}$  fixed, the  $\alpha_x$ , the  $\kappa_t^{(\cdot)}$  processes and the cohort parameters  $\gamma_{t-x}$  can be estimated from the data for each individual population as well as for the total population. We propose a four-step approach for this estimation in Section 2.2. The resulting parameter values can then be used to determine volatilities, correlations and starting values for the projection of the  $\kappa_t^{(\cdot)}$  processes and the cohort parameters.

In order to estimate the volatilities of the  $\kappa_t^{(\cdot)}$  processes, we also need the weighting parameter  $h_\sigma$ . The smaller  $h_\sigma$ , the more influence most recent data points have on the volatility estimates. In general, we propose using a rather large value for  $h_\sigma$  in order to have a stable model calibration over time. For projecting  $\kappa_{t,\text{total}}^{(1)}$  for the total population, we also require the weighting parameter  $h$  which determines how sensitive trend estimates are to most recent data points compared to data points farther in the past. As we saw in Section 5, this is the most important parameter in our model. We propose calibrating it to extreme trend changes in the historical mortality evolution.

The projection of the  $\kappa_{t,p}^{(1)}$  for the individual populations requires projection of the AR(1) difference processes with parameters  $a_p$ ,  $b_p$  and  $\sigma_p^{(1)}$  as well as correlations between the innovations for different populations. The parameters  $b_p$  can be estimated from the historical differences  $\kappa_{t,p}^{(1)} - \kappa_{t,\text{total}}^{(1)}$  under the constraints that  $1 > b_p \geq 0$ . The parameters  $a_p$  determine the long-term differences between the mortality trends for the total population and the individual populations. There are several different ways how these parameters can be estimated, and we show in Section 2.3.1 that most suitable approaches vary between populations.

In order to ensure a reasonably conservative model calibration, we introduced the volatility add-ons  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$ . The add-on  $\bar{\sigma}^{(1)}$  determines the severity of short-term mortality events. Its impact on capital requirements is moderate though as long as  $h$  is adjusted to the choice of  $\bar{\sigma}^{(1)}$  (cf. Section 5). As for  $h$ , we show how the add-on can be calibrated to severe historical mortality events. The add-on  $\bar{\sigma}^{(2)}$  impacts the diversification between mortality and longevity risk. We propose determining it such that correlation between mortality rates for very young and very old ages is close to zero. This is, from our point of view, the most conservative of all biologically reasonable assumptions.

The calibration of the four parameters  $h$ ,  $h_\sigma$ ,  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$  involves a certain degree of subjectivity. Moreover, the parameters interact, and we therefore introduce an iterative calibration procedure for  $h$ ,  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$  in Section 3. As part of this calibration procedure, we make several simplifying assumptions. Therefore, our approach should be seen as a first step to estimating model parameters from severe historical mortality events. In future work, we intend to establish



more sophisticated methods, e.g. based on extreme value theory. Then, it might even be possible to estimate all model parameters simultaneously. Possible refinements of these parameters based on mortality/longevity threat scenarios are discussed in Section 6.

## 8. REMAINING RISKS

No matter how sophisticated a mortality model is, there always remain some uncertainties which can roughly be divided into two classes: uncertainties linked to the mortality model itself and uncertainties which arise from a particular application of the model. In this section, we summarize these uncertainties for our modeling framework and discuss their magnitude and relevance.

The most relevant uncertainty in our framework is the parameter uncertainty contained in the weighting parameter  $h$ . This parameter determines the magnitude of potential changes in the long-term mortality trend within one year. Therefore, special attention has to be paid to the calibration of this parameter. Less crucial but still important are the volatility add-ons  $\bar{\sigma}^{(1)}$  and  $\bar{\sigma}^{(2)}$ . Also these parameters have to be chosen carefully. In comparison, we regard the uncertainty surrounding all other parameters as less critical.

One of our core assumptions is that the simulated realization of  $\kappa_{t,total}^{(1)}$  determines the changes in the expected future mortality trend. Even though this assumption is in line with actuarial practice where mortality projections are typically derived as reasonable extrapolations of trends in the most recent historical data, it can still be counterintuitive sometimes. For instance, if a cure for cancer was found in a year with a strong flu wave, the long-term mortality trend and the realized mortality rates should move in opposite directions. However, this is a general issue for any extrapolative mortality model.

On the application side, the modeling of dependencies is a source of uncertainty. As indicated in the section on threat scenarios, there might be some persistent divergence in mortality trends for different populations which our model does not allow for. If we consider mortality or longevity risk only, the assumption of a global long-term mortality trend is conservative. However, if we mix mortality and longevity risk for different populations, diverging trends could be critical. In case of a mixed portfolio, a threat scenario could be used to enrich the model outcomes and to thus ensure reasonably conservative results.

Basis risk may be seen in the use of mortality data for the general population when assessing mortality or longevity risk in a portfolio of insured. Even if the model perfectly described the future mortality evolution of the general population, the resulting mortality deviations might not be applicable to a possibly very specific population of insured. However, we regard this risk as minor since the mortality trend of the insured, as a subpopulation of the general population, can hardly diverge from the trend of the general population until infinity (see also Cairns *et al.*, 2011).

Another issue arises from the delayed availability of data. In almost every country, mortality rates are currently only available with a time lag of at least one year. For the purpose of modeling only mortality deviations, we propose shifting the historical data up to the present year. While this obviously introduces a shift in the absolute values of simulated future mortality rates, the relative deviations remain mostly unaffected. Recall that the weighting parameter  $h_\sigma$  is supposed to be large such that a time lag of only a few years does not substantially change the measured volatilities. When modeling mortality rates, we propose filling up the missing years with best estimate data. Data sets to which the volatilities are calibrated obviously must not contain these best estimates.

Finally, the one-year time horizon under Solvency II and the SST implies the assumption that mortality trends and the resulting mortality tables are updated annually. In practice, this is usually not the case. Trend assumptions only get updated once several years of additional data clearly indicate the need for an update. The resulting adjustments are then typically larger than those we allow for in our model. However, we do not regard these adjustments as one-year changes but aggregated changes for several years which our model does not need to cover within one year. Instead, we see the whole issue as a shortcoming of out-of-date mortality tables, and hence, it should be accounted for by loadings in the mortality tables.

## 9. CONCLUSION

A comprehensive risk model for mortality and longevity meeting the Solvency II and SST criteria requires an adequate stochastic mortality model. In particular, the one-year time horizon and a focus on extreme mortality trend deviations deserve special attention. Moreover, an adequate stochastic model needs to be efficient and sufficiently simple to maintain.

In this paper, we specify a mortality model which meets these criteria and we show how it can be calibrated and applied. The model covers the full age range, thus making it suitable for determining both mortality and longevity risk capital charges. Furthermore, it includes cohort effects, and its five stochastic drivers offer a great variability in the mortality scenarios it generates. Moreover, all model parameters possess a clear interpretation. The dependency structure between the mortality rates for different ages is non-trivial such that mortality and longevity risk can be assessed simultaneously.

In order to be able to quantify risk over a one-year time horizon, we introduce a new stochastic process for the general level and trend of mortality, i.e.  $\kappa_t^{(1)}$  in our model. This process follows the actuarial practice and intuition of annually re-estimating the mortality trend based on the data available up to that respective point in time. It can be implemented as part of basically any extrapolative mortality model and it is not only applicable for the one-year view but also for run-off simulations. We show that, in the latter case, it offers significant advantages compared to the commonly used random walk with drift. In

particular, we are able to disentangle the magnitude of one-year random fluctuations from the long-term trend uncertainty, thus providing more realistic confidence bounds.

Furthermore, we establish our mortality model in a multi-population setting. Thus, we are able to simulate the future mortality evolution for several populations simultaneously and consistently. This is particularly relevant in practice as basically any insurance portfolio consists of at least two populations, i.e. males and females in the same country. Also for exploiting diversification effects, a multi-population setting is essential. Moreover, even if only one population is considered, it can be worthwhile to use a multi-population model. As we show, the information provided by the mortality data from other populations can significantly reduce the trend uncertainty and thus improve the assessment of mortality/longevity risk.

As we aim for a mortality model for solvency purposes, we pay special attention to a conservative specification and calibration of the model. In particular, we explain where margins should be included in the model, e.g. by volatility add-ons, and for which reason. We also show how these margins can be specified and calibrated to extreme mortality events in the historical data. Moreover, we explain how expert judgment in form of mortality/longevity threat scenarios can be used to check the plausibility of a mortality model and to enrich the outcomes it generates. In order to illustrate our model, we compare the resulting capital requirements under Solvency II to those from the Solvency II standard formula. We observe that our model may require significantly less capital and that the diversification assumed in the standard formula seems to be prudent.

## NOTES

1. For a detailed overview and a discussion of the Solvency II proposals, we refer to Eling *et al.* (2007), Steffen (2008), and Doff (2008). For the SST, we refer to the Swiss Federal Office of Private Insurance (2006) and Eling *et al.* (2008). A comparison of these solvency regimes also with other regimes can be found in Holz Müller (2009).

2. We obtain mortality rates for the total population as weighted averages of the mortality rates of the individual populations, with weights according to the population size.

3. The logit  $q_{x,t}$  have already been modeled by other authors as well, see, e.g., Cairns *et al.* (2006) whose model is a special case of ours.

4. We considered the 37 countries listed in the Human Mortality Database, <http://www.mortality.org>, as of 17/11/2010: Australia, Austria, Belarus, Belgium, Bulgaria, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Russia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Taiwan, UK, USA, Ukraine.

5. The expression  $a \leftarrow b$  differs from  $a = b$  in the sense that, e.g.,  $a \leftarrow a + 1$  is valid and increases  $a$  by 1.

6. Note that, by this approach, the mortality trend at a future point in time is derived from data which are simulated according to the trend estimated today. Thus, the future trend estimate depends on today's trend estimate which is not fully intuitive. However, this situation also occurs when re-estimating a full mortality model at a future point in time as part of nested simulations. We plan to discuss this issue in detail in future work.

7. The weights are specified such that

$$\sum_{t \leq t_{\max}} \frac{w_t(t_{\max} - t)}{\sum_{t' \leq t_{\max}} w_{t'}} = h,$$

i.e. the average number of considered data points is  $h$ . Or, to put it in another way, the data points up to  $h$  years into the past have a cumulative weight of about  $1 - 1/e$ , i.e. roughly  $2/3$ , while the remaining older data points only have a cumulative weight of about  $1/e$ , i.e. roughly  $1/3$ .

8. For the calculation of the sample standard deviation, the mean of the errors must not be estimated from the data but must be set equal to zero. Otherwise, the volatility  $\sigma^{(1)}$  would be systematically underestimated for a projection of future innovations with mean zero.

9. An alternative approach would be to use a mixture of two normal distributions with different variances for  $\varepsilon_t^{(1)}$ . For each year of simulation, a Bernoulli distributed random variable would indicate whether the add-on is to be applied or not.

10. The volatility  $\sigma^{(1)}$  is derived from the historical errors based on a weighting parameter  $h_\sigma = 30$  (cf. Section 3 for details on this parameter choice) and the volatility add-on  $\bar{\sigma}^{(1)}$  is set to zero.

11. To be more precise, equality between the two formulations only holds if a fixed number  $n$  of  $\kappa^{(1)}$  values are used to determine  $\ell_t$  for  $t > t_{\max}$  i.e.  $\kappa_{t-n+1}^{(1)}, \dots, \kappa_t^{(1)}$ . However, due to the weights applied in the fitting of  $\ell_t$ , this essentially holds for a sufficiently large value of  $n$ .

12. For more details on exponential smoothing, we refer to Gardner (2006) and references therein.

13. For this example, the drift is determined as the weighted average of the first differences in the historical  $\kappa_t^{(1)}$ , with weighting parameter  $h = 5$ . The volatility of the random walk is equal to the volatility in the other processes in Figure 3.

14. The calibration of the weighting parameters  $h$  and  $h_\sigma$  as well as the volatility add-on  $\bar{\sigma}^{(1)}$  for the total population is discussed in Section 3.

15. In fact, the uncertainty in the long-term differences could be accounted for by modeling them stochastically. However, a reasonable distribution for these differences is not obvious, and the additional modeling complexity would be enormous. We therefore do not pursue this idea.

16. Note that, in case of simultaneous estimation, the  $\gamma_{t-x}$  are set equal to zero for cohorts with current age below 45 or above 95. Due to insufficient data, the parameter estimates for these cohorts are very volatile, which leads to instability of the overall model estimation.

17. Following an analogous line of argument, the Russia scenario might even be seen as a 1000-year event since we have, on average, about 100 years of data for each population. However, for the sake of conservatism and due to the fact that the observed event may not be representative for the true distribution of the mortality evolution, we stick to the 500-year event.

18. The omitted male and female populations of Chile, Israel, Luxembourg, Slovenia and Taiwan are all rather small and thus do not have a significant impact on the total population.

19. For Russian males, the parameter values underlying the life expectancy simulation are  $h_\sigma = 30$ ,  $h = 4$ ,  $\sigma^{(1)} = 0.06$ ,  $\bar{\sigma}^{(1)} = 0.05$ ,  $\bar{\sigma}^{(2)} = 0.0005$ ,  $\sigma_\gamma = 0.048$  and  $0.0001 \begin{pmatrix} 0.02 & 0.00 & -0.01 \\ 0.00 & 0.02 & 0.0178 \\ -0.01 & 0.02 & 0.24 \end{pmatrix}$  as covariance matrix of  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$ . For Dutch males, the respective parameter values are  $h_\sigma = 30$ ,  $h = 4$ ,  $\sigma^{(1)} = 0.02$ ,  $\bar{\sigma}^{(1)} = 0.05$ ,  $\bar{\sigma}^{(2)} = 0.0005$ ,  $\sigma_\gamma = 0.01$  and  $0.0001 \begin{pmatrix} 0.05 & 0.07 & -0.02 \\ 0.07 & 0.14 & -0.07 \\ -0.02 & -0.07 & 0.27 \end{pmatrix}$  as covariance matrix of  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$ .

20. The random walks are fitted to the same data sets as the stochastic trend processes. The drift parameters are equal to 0.0011 for Russian males and 0.0114 for Dutch males, the respective volatilities are 0.07 and 0.019.

21. The other parameters are  $\sigma_{total}^{(1)} = 0.014$ ,  $a_p = 0.003$ ,  $b_p = 0.97$  and  $\sigma_p^{(1)} = 0.02$  in the multi-population model and  $\sigma^{(1)} = 0.01$  in the single population model. For both model variants, the covariance matrix for  $\varepsilon_t^{(2)}$ ,  $\varepsilon_t^{(3)}$  and  $\varepsilon_t^{(4)}$  is  $0.0001 \begin{pmatrix} 0.02 & 0.01 & 0.01 \\ 0.01 & 0.03 & 0.01 \\ 0.01 & 0.01 & 0.10 \end{pmatrix}$  and  $\sigma_\gamma = 0.026$ .

22. For these simple contracts, the BEL can be calculated directly from the best estimate mortality rates. For contracts which contain asymmetric structures such as option-type payoffs, the BEL can only be determined via nested simulations.

23. Note that the capital requirements which we state here for illustrative purposes do not contain any contributions from random risk, parameter risk or mortality/longevity threat scenarios (see Section 6).

24. The alternative scenarios for  $h$  and  $\bar{\sigma}^{(1)}$  are specified such that we can draw comparisons with the second set of sensitivity analyses later on.

25. Note that this should not be misinterpreted in the sense that scaling up volumes of both mortality and longevity portfolios by the same factor does not have an impact on the total SCR.

## REFERENCES

- BÖRGER, M. (2010) Deterministic shock vs. stochastic value-at-risk – an analysis of the Solvency II standard model approach to longevity risk. *Blätter der DGVMF*, **31**, 225–259.
- CAIRNS, A., BLAKE, D. and DOWD, K. (2006) A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, **73**, 687–718.
- CAIRNS, A., BLAKE, D. and DOWD, K. (2008) Modelling and management of mortality risk: A review. *Scandinavian Actuarial Journal*, **2**, 79–113.
- CAIRNS, A., BLAKE, D., DOWD, K., COUGHLAN, G. and KHALAF-ALLAH, M. (2011) Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin*, **41**, 29–59.
- COUGHLAN, G., KHALAF-ALLAH, M., YE, Y., KUMAR, S., CAIRNS, A., BLAKE, D. and DOWD, K. (2011) Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, **15**, 150–176.
- CURRIE, I., DURBAN, M. and EILERS, P. (2004) Smoothing and forecasting mortality rates. *Statistical Modelling*, **4**, 279–298.
- DEVINEAU, L. and LOISEL, S. (2009) Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula? *Bulletin Français d'Actuariat*, **18**, 107–145.
- DOFF, R. (2008) A critical analysis of the Solvency II proposals. *The Geneva Papers on Risk and Insurance Issues and Practice*, **33**, 193–206.
- DOWD, K., CAIRNS, A., BLAKE, D., COUGHLAN, G. and KHALAF-ALLAH, M. (2011) A gravity model of mortality rates for two related populations. *North American Actuarial Journal*, **15**, 334–356.
- ELING, M., GATZERT, N. and SCHMEISER, H. (2008) The Swiss Solvency Test and its market implications. *The Geneva Papers on Risk and Insurance Issues and Practice*, **33**, 418–439.
- ELING, M., SCHMEISER, H. and SCHMIT, J. (2007) The Solvency II process: Overview and critical analysis. *Risk Management and Insurance Review*, **10**, 69–85.
- ENGLE, R. and GRANGER, C. (1987) Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, **55**, 251–276.
- GARDNER, E. (2006) Exponential smoothing: The state of the art – Part II. *International Journal of Forecasting*, **22**, 637–666.
- HARI, N., DE WAEGENAERE, A., MELENBERG, B. and NIJMAN, T. (2008) Longevity risk in portfolios of pension annuities. *Insurance: Mathematics and Economics*, **42**, 505–519.
- HOLZMÜLLER, I. (2009) The United States RBC standards, Solvency II and the Swiss Solvency Test: A comparative assessment. *The Geneva Papers on Risk and Insurance Issues and Practice*, **34**, 56–77.
- HUNT, A. and BLAKE, D. (2013) *A general procedure for constructing mortality models*. Working Paper n. 1301, Pensions Institute.
- HYNDMAN, R., BOOTH, H. and YASMEEN, F. (2011) *Coherent mortality forecasting: The product-ratio method with functional time series models*. Working Paper n. 01/11, Monash University.
- JARNER, S. and KRYGER, E. (2011) Modelling adult mortality in small populations: The SAINT model. *ASTIN Bulletin*, **41**, 377–418.
- LEE, R. and MILLER, T. (2001) Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, **38**, 537–549.
- LI, J. and HARDY, M. (2011) Measuring basis risk in longevity hedges. *North American Actuarial Journal*, **15**, 177–200.

- LI, N. and LEE, R. (2005) Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, **42**, 575–594.
- OLIVIERI, A. (2011) Stochastic mortality: Experience-based modeling and application issues consistent with Solvency 2. *European Actuarial Journal*, **1**, 101–125.
- OLIVIERI, A. and PITACCO, E. (2008a) *Solvency requirements for life annuities: Some comparisons*. Working Paper, University of Parma and University of Trieste.
- OLIVIERI, A. and PITACCO, E. (2008b) *Stochastic mortality: The impact on target capital*. CARFIN Research Paper, University Bocconi.
- PLAT, R. (2009) On stochastic mortality modeling. *Insurance: Mathematics and Economics*, **45**, 393–404.
- PLAT, R. (2011) One-year value-at-risk for longevity and mortality. *Insurance: Mathematics and Economics*, **49**.
- RICHARDS, S., CURRIE, I. and RITCHIE, G. (2012) *A value-at-risk framework for longevity trend risk*. Working Paper, Longevitas Ltd, Edinburgh.
- STEFFEN, T. (2008) Solvency II and the work of CEIOPS. *The Geneva Papers on Risk and Insurance Issues and Practice*, **33**, 60–65.
- STEVENS, R., DE WAEGENAERE, A. and MELENBERG, B. (2010) *Calculating capital requirements for longevity risk in life insurance products. Using an internal model in line with Solvency II*. Working Paper, Tilburg University.
- SWISS FEDERAL OFFICE OF PRIVATE INSURANCE. (2006) Technical Document on the Swiss Solvency Test. Available at [www.finma.ch](http://www.finma.ch).

MATTHIAS BÖRGER (Corresponding author)  
*Institute of Insurance, Ulm University*  
& *Institute for Finance and Actuarial Sciences (ifa), Ulm*  
*Helmholtzstraße 22, 89081 Ulm, Germany*  
*Phone: +49 731 50 31257, Fax: +49 731 50 31239*  
*E-Mail: m.boerger@ifa-ulm.de*

DANIEL FLEISCHER  
*Swiss Reinsurance Company Ltd*  
*Mythenquai 50/60, 8022 Zurich, Switzerland*

NIKITA KUKSIN  
*Swiss Reinsurance Company Ltd*  
*Mythenquai 50/60, 8022 Zurich, Switzerland*