Modeling Mortality Trend under Solvency Regimes

Matthias Börger, Daniel Fleischer und Nikita Kuksin

Preprint Series: 2011-01

Fakultät für Mathematik und Wirtschaftswissenschaften

UNIVERSITÄT ULM

Modeling Mortality Trend under Modern Solvency Regimes

Matthias Börger*

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
Email: 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

February 11, 2011

PRELIMINARY VERSION

Abstract

New solvency regimes in Europe, Solvency II and the Swiss Solvency Test, necessitate the stochastic modeling of mortality/longevity trend risk. 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 mortality and longevity risk for several populations.

To account for the one-year time horizon of the solvency regimes, we propose a specification of the model parameters which implies stochasticity in the long-term mortality trend. This approach spares the common re-estimation of the mortality model at the end of the one-year time horizon and, at the same time, provides highly plausible run-off scenarios. Finally, we explain how expert judgment in form of mortality/longevity threat scenarios can be used to test and enrich a mortality model.

*Corresponding Author

1 Introduction

In Europe, new solvency regimes such as of Solvency II and the Swiss Solvency Test (SST) will be introduced in the near future or have already been introduced. Both regimes are conceptually similar and follow the common goal of a comprehensive modeling and assessment of the risks insurance companies are exposed to, both in terms of number of risk factors as well as severity of risks. 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 Holzmüller (2009).

In both solvency regimes, capital requirements can either be computed via a standard formula or an internal model. However, companies are encouraged to develop internal models which cover the company specific risk profile as closely as possible. In contrast to the scenario based standard formula, these internal models are generally stochastic and thus are often more complex than the standard formula. They do share however with the standard formula the split of the overall risk into several modules for which separate capital requirements are computed and finally aggregated. For a life insurance company, one of these modules is devoted to the treatment of mortality and/or longevity risk.

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 future mortality evolution. This risk is relevant for insurance products which pay upon death of the insured person as well as products which pay upon survival. 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, e.g. the risk arising from natural disasters, pandemics, etc. or random risk, i.e. the risk of random fluctuations in a small 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 the true 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 appropriate for the computation of capital requirements under Solvency II or the SST as certain features of these solvency regimes impose strong requirements on the models.

In this paper, we propose a mortality model which satisfies all these requirements. It can be applied in basically all relevant settings where the most complex one is generally seen as the simultaneous modeling of mortality and longevity risk for several different populations. We discuss all aspects of the model specification, calibration, and application and provide ideas and practical advice for the implementation of all three steps.

Our model covers the full range of ages prevalent in typical risk products and annuity/pension portfolios, and includes cohort effects. The fitting of the model via Generalized Linear Model (GLM) techniques is simple and fast. Moreover, the model structure is such that it permits the modeling of deviations from today's best estimate mortality rates. This overcomes the typical issue of insufficient data for the calibration of a stochastic mortality model for an insured portfolio. When modeling mortality deviations, the model can be fitted to general population data, which is generally sufficiently available, but still be applied to an insurance portfolio. Five stochastic drivers ensure a large variety in simulated mortality scenarios and, since they apply differently to different age groups, the correlation structure of the model is non-trivial. This is particularly relevant when mortality and longevity risk are considered simultaneously such that diversification effects between these risks can be exploited. The correlation in the mortality evolution for adjacent ages is typically very strong but for larger age differences, this is not necessarily true. Therefore, since mortality risk is usually relevant for younger ages and longevity risk is most significant for older ages, a perfect correlation implies an overestimation of the diversification between mortality and longevity risk.

Under Solvency II and the SST, mortality/longevity risk is to be quantified over a one-year time horizon, i.e. an insurance company needs to hold enough capital that economic losses which may occur over one year can be covered with sufficiently large probability. As Börger (2010) points out, this does not only affect the simulation of one-year-ahead mortality but also the reassessment of assumptions with respect to future expected mortality. Hence, a mortality model for solvency purposes does not only have to be specified such that it provides plausible realized mortality rates but also realistic changes in the long-term mortality trend assumption. A common approach to implementing this *one-year view* is to simulate next year's mortality rates and then to re-estimate the mortality model based on a data set which contains the historical data and the newly generated data point. This approach is plausible in that an update of the mortality trend assumption must rely on the new information obtained in the current year. The application of this approach, however, can become extremely time consuming, in particular when several populations are considered, as the model parameters have to be re-estimated in each simulation path.

In order to avoid these re-estimations, we propose and implement stochastic processes in our model which do not only provide plausible mortality rates for the current year but, at the same time, imply the corresponding changes in the (long-term) mortality trend. In other words, the long-term mortality trend is stochastic. Thus, no re-estimation of the model is required anymore which is particularly helpful when consistent mortality scenarios for several populations are required. Moreover, our approach implies consistency between the *one-year view* and the *run-off view* as the model's run-off view is obtained via iterative one-year simulations. When re-estimating model parameters each year, the resulting iterative run-off does generally not coincide with any possible run-off simulation without re-estimation, which is counterintuitive. Another appealing feature of our choice of the stochastic processes is best estimate consistency. This means that the mortality trend assumption does not change if next year's realized mortality rates coincide with their best estimates today. In case of the re-estimation approach, this consistency is usually not fully guaranteed. Even though we only implement our "implicit re-estimation approach" in the model setting of this paper, it can be applied to basically any mortality model which includes a linear trend assumption, e.g. in form of a random walk with drift.

An issue which needs careful consideration in any re-estimation approach is the allowance for reasonably significant trend changes. For instance, when a (linear) trend is fitted to, say, 50 data points the addition of another data point – even though it may be quite an extreme one – hardly impacts the trend. The equal weight assigned to each data point is critical in this case. Figure 1 illustrates this point for the (partially linear) trend in *life expectancy*, i.e. period life expectancy at birth. For all countries, more or less obvious breaks in the trend can be observed. This underlines the necessity of weighting data points in order to capture the most recent trend slope adequately when extrapolating trends into the future. Therefore, we include weights in the fitting of trends which fade out exponentially going backwards in time.

In practice, the simultaneous modeling of mortality for several populations is inevitable as basically any portfolio of contracts consists of at least two populations: males and females of the same geographical region/country.

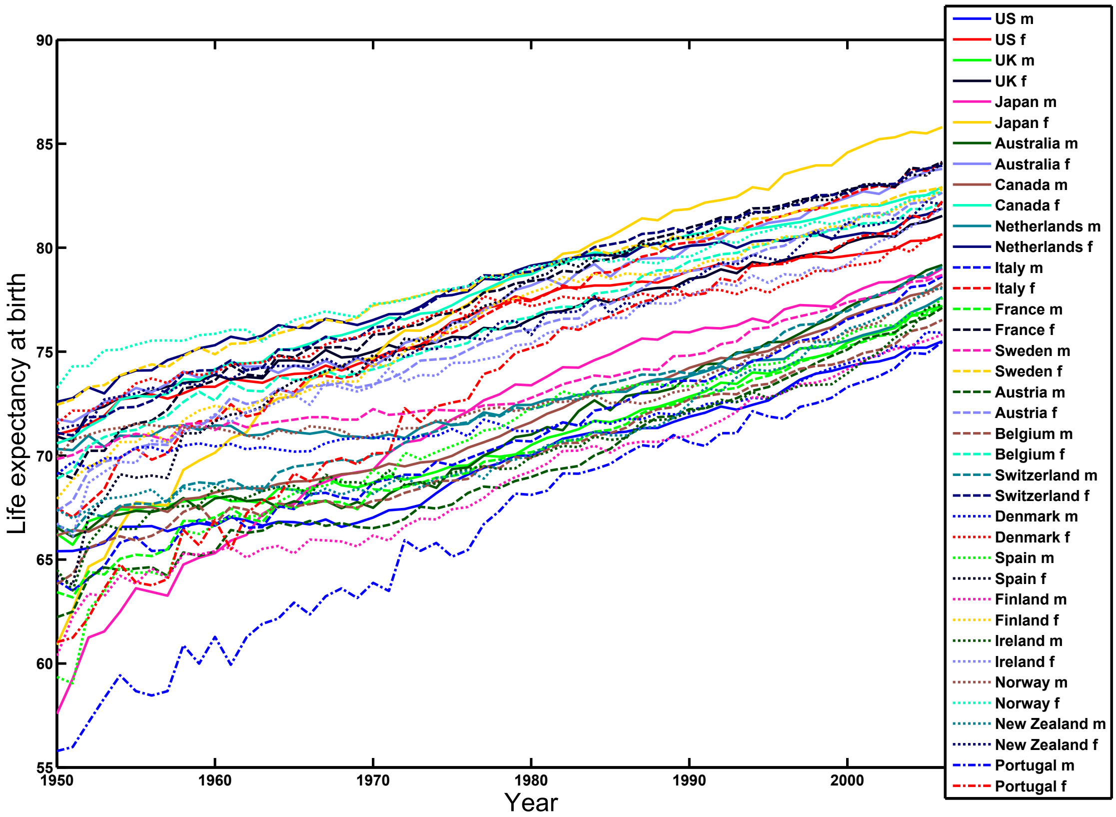

Figure 1: Period life expectancies at birth of various populations

Often, also portfolios from different countries are combined which automatically adds another two populations for each additional country. Therefore, after a full description of our one-population model, we extend it to a multi-population model. We ensure consistency between the mortality evolutions of different population by modeling a common stochastic (long-term) mortality trend for all populations. Differences between the common trend and the population specific mortality evolutions are described by additional stationary and correlated processes.

Under Solvency II, risk is specified as the 99.5% Value-at-Risk and in the SST, it is the 99% Expected Shortfall. Thus, in both cases only the most extreme model outcomes are particularly relevant and therefore a mortality model has to be specified with a focus on tail events. This refers to the choice of the stochastic processes and the estimation of their parameters in particular. Stochastic models, which are based on historical data, are only extrapolations of the historical evolution. This has two implications for mortality models: When a model is fitted to a (rather short) period of low volatility without extreme events, the variation in the simulation outcomes generated by the model will be rather small, too. We tackle this issue by introducing a volatility add-on which implies a minimum volatility stochastic processes cannot undercut. We also show how volatility add-ons can be calibrated in practice. The second implication of the extrapolative nature of a mortality model is that, loosely speaking, mortality models only allow for structural changes in (best estimate) mortality which have been observed in the past. They do usually not allow for changes due to unprecedented events like, e.g. the possible finding of a cure for cancer. However, such an event may well become reality hence representing a major risk for longevity business. We show how such events can – based on expert judgment – be incorporated in a mortality model via so-called mortality/longevity threat scenarios. Such scenarios can also be used to check the plausibility of a model and/or to adjust the range of simulation outcomes.

The stochastic modeling of mortality for the computation of solvency capital requirements has already been discussed by other authors. For instance, Plat (2010) proposes a stochastic model for mortality reduction factors and thus takes a different approach to updating the best estimate assumption at the end of the one-year time horizon. However, he pays less attention to the plausibility of the realized mortality rates, in particular in a model run-off. Stevens et al. (2010) derive closed form approximations for the capital requirement under Solvency II in order to avoid the re-estimation of the mortality model. But their approximations are based on the model of Lee and Carter (1992) which is in many cases too simple to fulfill the needs of a mortality risk model for Solvency II or SST purposes, e.g. its correlation structure is trivial. Hari et al. (2008) and Olivieri (2009) use more sophisticated mortality models but focus on longevity risk and the impact of systematic (trend) and non-systematic (random) risk. Moreover, Olivieri (2009) 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 model. Moreover, all these papers allude only briefly to the simultaneous modeling for several populations or the inclusion of expert judgment.

The remainder of this paper is organized as follows: In the following section, we introduce our mortality model for one population. We describe the model structure, the fitting process and the simulation of future mortality scenarios. Implementations of the respective algorithms are provided in the Appendix. Moreover, we go into the caveats of calibrating a mortality model for solvency purposes in detail. In Section 3, we then extend our model to a multi-population model. Besides the derivation of a stochastic common trend, this includes a discussion of how differences between different populations can be modeled. In the subsequent section, we introduce the notion of mortality/longevity threat scenarios. Such scenarios can be used to incorporate further judgment, e.g. on unprecedented events, into a mortality model and its outcomes. Section 5 contains a detailed discussion and assessment of all risks inherent in our modeling framework. Finally, Section 6 concludes.

2 Mortality Trend Model

In this section, we introduce a mortality (trend) model which is an extension of the model proposed by Plat (2009). The extension is necessary for two reasons: modeling of (extreme) long-term trend deviations and modeling under the one-year view. Both require changes to the underlying parameter processes. After a comment on data issues, we show how the model can be calibrated and used for simulating future mortality evolutions. Algorithms for the model fitting and simulation are provided in the Appendix.

2.1 Data Availability and Modeling of Deviations

In view of simulation, a mortality model should in principle be calibrated to the historical mortality evolution of the portfolio-specific population. However, this is not always possible. A population of insured is often not large enough to allow for a reliable calibration of a mortality model. Moreover, the data history may not be sufficiently long or data may not yet be available for ages for which simulation outcomes are required later on, e.g. old ages when the portfolio currently only consists of annuities in their deferment period. A population for which sufficient data is typically available is the general population of a country. However, here one is often faced with a strong bias, i.e. the level of mortality rates in the general population may differ significantly from the level observed in the population of insured.

Nevertheless, the mortality evolution of the general population is still useful for generating mortality scenarios for a population of insured. Even though the mortality rate levels may differ, their variation is likely to be similar as one population is a subpopulation of the other.[1] Thus, it is ade-

[1] Note again that we disregard any random or small sample risk in the insurance portfolio.

quate to calibrate a mortality model to the historical mortality evolution of the general population and to adjust the level of mortality rates thereafter. This can be accomplished by simulating mortality deviations, i.e. deviations of the mortality rates from their best estimates.

To be more specific, let $\hat{q}_{x,t}$ be the best estimate mortality rates according to our mortality model for the general population for all ages x and future years t. For each simulation/scenario s, a realization of future mortality rates $q_{x,t,s}$ is drawn, and (relative) deviations from the best estimate mortality rates are then computed as

$$r_{x,t,s} = \frac{q_{x,t,s}}{\hat{q}_{x,t}} \ . \tag{1}$$

By multiplying these deviations by (a set of) best estimate mortality rates for the population of insured, $\hat{q}_{x,t}^{(\text{insured})}$, mortality scenarios for this population can finally be obtained.

2.2 Model Structure

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

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

where $\operatorname{logit}(\cdot) = \ln(\cdot/(1-\cdot))$ is the logit function and $x^+ = \max\{x, 0\}$. In this specification, the process $\kappa_t^{(1)}$ describes the general tendency of the mortality evolution independent of age. The term $\kappa_t^{(2)}(x - x_{\text{center}})$ reflects the "slope" of the logit $q_{x,t}$ or the mortality steepness and the parameter x_{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 $x_{\min} = 20$ and $x_{\max} = 105$.[2] The terms $\kappa_t^{(3)}(x_{\text{young}} - x)^+$ and $\kappa_t^{(4)}(x - x_{\text{old}})^+$ account for a larger volatility typically observed in the mortality rates of young and old ages where the number of deaths $d_{x,t}$ is low compared to ages around x_{center}. Moreover, the terms allow for flexibility regarding the correlation between the mortality

[2] Note that this choice does not imply that only rotations around $x_{\text{center}} = 60$ are simulated but the combination of $\kappa^{(1)}$ and $\kappa^{(2)}$, in principle, allows a rotation around every age x.

evolutions for different ages. We set $x_{\text{young}} = 55$ and $x_{\text{old}} = 85$.[3] Possible cohort effects are represented by the process γ_{t-x} and, finally, the parameters α_x account for age-dependent differences in mortality which are not covered by the other components of Equation (2).

Our model structure is fairly close to the one of 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 do this because, firstly, the $q_{x,t}$ are the quantities actually required for the computation of cashflows and secondly, because the logit function is the canonical link function in a generalized linear model with binomial distributions. The latter is very practical for parameter estimation of the model. Even though the modeling of the log $m_{x,t}$ appears to be more common in the academic literature (see, e.g. the vast literature on the model of Lee and Carter (1992)), the logit $q_{x,t}$ have already been modeled by other authors as well, see, e.g. Cairns et al. (2006). A term of Equation (2) which is not incorporated in the model of Plat (2009) is the term for old ages. We have conducted analyses of various populations[4] and have found this parameter to be significant. Without this term the volatility of old ages, where very little data is available, would be underestimated.

2.3 Parameter Estimation

In the following, we describe how all parameters are estimated. For the remainder of this paper let historical mortality rates $q_{x,t}$ of a given population be available for ages $x = x_{\min}, \ldots, x_{\max}$ and years $t = t_{\min}, \ldots, t_{\max}$. Instead of estimating α_x together with the remaining parameters we set

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

[3] Obviously, for a different age range or for some specific population it might be reasonable to use other values for the parameters x_{center}, x_{young}, and x_{old}. However, the concept of separating the ages into young, medium and old ages should be preserved.

[4] We considered the 37 countries available at 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, US, Ukraine.

i.e. the mean of the logit $q_{x,t}$ over time. This is mainly done for reasons of parameter uniqueness and computation time.

For the cohort parameters γ_{t-x} we neglect the $c_{\text{excl}} = 10$ oldest cohorts as well as the youngest cohorts whose present age is below $x_{\text{cutoff}} = 45$ since there is too little historical data available to reasonably estimate these parameters.[5] Together with the κ_t-parameters we can set up a predictor matrix M that has full rank by construction, and formulate the corresponding generalized linear model with logit link function and binomial distributions. This can be solved by iteratively reweighted least squares, see Nelder and Wedderburn (1972).[6] Binomial distributions obviously require the number of deaths $d_{x,t}$ and lives $\ell_{x,t}$ instead of mortality rates $q_{x,t}$. The reconciliation takes actual demographic data of the given population (to be used as actual $\ell_{x,t}$) into account.

The underlying equation of the generalized linear model is

$$\begin{pmatrix} \operatorname{logit}(q_{x_{\min},t_{\min}}) - \alpha_{x_{\min}} \\ \operatorname{logit}(q_{x_{\min}+1,t_{\min}}) - \alpha_{x_{\min}+1} \\ \vdots \\ \operatorname{logit}(q_{x_{\max},t_{\min}}) - \alpha_{x_{\max}} \\ \operatorname{logit}(q_{x_{\min},t_{\min}+1}) - \alpha_{x_{\min}} \\ \vdots \\ \operatorname{logit}(q_{x_{\max},t_{\max}}) - \alpha_{x_{\max}} \end{pmatrix} \approx M \begin{pmatrix} \kappa^{(1)}_{t_{\min}} \\ \kappa^{(1)}_{t_{\min}+1} \\ \vdots \\ \kappa^{(4)}_{t_{\max}} \\ \gamma_{t_{\min}-x_{\max}+c_{\text{excl}}} \\ \vdots \\ \gamma_{t_{\max}-x_{\text{cutoff}}} \end{pmatrix}, \tag{3}$$

where M is the coefficient matrix induced by Equation (2).

Basically, the fitting process could be stopped at this stage. However, we apply the following operations to the parameters. Note that these operations do not modify the response, i.e. the resulting logit $q_{x,t}$. The idea is to have parameters $\kappa^{(1)}$ and $\kappa^{(2)}$ that possess a clear interpretation and are thus more handy to monitor, e.g. (the negative of) $\kappa^{(1)}$ evolves very similar to life expectancy. To this end, let φ_1 be the slope of the regression line to α_x for $x \in \{x_{\text{young}}, \ldots, x_{\text{old}}\}$. Note that in the following $a \leftarrow b$ denotes that a gets assigned the value of b, which differs from $a = b$ in the sense that, e.g. $a \leftarrow a + 1$ is reasonable and increases a by 1. For all x and t set

[5] Regarding the choice of $x_{\text{cutoff}} = 45$, we have observed that the cohort parameters become stable over time for current ages roughly above age 45.

[6] For example, MATLAB's Statistics Toolbox and the freely available statistical software R provide ready-to-use implementations.

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

Moreover, we apply an additional re-definition of the parameter α_x. The idea is to have best fitting results at most recent past in order to be able to forecast near future more accurately. We therefore redefine

$$
\begin{aligned}
\alpha_x \leftarrow \sum_{t=t_{\min}}^{t_{\max}} \frac{w_t}{\sum_{t'} w_{t'}} \Big(&\operatorname{logit} q_{x,t} - \kappa_t^{(1)} - \kappa_t^{(2)}(x - x_{\text{center}}) \\
&- \kappa_t^{(3)}(x_{\text{young}} - x)^+ - \kappa_t^{(4)}(x - x_{\text{old}})^+ - \gamma_{t-x} \Big) \ ,
\end{aligned}
$$

where $w_t = (1 + 1/h)^t$ are weights to emphasize the near past and fade out the more distant past, see also Sections 2.4.1 and 2.5 for a discussion of the choice of the weights and the parameter h, respectively.

2.4 Simulation of Future Mortality Rates

Once all parameters have been estimated the time dependent parameters are projected into future years $t_{\max} + 1, t_{\max} + 2, \ldots$ as follows.

2.4.1 Projection of the $\kappa_t^{(1)}$

Reasonably forecasting $\kappa^{(1)}$ in the context of the one-year view is crucial. It is $\kappa^{(1)}$ that is the most important parameter of the model, where important is both understood in the sense of the ability to explain historical movements of mortality rates, and in the sense of economic impacts of trend risk.

Figure 3 shows the historical evolution of the process for some example populations. Just like for life expectancies (see Figure 1) we observe rather linear trends with changes of the slope over time. Therefore, it is obvious to model the future evolution of $\kappa^{(1)}$ as a linear trend with stochastic drift (and intercept) and some additional noise around this stochastic trend. Following the idea of re-estimating the model parameters in the *one-year view*, we re-estimate the linear trend yearly based on a data set which includes an additional realization of $\kappa^{(1)}$ for the current year. This approach to simulating $\kappa^{(1)}$ can be formulated in two different ways which are either more intuitive or mathematically elegant.

In the first formulation, we fit a regression line by weighted least squares to the historical $\kappa_t^{(1)}$. This line constitutes the current best estimate trend. To obtain stochastic forecasts we then add a (normally distributed) random variable on top of the regression line at $t_{\max}+1$. This yields the stochastic forecast at time $t_{\max}+1$. 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 times $t_{\max}+2, t_{\max}+3, \ldots$

For $t = t_{\min}+2, \ldots, t_{\max}$ let $\ell_t \colon [t_{\min}, \infty[\longrightarrow \mathbb{R}$ denote the regression line obtained for $\kappa_{t_{\min}}^{(1)}, \ldots, \kappa_t^{(1)}$ with respect to the introduced weights w. The general form of these weights $w_t = (1+1/h)^t$ is chosen such that

$$\sum_{t \le t_{\max}} \frac{w_t}{\sum_{t' \le t_{\max}} w_{t'}} = h \ ,$$

i.e. the average number of considered historical years is h. Or, to put it in another way, the years up to h in the past have a cumulative weight of about $1-1/e$, i.e. roughly $2/3$, while the remaining older years only have a cumulative weight of about $1/e$, i.e. roughly $1/3$. The value of h is crucial since it defines how "fast" best estimates will be readjusted from year to year. This parameter can be considered as a measure of inertia. A large value of h means that the new data has a comparably small weight and has little effect when updating best estimates in year $t_{\max}+1$. If h on the other hand is small it is the other way round.

Let $\varepsilon_t^{(1)} \sim \mathcal{N}(0,1)$ be independent and identically, normally distributed. For $t > t_{\max}$ we assume

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

where $\sigma^{(1)}$ denotes the sample standard deviation of the empirical errors $\kappa_t^{(1)} - \ell_{t-1}(t)$ for $t = t_{\min}+2, \cdots, t_{\max}$ weighted by $w_t^* = (1+1/h^*)^t$, and $\overline{\sigma}^{(1)}$ is an add-on to ensure a reasonably conservative volatility. In Section 2.5 we discuss the volatility add-on $\overline{\sigma}^{(1)}$ and the weights w_t^* in more detail.

In the alternative formulation, an autoregressive time series process is used for the projection of $\kappa^{(1)}$. Let $n = t_{\max} - t_{\min} + 1$ be the number of years of historical data. Then the linear trend fitting approach is equal to applying an AR(n) process[7]

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

where $a_j = (1 + 1/h)^j (c_1 + c_2 j)$ for values c_1, c_2 that can be expressed in terms of n and h.

In the literature, the process $\kappa^{(1)}$ is usually projected as a random walk with drift. However, our approach of a stochastic linear trend offers several advantages compared to that formulation. First, our $\kappa^{(1)}$ process is feasible for the *one-year view* and the *run-off view* at the same time and we have consistency in the sense that the *run-off view* is nothing else but the iteration of the *one-year view*.[8] For a random walk with drift, the whole model has to be re-estimated in the *one-year view* which, in general, prohibits consistency with the *run-off view* and requires significant additional computation time.

Also for the *run-off view* alone the stochastic linear trend approach seems more adequate. The random walk with drift has come under some criticism for yielding, e.g. in the model of Lee and Carter (1992), rather wide confidence bounds in the short run and implausibly narrow confidence bounds in the long run. In Figure 2, we see that our $\kappa^{(1)}$ process overcomes this issue. Moreover, if the current slope of the linear trend is positive, for the random walk with drift, this right away excludes the possibility that there might exist a natural upper bound for life expectancy because the probability that life expectancy exceeds any given upper bound tends to 1 as $t \to \infty$.

2.4.2 Projection of the $\kappa_t^{(2)}, \kappa_t^{(3)}$, and $\kappa_t^{(4)}$

Figure 3 presents historical evolutions of the processes $\kappa_t^{(2)}, \kappa_t^{(3)}$, and $\kappa_t^{(4)}$ for some example populations. We observe that the processes evolve quite differently for those populations and it does not seem biologically reasonable to assume any kind of persistent (linear) trend for the projection of these processes. Therefore, we project $\kappa^{(2)}, \kappa^{(3)}, \kappa^{(4)}$ as three-dimensional random walk. The choice of a random walk also fits the idea of the *run-off view* being

[7] To be more precise, equality between the two formulations only holds if a fixed number n of $\kappa^{(1)}$ values to determine ℓ_t for $t > t_{\max}$ are used, i.e. $\kappa_{t-n+1}^{(1)}, \ldots, \kappa_t^{(1)}$. However, due to the weights applied in the fitting of ℓ_t this is approximately true for a reasonably large value of n.

[8] To be more precise, a constant volatility over time is needed for equivalence.

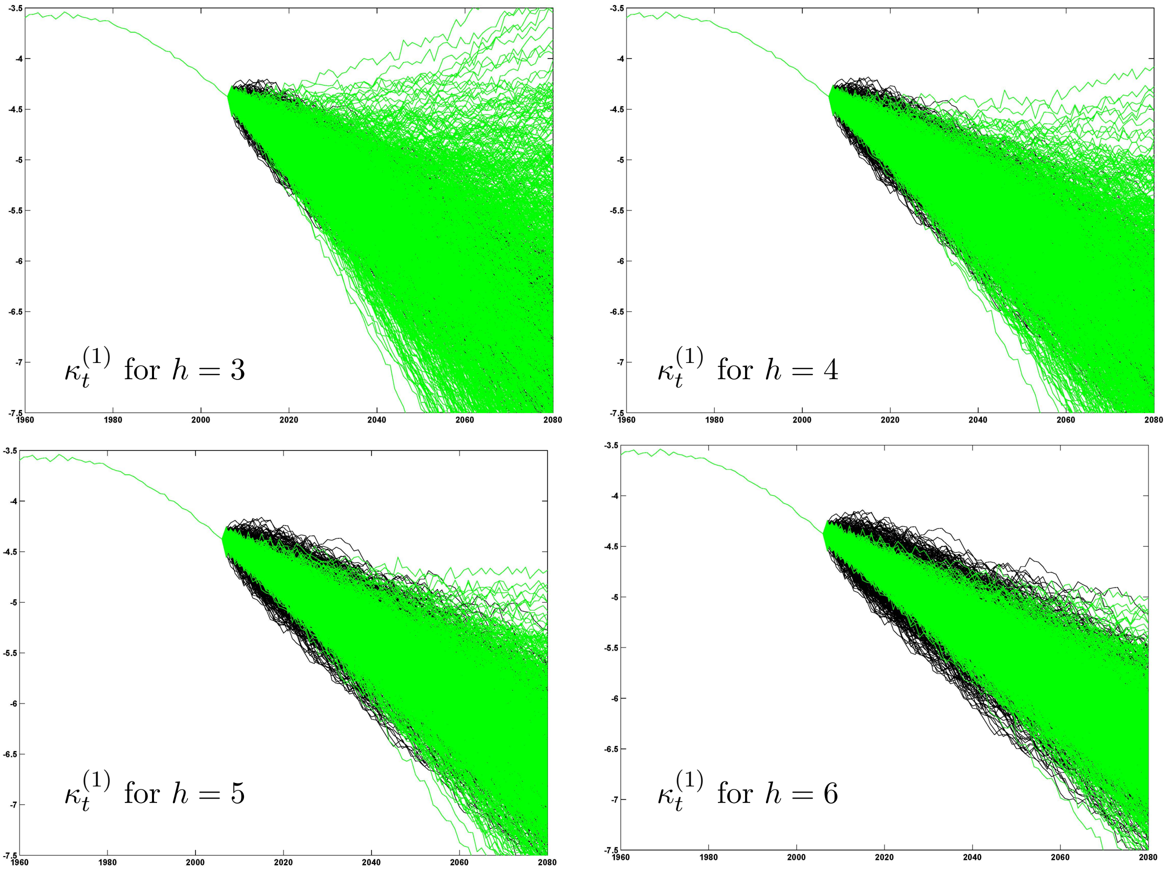

Figure 2: Comparing the different approaches for modelling $\kappa_t^{(1)}$, i.e. the common random walk with drift approach (black lines) and the stochastic linear trend approach (green lines) for different values of the weighting parameter h from 3 to 6

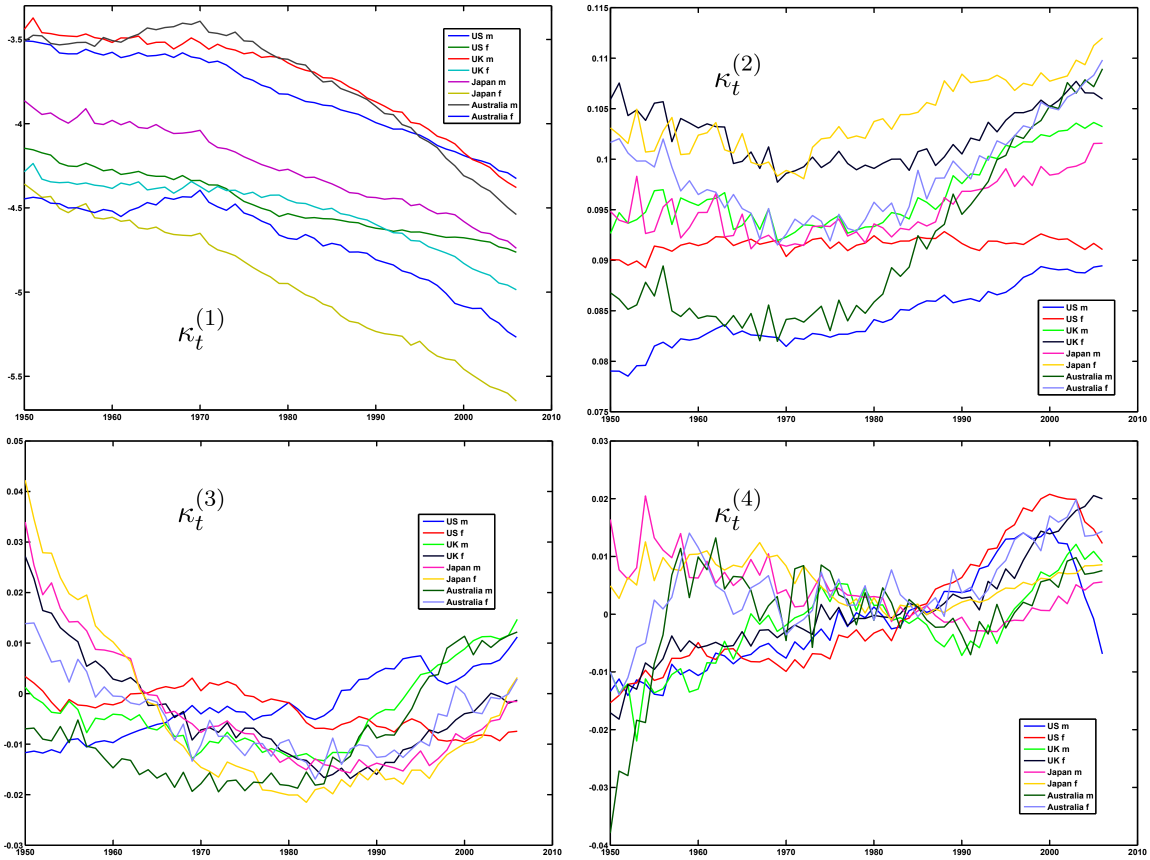

Figure 3: Parameters $\kappa^{(1)}, \kappa^{(2)}, \kappa^{(3)}$, and $\kappa^{(4)}$ for some example populations

an iterative application of the *one-year view* which we discussed for $\kappa^{(1)}$ in the previous section.

Additionally, for $\kappa^{(2)}$ we increase the empirical standard deviation by $\overline{\sigma}^{(2)}$, analogous to $\kappa^{(1)}$, i.e.

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \varepsilon_t^{(2)}(\overline{\sigma}^{(2)} + \sigma^{(2)}) \ . \tag{4}$$

An analysis of the historical data of various populations reveals mostly significant and very similar correlations between the $\varepsilon_t^{(2)}, \varepsilon_t^{(3)}$, and $\varepsilon_t^{(4)}$. Therefore, we carry over these correlations into the projection. The correlations between $\kappa_t^{(1)}$ on the one hand, and $\kappa_t^{(2)}, \kappa_t^{(3)}, \kappa_t^{(4)}$, on the other hand, are mostly insignificant and vary in sign. We thus interpret them as not persistent and assume independence between $\kappa_t^{(1)}$ and the remaining parameters for the projections.

2.4.3 Projection of the γ_{t-x}

The cohort parameter γ_{t-x} is projected into the future as imposed stationary AR(1) process $\gamma_{t-x} = a + b\gamma_{t-x-1} + \mathcal{N}(0, \sigma^2)$, i.e. $b \leftarrow \min\{c_{\text{damp}}, b\}$, for a positive constant $c_{\text{damp}} < 1$,[9] and the mean is redefined to be 0, i.e. $a \leftarrow 0$. The actual projection starts at cohorts with present age x_{cutoff}. The concept is analogous to Plat (2009). Furthermore, in order to avoid a jump at $q_{x,t}$ from $t_{\max}$ to $t_{\max} + 1$ for $x < x_{\text{cutoff}}$ we linearly fade in these cohort parameters during $x_{\text{cutoff}} - x_{\min}$ years. This obviously raises the question whether cohort effects have a significant impact for the *one-year view* at all and whether they should be included in a model for solvency purposes in the first place. We believe that the cohort parameters are still important in the initial parameter estimation because they support the correct identification of all other parameters.

2.5 Volatility Add-ons and Weighting Parameters

In this section we discuss all additional fixed model parameters, i.e. parameters that have to be calibrated once but are not estimated within each model

[9] Additionally, for the very unlikely case that $b < -c_{\text{damp}}$ we also use $b \leftarrow \max\{-c_{\text{damp}}, b\}$. The actual choice of c_{damp} should be rather close to 1. The effect is that it takes $\ln(1/2)/\ln(c_{\text{damp}})$ years until future expected cohort parameters halve, e.g. roughly 23 years for $c_{\text{damp}} = 0.97$.

run. This includes the volatility add-ons $\overline{\sigma}^{(1)}$ and $\overline{\sigma}^{(2)}$ for the $\kappa^{(1)}$ and $\kappa^{(2)}$ processes as well as the weighting parameters h and h^* which define the weights w and w^*. For the modeling of long-term trend deviations the parameter h is the most important one.

The general procedure how these parameters are calibrated is as follows.

- First, the volatility add-on $\overline{\sigma}^{(1)}$ which is added to the (observed) volatility $\sigma^{(1)}$ is calibrated to historically observed sudden and extreme changes in the mortality regime, e.g. a sudden drop of life expectancy after a period of low volatility.[10] Because of the low volatility such a drop would be considered extremely unlikely, e.g. only with a probability of 1 in a million. Currently, in most industrialized countries volatilities have been rather low over the last decades.[11] The volatility add-on $\overline{\sigma}^{(1)}$ ensures that all historically observed *short-term deviations* were to be expected with a reasonably high probability that corresponds to the size of available historical data, i.e. the more data is available the more it is likely that it contains possible extreme events.

- After $\overline{\sigma}^{(1)}$ has been defined it is mainly the parameter h that determines how volatile simulated long-term trend deviations are. To be more precise, while for modeling $\kappa^{(1)}$ as random walk with linear trend it would only be $\sigma^{(1)}$ (and $\overline{\sigma}^{(1)}$) that determine the long-term behavior of $\kappa^{(1)}$, for our stochastic linear trend additionally h comes into play. Figure 2 illustrates the impact of the parameter h. Together with $\overline{\sigma}^{(1)}$ the parameter h ensures that all historically observed *long-term trend deviations* were to be expected with a reasonably high probability.

- The volatility add-on $\overline{\sigma}^{(2)}$ is calibrated analogously to $\overline{\sigma}^{(1)}$. However, this add-on cannot be set to a very large value because otherwise the correlation of simulated $q_{x,t}$ between very old and very young ages might become significantly negative which is not plausible.

[10]Recall that catastrophe risk still is excluded in our considerations. Extreme shock events like for example the Spanish Flu hence are not taken into account for the calibration of $\overline{\sigma}^{(1)}$. However, the drop in life expectancy of Russian males during the fall of communism in the beginning of the 1990's does not belong to the category of catastrophe risk and thus we take it into account.

[11]For instance, Bauer and Kramer (2010) show that volatility in the US was significantly larger during the first half of the previous century than during the second half.

- Finally, h^* determines the weighting of the (observed) volatility $\sigma^{(1)}$ over time. A small value means that the very recent volatility is considered with more weight for the calculation of $\sigma^{(1)}$.

Note that the four parameters cannot be calibrated independently from each other. For instance, for a larger volatility add-on $\overline{\sigma}^{(1)}$, also a larger value for h would be required/sufficient in order to allow for comparable trend changes. Thus, $\overline{\sigma}^{(1)}$ obviously impacts the calibration of h. Also, changes of h and h^* impact the (observed) volatility $\sigma^{(1)}$ such that the add-on $\overline{\sigma}^{(1)}$ would have to be calibrated again. However, the calibration of $\overline{\sigma}^{(1)}$ is only very weakly dependent on the other parameters, such that the calibration of this parameter is the best starting point for a potentially iterative calibration process. We provide an example.

We propose to define $\overline{\sigma}^{(1)}$ such that the most severe *short-term*/one-year mortality deviations of basically all industrialized countries[12] after World War II are still covered by the model, the most severe one being the drop of life expectancy of Russian males in the 1990's during the fall of communism. By "covered" we mean that deviations should be close to the 99% quantile with respect to life expectancy, or nearly equivalent with respect to $\kappa^{(1)}$. We use the 99% quantile since this corresponds to events that take place once in 100 years, which is roughly the average number of years of available historical data, some countries offering a lot more years, some only very few years. For $\overline{\sigma}^{(1)} = 0.08$, which has to be seen in relation to an observed volatility of $\sigma^{(1)} \approx 0.06$ in Russia at the beginning of the 1990's, this criterion is met.

For *long-term* trend deviations it is Dutch males that provide the most severe historical example, where in the 1950's and 1960's life expectancy was slightly decreasing and volatility has been rather low, and from the 1970's on life expectancy began to increase very rapidly.[13] Any model that only uses historical data from 1950 to 1970 fails to forecast this rapid increase of life expectancy. But, for example, $h = 6$ (together with $\overline{\sigma}^{(1)} = 0.08$) still sees this rapid increase at the 95% quantile with respect to life expectancy in 2006.[14] For $h = 4$ the rapid increase would be at the 90% quantile. In contrast to calibrating $\overline{\sigma}^{(1)}$ such that the maximum observed deviation matches a specific

[12] We considered the countries available at the Human Mortality Database

[13] The same effect can be seen in various other male populations, but in terms of unexpected trend deviation Dutch males provide the most severe example.

[14] The same holds with respect to the sum of life expectancies (year by year) from the 1970's until 2006.

quantile, where about 100 (independent) historical samples of short-term/ one-year deviations are available, the calibration of h to a specific quantile is more uncertain because it requires at least, say, 10 years to observe a trend deviation and hence the number of (independent) historical samples is substantially lower. Slicing history into intervals of length, say, 10 years yields 10 samples. The "effective" number of samples might be increased to, say, 20 samples by partly overlapping intervals and considering different populations, which leads to the 95% quantile. However, given the still small sample number expert judgment is required, see also discussion in Section 5.

Figures 4 and 5 show the simulated short-term and long-term deviations.

Similarly, $\overline{\sigma}^{(2)}$ can be calibrated from the Russian males scenario. A value of $\overline{\sigma}^{(2)} = 0.0005$ (compared to an observed volatility of $\sigma^{(2)} \approx 0.0015$) is sufficient and at the same time the correlation between very young and very old ages stays positive or becomes only very slightly negative for all populations available from the Human Mortality Database.

Finally, for h^* we propose a large value, say from 30 on, or even $w^* \equiv 1$. If h^* is small we would implicitly assume that trend risk is time dependent in the sense that there might be large differences in trend risk within only very few years in which the volatility changes. Given the long time horizons when trend risk manifests itself this is counterintuitive. Thus, a large value for h^*, and also a significantly large value for $\overline{\sigma}^{(1)}$ prevent 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. Note that factors of 2 and more between observed volatilities of different regimes can be found in history.

Figure 6 shows some example backtesting results, i.e. life expectancies, some specific $q_{x,t}$ and parameters, for US males where the model uses data from 1950 to 1990 and forecasts up to 2006.

3 Simultaneous Modeling of Several Populations

So far, we have only considered mortality trend deviations of one specific population. In practice, however, one has to model deviations for at least two populations simultaneously: males and females of the same region/country. Moreover, portfolios for populations from different countries are often com-

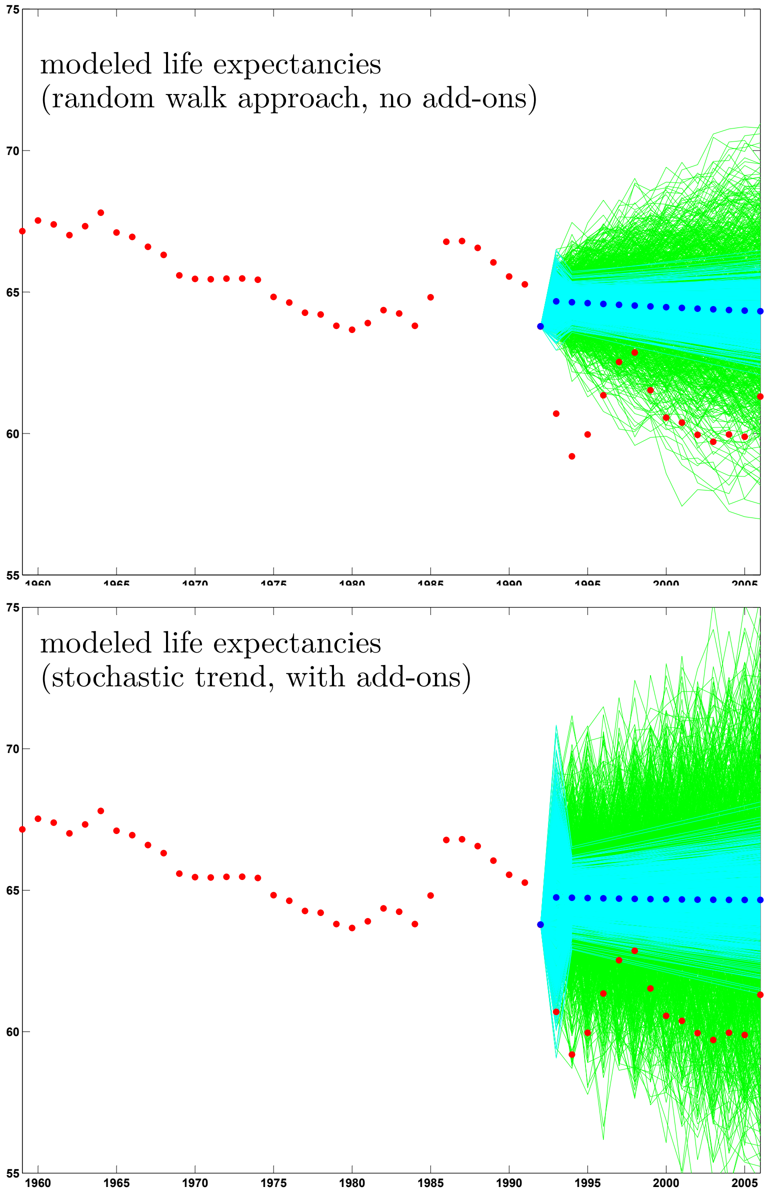

Figure 4: Backtesting on historical mortality rates of Russian males, using data from 1959 to 1992, forecasting up to year 2006. The color red shows historical life expectancies at birth from 1959 to 2006, dark blue are best estimates as calculated from data up to year 1992, green are 1000 realizations of the run-off view, light blue are the corresponding one-year view realizations. The upper plot shows outputs with the common random walk approach for $\kappa^{(1)}$ without volatility add-ons $\overline{\sigma}^{(1)} = 0.08$ (and $\overline{\sigma}^{(2)} = 0.0005, h = 6$), the lower one includes the add-ons and uses the proposed process for $\kappa^{(1)}$

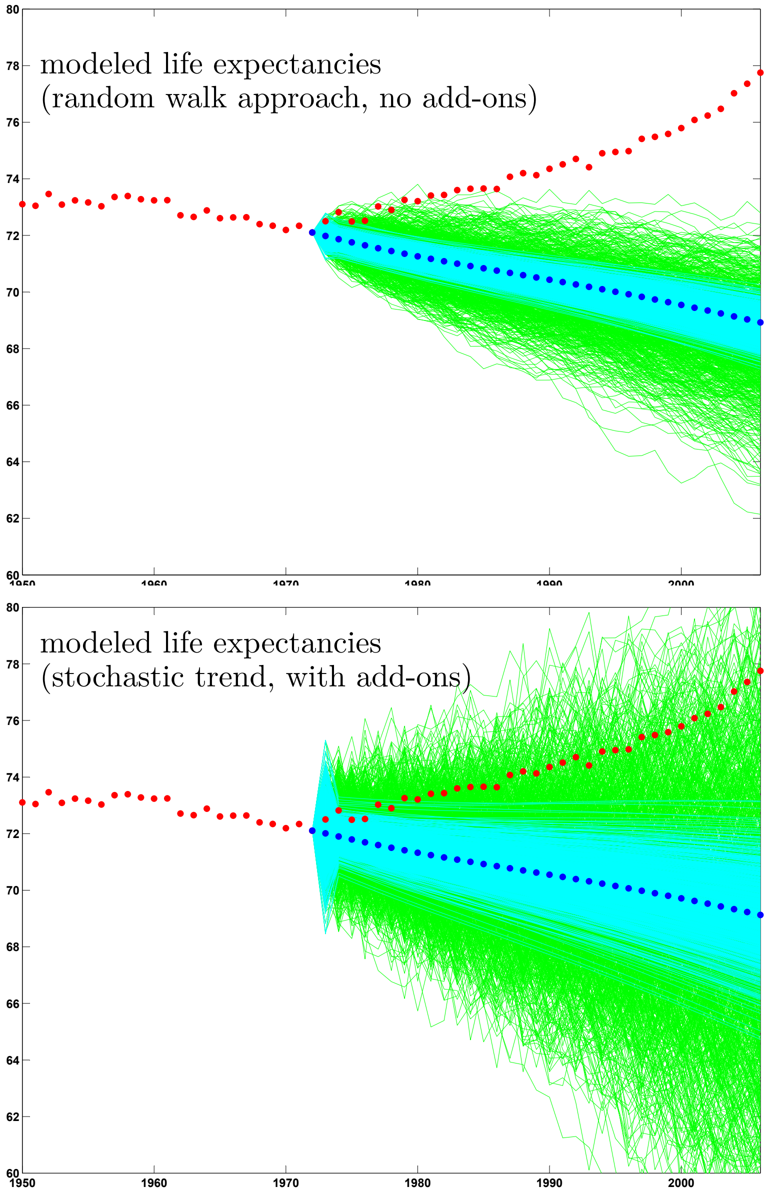

Figure 5: Backtesting on historical mortality rates of Dutch males, using data from 1950 to 1972, forecasting up to year 2006. The color red shows historical life expectancies at birth from 1950 to 2006, dark blue are best estimates as calculated from data up to year 1972, green are 1000 realizations of the run-off view, light blue are the corresponding one-year view realizations. The upper plot shows outputs with the common random walk approach for $\kappa^{(1)}$ without volatility add-ons $\overline{\sigma}^{(1)} = 0.08$ (and $\overline{\sigma}^{(2)} = 0.0005, h = 6$), the lower one includes the add-ons and uses the proposed process for $\kappa^{(1)}$

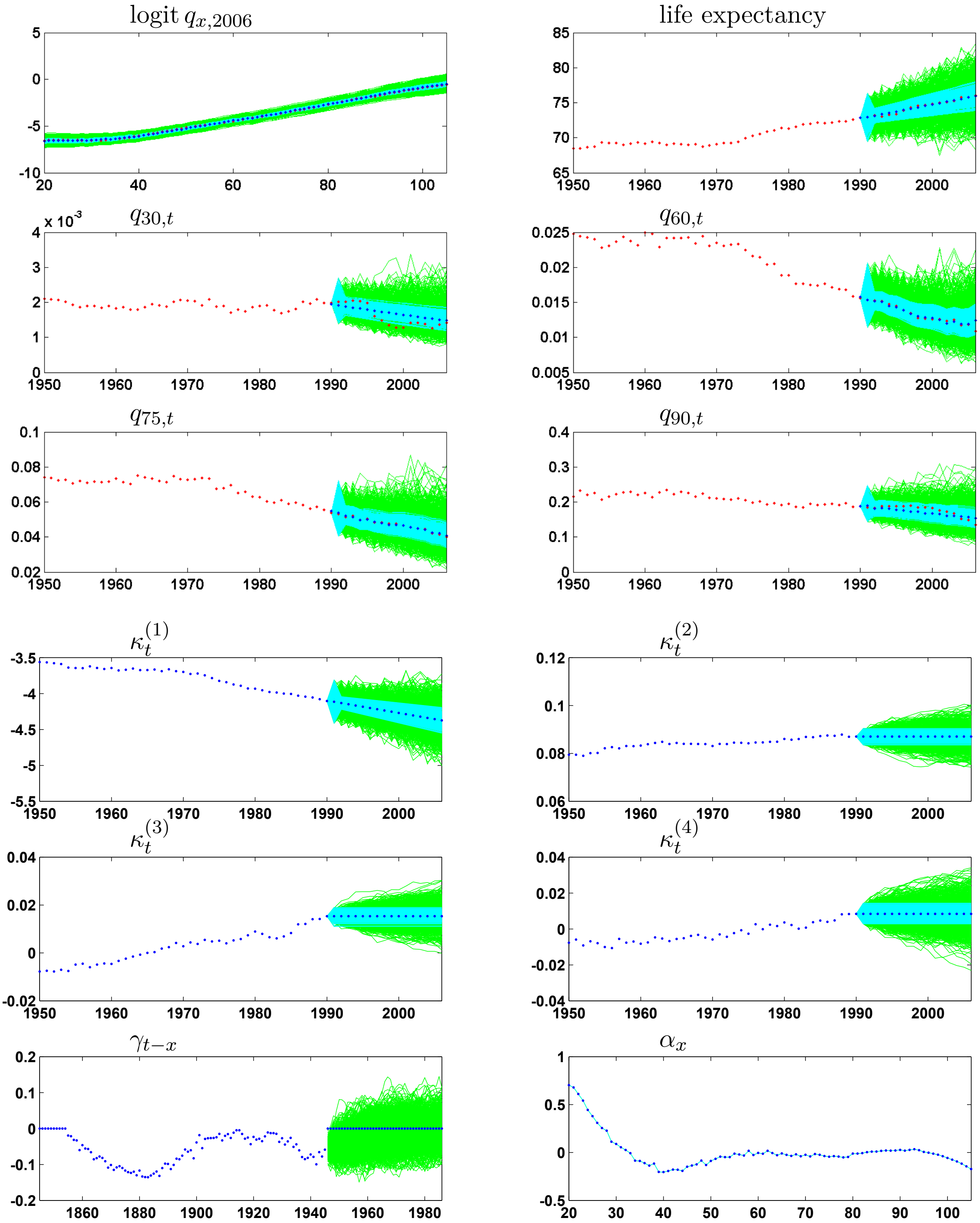

Figure 6: Backtesting on historical mortality rates of US males, using data from 1950 to 1990, forecasting up to year 2006. The color red denotes historical data, dark blue are best estimates, green are 1000 realizations of the run-off view, light blue are the corresponding one-year view realizations

bined to exploit diversification effects. Thus, an extension of the model from the previous section is indispensable. Note also that, even if mortality scenarios for only one or two populations are required, it is generally worthwhile to consider other populations in the process as well. As Jarner and Kryger (2008) show these populations can provide additional information and thus improve forecasts quite substantially and reduce uncertainty of forecasts.

Modeling only next year's mortality rates for several populations (without considering the years after) is easily done via the sample correlation matrices of the innovations of $\kappa_{t,p}^{(1)}$ and $\kappa_{t,p}^{(2)}$ where p defines a population.[15] However, these correlations are not necessarily suitable for trend deviations. The correlations of the innovations are usually quite low and hence the resulting mortality trend deviations would carry over the weak dependency (see also Plat (2010) and Coughlan et al. (2010)). But this clearly contradicts the historically observed strong parallelism of mortality evolutions, see Figure 1.

To account for this, in this section we introduce a method that is a special case of cointegration and an error correction model, respectively. Similar concepts have been introduced in Li and Lee (2005) and Jarner and Dengsøe (2009).

3.1 Modeling Differences to Common Trend

Let P be a set of populations p, e.g. as given in Figure 1. For each population p we obtain estimated parameters $\alpha_{x,p}, \kappa_{t,p}^{(1)}, \kappa_{t,p}^{(2)}, \kappa_{t,p}^{(3)}, \kappa_{t,p}^{(4)}, \gamma_{t-x,p}$. Additionally, by using the combined mortality rates of all populations in P, i.e.

$$q_{x,t,P} = \frac{\sum_{p \in P} S_p q_{x,t,p}}{\sum_{p \in P} S_p} \ ,$$

where $q_{x,t,p}$ denote mortality rates of population p and S_p is its population size, we obtain corresponding parameters $\alpha_{x,P}, \kappa_{t,P}^{(1)}, \kappa_{t,P}^{(2)}, \kappa_{t,P}^{(3)}, \kappa_{t,P}^{(4)}, \gamma_{t-x,P}$ of the combined/total population. While for the total population the same projection of these parameters into the future as already introduced in the previous sections is used, the projections for each specific population differ. Consider the differences

[15] Across populations we only consider the correlations between the most important parameters, i.e. $\kappa_{t,p}^{(1)}$ and $\kappa_{t,p}^{(2)}$. However, the correlation between $\kappa_{t,p}^{(2)}, \kappa_{t,p}^{(3)}, \kappa_{t,p}^{(4)}$ of the same population also induces non-trivial correlations between $\kappa_{t,p}^{(3)}, \kappa_{t,p}^{(4)}$ of different populations.

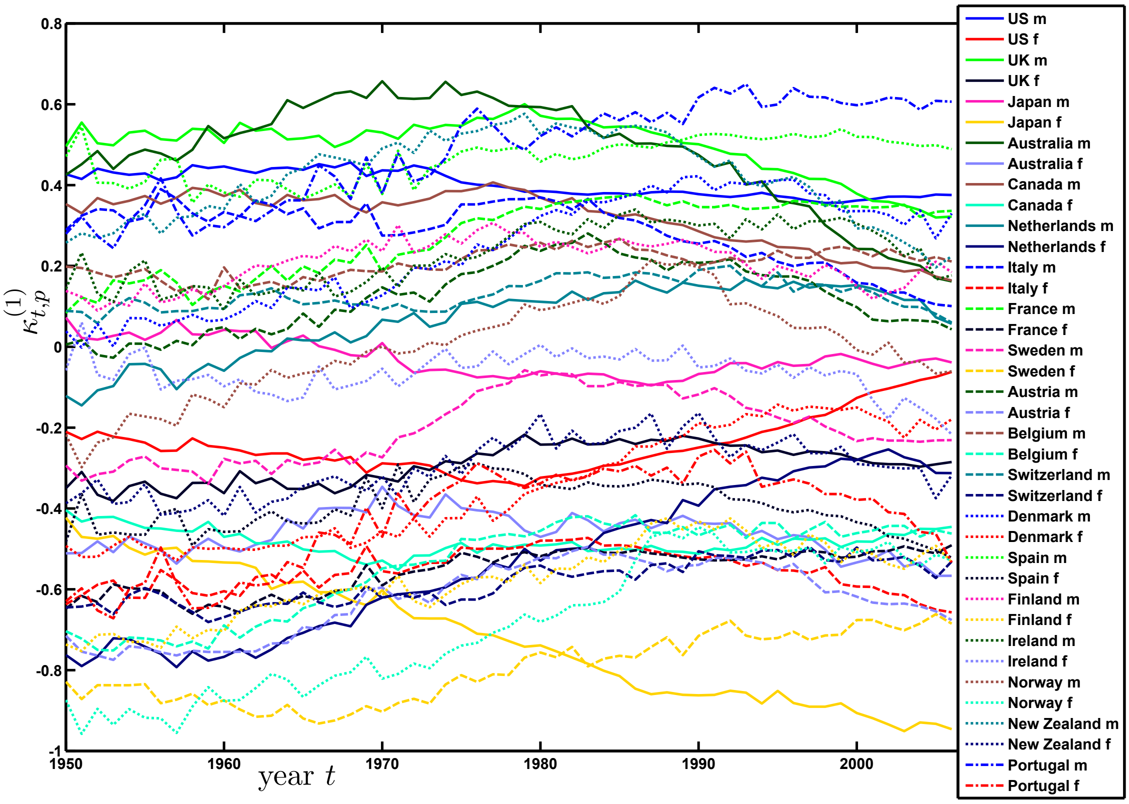

Figure 7: Differences $\kappa^{(1)}_{t,p} - \kappa^{(1)}_{t,P}$ for various populations p. The evolutions of $\kappa^{(1)}$ are very similar to life expectancy. The maximum difference of about 1.5 between Japanese females and Portugal males in 2006 corresponds to a gap in life expectancy of about 10 years

$$\kappa^{(1)}_{t,p} - \kappa^{(1)}_{t,P} \ ,$$

which are shown in Figure 7. Obviously, the $\kappa^{(1)}_{t,p}$ of a population p might deviate from the $\kappa^{(1)}_{t,P}$ quite substantially, but on the long-term we can expect that the differences will not become arbitrarily large, see also Jarner and Kryger (2008). The aim of the following is to estimate how large the differences might become, what the expected long-term difference is, and how fast convergence to the long-term difference could be expected. We model the differences as AR(1) processes

$$\kappa^{(1)}_{t,p} = \kappa^{(1)}_{t,P} + a_p + b_p(\kappa^{(1)}_{t-1,p} - \kappa^{(1)}_{t-1,P}) + \varepsilon_{t,p} \ , \tag{5}$$

where $\varepsilon_{t,p}$ are $\mathcal{N}(0, \sigma_p)$ distributed and serially independent. Across populations we allow for dependencies via the sample correlation matrix of the $\varepsilon_{t,p}$.

The typical size of the difference $\kappa^{(1)}_{t,p} - \kappa^{(1)}_{t,P}$ is defined by the parameters b_p and σ_p, $a_p/(1-b_p)$ defines the expected long-term difference, and b_p defines how fast the expected long-term difference is approached.

Note that this can be considered as a special case of cointegration (see, e.g. Engle and Granger (1987)), where we explicitly use the population sizes S_p for each population as well as the (negative of the) total population size as cointegration vector. Furthermore, all coefficients corresponding to populations' year-to-year differences are zero.

We mention some refinements of the estimators a_p, b_p, σ_p that should be used. Having a too small dataset for example might yield an estimator $b_p > 1$ (or even $b_p < -1$) such that (best estimate) forecasts diverge. An upper bound $c_{\text{damp}} < 1$ should be used. Furthermore, the expected long-term difference $a_p/(1-b_p)$ should be monitored, obviously in particular if b_p has been modified beforehand, a_p also has to be modified. Several methods, e.g. a (weighted) average of the differences $\kappa^{(1)}_{t,p} - \kappa^{(1)}_{t,P}$ can be used to determine various values for $a_p/(1-b_p)$. But eventually, since the data basis for these kind of long-term assumptions is very small, expert judgment is necessary. In case P contains populations with very different trend evolutions, e.g. Russia and some Eastern European populations vs. US and Western European populations, where the latter show very constant improvements while the former remain pretty much at the same level, it becomes even more uncertain what the expected long-term differences are. Again, expert judgment is needed in such cases.

In Table 1 we present estimated long-term differences based on 4 different methods using historical data from 1950 to 2006:

1. Directly use the estimated value $a_p/(1-b_p)$.
2. Use (unweighted) average of the differences.
3. Use w-weighted average of the differences, i.e. weighted by w.
4. Extrapolate the current linear trend (w-weighted linear regression) of the differences 5 years into the future.

The first method might right away be appropriate if enough data is available but note that usually this is not the case. For example, currently there appears to be a linear trend in the evolution of the differences of Japanese females. Here, only lots of additional future years might reveal the "true

Table 1: Expected long-term differences $\kappa_{t,p}^{(1)} - \kappa_{t,P}^{(1)}$ based on 4 methods as described in the Section 3.1. The set P contains the populations listed in Figure 7

Population	1.	2.	3.	4.
US males	0.339	0.394	0.347	0.343
US females	-0.314	-0.251	-0.144	-0.037
Japan males	-0.067	-0.036	-0.064	-0.053
Japan females	-1.298	-0.726	-0.940	-1.009
Australia males	0.629	0.472	0.223	0.045
Australia females	-0.488	-0.470	-0.550	-0.619
Dutch males	0.087	0.056	0.090	0.048
Dutch females	0.257	-0.542	-0.327	-0.272
Swedish males	-0.200	-0.212	-0.238	-0.298
Swedish females	-0.762	-0.806	-0.717	-0.682
UK males	0.653	0.489	0.346	0.241
UK females	-0.296	-0.300	-0.304	-0.331

behavior". A continuation of the linear trend on the long-term seems very unlikely. Thus, the fourth method seems most appropriate for Japanese females. For UK females, where the differences remain pretty much constant over time, the 4 methods produce very similar results.

Also, to avoid the immediate start of reversion of the process to its mean (in the case of best estimate forecasts) the current short-term linear trend might be continued for some years until the process reverts to its mean.

Algorithm 4 in the Appendix presents the multi-population model.

3.2 Recalibration of the Model to Common Trend

In Section 2.5 we calibrated the parameters $\overline{\sigma}^{(1)}$ and h by the principle that $\overline{\sigma}^{(1)}$ takes into account all historically observed short-term deviations, and h takes into account historically observed long-term trend deviations.

In this section we discuss how these parameters have to be modified for the multi-populations model, i.e. the extension of the model to several populations as described in the previous section.

If we apply the ideas of Section 2.5 to the total population we come across the very same years where extreme scenarios can be found. The short-term mortality deviations of Russian males are also present for Ukrainian males and other populations, such that these deviations are carried over to the common trend in a weaker form. For the common trend using historical data from year 1959 on leaves 24 of the 37 countries available at the Human Mortality Database.[16] This leaves around 50 years of historical data, such that $\overline{\sigma}^{(1)} = 0.02$ (while $\sigma^{(1)} = 0.015$) shows the deviations during the beginning of the 1990's at the 98% quantile.

For the calibration of h we observe the same "inflection point" in the common trend in the beginning of the 1970's but in a weaker form. A value of h between 4 and 5 would see the deviations between the 90% and 95% quantile. These values should be understood as an example how the model can be calibrated, see discussion in Section 5.

If we compare the multi-populations model (with $h = 4$ and $\overline{\sigma}^{(1)} = 0.02$) with the single-population model (with $h = 6$ and $\overline{\sigma}^{(1)} = 0.08$) we typically observe improved forecasts while at the same time the uncertainty of forecasts is slightly reduced, see Figure 8 for a typical example.

4 Mortality/Longevity Threat Scenarios

For the computation of risk measures such as the 99.5% Value-at-Risk or the 99% Expected Shortfall, we are particularly interested in tail mortality/longevity scenarios. However, mortality models are typically calibrated to data series which rarely contain such extreme scenarios. Thus, a fair amount of uncertainty remains whether tail scenarios generated by a mortality model are adequate and sufficiently severe. Therefore, we recommend using epidemiological or demographic expert judgment 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 or longevity stress scenarios in the Solvency II standard model, even though their adequacy has been questioned (cf. Börger (2010)).

Given some specific threat scenario, one needs to answer the question

[16] From the 37 countries available at the Human Mortality Database for the 5 countries Chile, Israel, Luxembourg, Slovenia, Taiwan there is no historical data available back until 1959. Furthermore, for the countries Belarus, Estonia, Germany, Latvia, Lithuania, Poland, Russia, Ukraine there is no data available back until 1950.

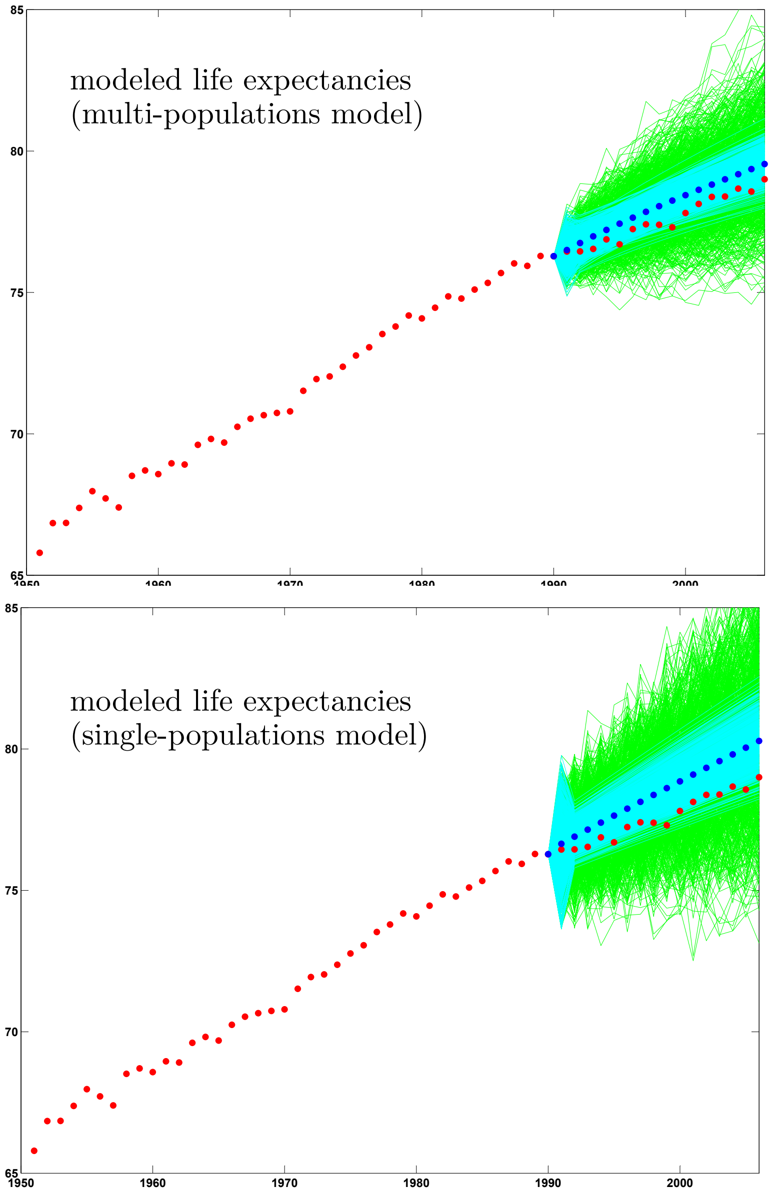

Figure 8: Comparing outcomes of the multi-populations model (above, 4. method, $h = 4$ and $\overline{\sigma}^{(1)} = 0.02$) to the single-populations model (below, $h = 6$ and $\overline{\sigma}^{(1)} = 0.08$) for Japanese males

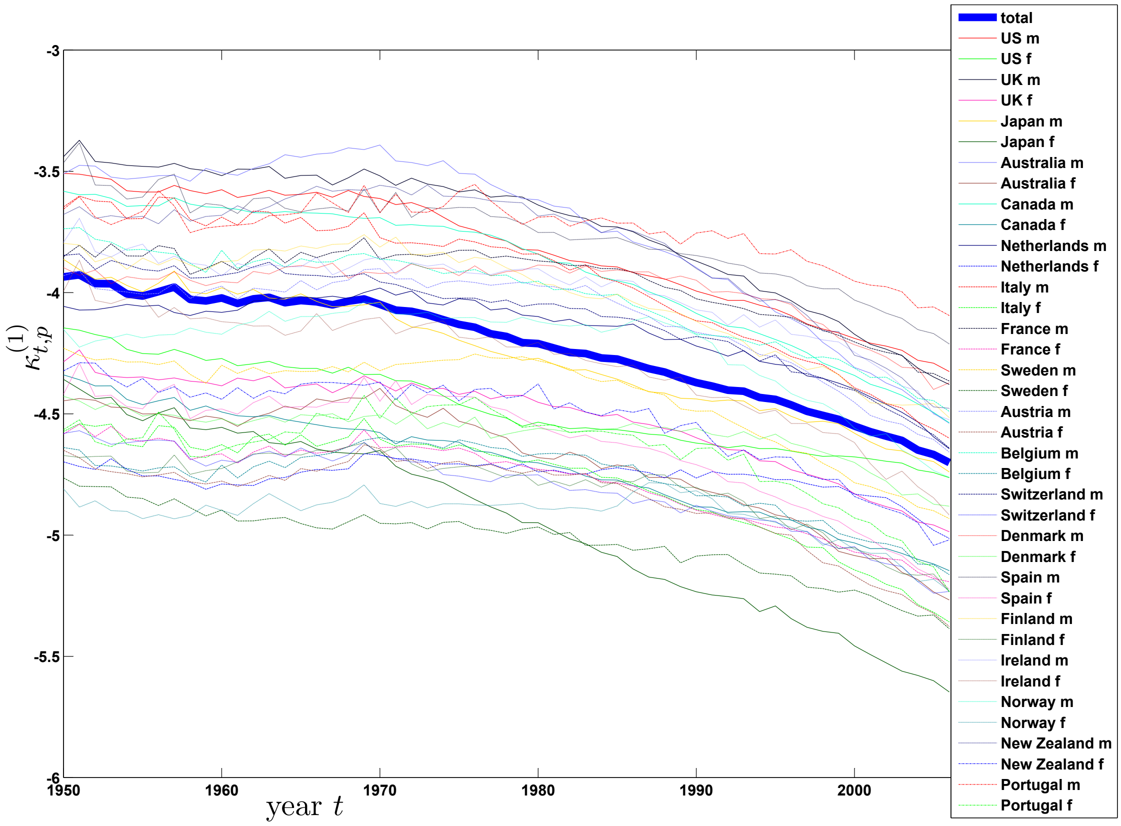

Figure 9: Parameters $\kappa^{(1)}$ of some populations p and the total population P

whether the scenarios from the stochastic model should already cover the threat scenario in terms of severeness. If this should be the case but is not, the model needs to be modified. For instance, larger volatility add-ons or a smaller weighting parameter h might be calibrated to the threat scenario instead of the historical data.

In case the calibration of the stochastic model seems appropriate the threat scenario can still be used to enrich the model outcomes, e.g. if the stochastic model's structure does not permit the generation of such a scenario. If the risk measure is the Expected Shortfall, the threat scenario could simply be added to the set of scenarios generated by the mortality model. Obviously, here the number of threat scenarios, i.e. the probability assigned to each threat scenario is crucial.

If risk is measured by the Value-at-Risk, the simple inclusion of a threat scenario into the set of scenarios from the stochastic model does not seem a valid approach. Threat scenarios – no matter how severe they may be compared to the scenarios from the stochastic model – only would shift the Value-at-Risk at 99.5% very slightly (as long as their occurrence probability is lower than 0.5%). Thus, in order to assign appropriate weight to the threat scenario it seems more reasonable to increase the volatility add-ons or decrease the parameter h until the most severe scenarios from the model are similar to the desired threat scenarios.

5 Remaining Risks

When modeling mortality or longevity risk one thing is for sure, nothing is absolutely certain. No matter how sophisticated a mortality model is, there are always some risks left which can roughly be divided into two classes, risks linked to the mortality model itself and risks that arise from a certain application of the model. In this section, we summarize these risks for our modeling framework and discuss their magnitude and relevance.

The most relevant risk in our framework certainly is the parameter risk 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. Slightly less crucial but still important are the volatility add-ons $\overline{\sigma}^{(1)}$ and $\overline{\sigma}^{(2)}$ as well as the volatility weighting parameter h^*. Also these parameters have to be chosen carefully. In comparison, we regard the uncertainty around all other parameters as less critical.

We consider the assumption that simulated realizations of next year's mortality rates determine the changes in the future mortality trend to be the main modeling risk as it may be very different from reality. 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. Such situations will not be captured by the model proposed in this paper. However, usually, changes in the mortality trend have a much more significant impact on capital requirements than the annual random variation of mortality rates (excluding short-term shock events). Therefore, we consider this risk not to be material.

One of the important application related risks is the modeling of dependencies between different populations. Rather parallel mortality evolutions have been observed in the past for all industrialized countries but this does not have to hold true for the future. In the case of mortality or longevity risk only, a strong trend dependence as implemented in our multi-population model is a conservative assumption. If mortality and longevity risk are combined for the different populations with diverging long-term trends this would be critical. Still, such a scenario can be accounted for by an appropriate mortality/longevity threat scenario.

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 described the future mortality evolution of the general population perfectly it is uncertain whether the adoption of mortality deviations to a population of insured would be adequate. However, we consider this risk to be limited because the insured are a subpopulation of the general population and thus, it is unlikely that mortality trends in the tail of the distribution are different for both populations. Therefore, we think it would be inappropriate to assume that insured mortality could diverge from general population mortality until infinity (see also Cairns et al. (2010)). Consequently, the long-term mortality trend of the insured should be very similar to the one for the general population and thus, we regard the adoption of mortality deviations for the general population to be appropriate.

Another caveat arises from the delayed availability of data. In nearly every country, mortality rates are currently only available with a time lag of at least two years. For the purpose of only modelling deviations from best estimates we propose to shift the historical data up to the present year. While this obviously introduces a shift to the absolute values of simulated future mortality rates, the relative deviations remain mostly unaffected. Recall that the weighting parameter h^* is supposed to be large such that a time lag of only a few years does not substantially change the measured volatilities.

Finally, we are making the assumption that mortality trends or tables are updated annually. In practice, this is usually not the case. Trend assumptions get updated only after some years of providing new information have passed hence leading to larger adjustments than have been allowed for in the model. However, we see this as a shortcoming of out-of-date mortality tables and hence, it should be accounted for by loadings in the mortality tables.

6 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 trend deviations deserve special attention. An adequate stochastic model needs not only fulfill the specific Solvency II and SST requirements but has to be efficient and sufficiently simple to maintain as well.

In this paper, we proposed, calibrated, and applied a mortality model which meets these criteria. It covers a wide range of ages which makes it suitable for determining both mortality and longevity risk capital. Furthermore, it includes cohort effects and has five stochastic drivers hence offering a great variability in the mortality scenarios it generates. A correlation structure between different ages has been implemented which makes the model applicable to a simultaneous assessment of mortality and longevity risk. Moreover, we discussed why margins need to be incorporated in the model and why weights have to be applied in the fitting of several model parameters. We also showed how these margins and weights can be specified and calibrated based on historical data of rather extreme mortality events/evolutions.

In order to respect the one-year time horizon, we followed the idea of a yearly re-estimation of the model. However, we specified the stochastic processes in our model such that the re-estimation is done implicitly, i.e. a full re-estimation of the model is not required which reduces computational time significantly. The model run-off is then obtained as an iterative application of the *one-year view* and we outlined the advantages of our specification of the stochastic trend process compared to the commonly used random walk with drift process.

A highly relevant issue from a practical point of view is the simultaneous mortality modeling for several populations. Therefore, we extended our model to a multi-population model where the mortality evolution for each individual population is driven by a combination of changes in a common trend and random fluctuations around this common trend. We showed that, even if risk is to be quantified for only one population, multi-population modeling is worthwhile. The additional data from other populations helps in determining the long-term mortality trend and thus typically reduces uncertainty.

Finally, we explained how expert judgment in form of mortality/longevity threat scenarios can be used to check the plausibility of a mortality model and the scenarios it generates. Moreover, these threat scenarios can also be applied to enrich the simulations by scenarios which the structure of the stochastic model does not allow for.

References

D. Bauer and F. Kramer. Risk and valuation of mortality contingent catastrophe bonds. *Working Paper, Georgia State University and Ulm University*, 2010.

M. Börger. Deterministic shock vs. stochastic value-at-risk – an analysis of the solvency II standard model approach to longevity risk. *Blätter der DGVFM*, 31:225–259, 2010.

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

A. Cairns, D. Blake, and K. Dowd. Modelling and management of mortality risk: A review. *Scandinavien Actuarial Journal*, 2:79–113, 2008.

A. Cairns, D. Blake, K. Dowd, G. Coughlan, and M. Khalaf-Allah. Bayesian stochastic mortality modelling for two populations. *Working Paper, The Pensions Institute, Cass Business School*, 2010.

G. Coughlan, M. Khalaf-Allah, Y. Ye, S. Kumar, A. Cairns, D. Blake, and K. Dowd. Longevity hedging: A framework for longevity basis risk analysis and hedge effectiveness. *Working Paper, The Pensions Institute, Cass Business School*, 2010.

R. Doff. A critical analysis of the solvency II proposals. *The Geneva Papers on Risk and Insurance Issues and Practice*, 33:193–206, 2008.

M. Eling, H. Schmeiser, and J. Schmit. The solvency ii process: Overview and critical analysis. *Risk Management and Insurance Review*, 10:69–85, 2007.

M. Eling, N. Gatzert, and H. Schmeiser. The swiss solvency test and its market implications. *The Geneva Papers on Risk and Insurance Issues and Practice*, 33:418–439, 2008.

R.F. Engle and C.W.J. Granger. Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 55:251–276, 1987.

N. Hari, A. De Waegenaere, B. Melenberg, and T. Nijman. Longevity risk in portfolios of pension annuities. *Insurance: Mathematics and Economics*, 42:505–519, 2008.

I. Holzmüller. 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, 2009.

S.F. Jarner and C. Dengsøe. Uncertainty and coherence of mortality projections. *Working Paper, The Danish Labour Market Supplementary Pension Fund (ATP)*, 2009.

S.F. Jarner and E.M. Kryger. Modelling adult mortality in small populations: The saint model. *Pensions Institute Discussion Paper*, PI 0902, 2008.

R. Lee and L. Carter. Modeling and forecasting us mortality. *Journal of the American Statistical Association*, 87:659–671, 1992.

N. Li and R. Lee. Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography*, 42(3):575–594, 2005.

J. A. Nelder and R. W. M. Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society, Series A*, 135:370–384, 1972.

A. Olivieri. Stochastic mortality: Experience-based modeling and application issues consistent with solvency 2. *Working Paper, University of Parma*, 2009.

A. Olivieri and E. Pitacco. Solvency requirements for life annuities: Some comparisons. *Working Paper, University of Parma and University of Trieste*, 2008a.

A. Olivieri and E. Pitacco. Stochastic mortality: The impact on target capital. *CARFIN Working Paper, University Bocconi*, 2008b.

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

R. Plat. One-year value-at-risk for longevity and mortality. *Working Paper, University of Amsterdam*, 2010.

T. Steffen. Solvency II and the work of ceiops. *The Geneva Papers on Risk and Insurance Issues and Practice*, 33:60–65, 2008.

R. Stevens, A. De Waegenaere, and B. Melenberg. Calculating capital requirements for longevity risk in life insurance products. using an internal model in line with solvency II. *Working Paper, Tilburg University*, 2010.

Swiss Federal Office of Private Insurance. Technical document on the swiss solvency test, 2006. Available at www.finma.ch.

A Algorithms

Algorithms 1 to 5 present the model described in this paper in full detail. When using 50 years of historical data an implementation of these algorithms typically results in running times of a few seconds for the single-population model and a few minutes for the multi-populations model with 50 populations.

Algorithm 1: Main algorithm to compute stochastic future mortality rates of one population. The subroutines for fitting and forecasting are given as Algorithms 2 and 3

Input: Historical mortality rates $q_{x,t}$ of a given population for $x = x_{\min}, \dots, x_{\max}, t = t_{\min}, \dots, t_{\max}$, year $t_{\text{fc}} > t_{\max}$ up to which forecasts are generated, argument view to calculate one-year view or run-off view

Output: n_s many realizations of future mortality rates $q_{x,t,s}$ (one-year view/run-off view) and future best estimate mortality rates $\hat{q}_{x,t}$ for $x = x_{\min}, \dots, x_{\max}$, $t = t_{\max} + 1, \dots, t_{\text{fc}}$ and $s = 1, \dots, n_s$

`Main(`q, t_{fc}, view`)` **begin**

define function $f_v \colon (x,t) \mapsto \text{logit}^{-1}\Big(\alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}(x - x_{\text{center}}) + \kappa_t^{(3)}(x_{\text{young}} - x)^+ + \kappa_t^{(4)}(x - x_{\text{old}})^+ + \gamma_{t-x}\Big)$, where v denotes the parameters $\alpha, \kappa^{(1)}, \kappa^{(2)}, \kappa^{(3)}, \kappa^{(4)}, \gamma$;

$\hat{v} = (\hat{\alpha}, \hat{\kappa}^{(1)}, \hat{\kappa}^{(2)}, \hat{\kappa}^{(3)}, \hat{\kappa}^{(4)}, \hat{\gamma}) \leftarrow$`Fit(`$q$`)`;

for $s = 1, \dots, n_s$ **do**

  foreach x *and* $t \le t_{\max}$ **do** $q_{x,t,s} \leftarrow q_{x,t}$;

  $v \leftarrow$`Forecast(`$\hat{v}, t_{\text{fc}}$, view`)`;

  foreach x *and* $t > t_{\max}$ **do** $q_{x,t,s} \leftarrow f_v(x,t)$;

$\hat{v} \leftarrow$`Forecast(`$\hat{v}, t_{\text{fc}}$, det`)`;

foreach x *and* $t > t_{\max}$ **do** $\hat{q}_{x,t} \leftarrow f_{\hat{v}}(x,t)$;

end

Algorithm 2: Parameter fitting

Input: Mortality rates $q_{x,t}$ for $x = x_{\min}, \ldots, x_{\max}, t = t_{\min}, \ldots, t_{\max}$ (these age and year ranges are assumed to be implicitly defined by the $q_{x,t}$ received as an argument)

Output: Estimated parameters $\alpha, \kappa^{(1)}, \kappa^{(2)}, \kappa^{(3)}, \kappa^{(4)}, \gamma$

`Fit(`q`)` **begin**

$n_x \leftarrow x_{\max} - x_{\min} + 1$, $n_t \leftarrow t_{\max} - t_{\min} + 1$;

$n_v \leftarrow n_x + 5n_t - 1 - c_{\text{excl}} - (x_{\text{cutoff}} - x_{\min})$;

foreach x **do** $\alpha_x \leftarrow n_t^{-1} \sum_{t=t_{\min}}^{t_{\max}} \operatorname{logit} q_{x,t}$;

set $r \in \mathbb{R}^{n_x n_t}, M \in \mathbb{R}^{n_x n_t \times n_v}, v \in \mathbb{R}^{n_v}$ according to Equation (3), i.e. response vector r (with offset induced by α_x), predictor matrix M, and parameter vector v;

solve the corresponding generalized linear model $r \approx Mv$ for v by iterative reweighted least squares to obtain parameters $\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \kappa_t^{(4)}$ and γ_{t-x} for each age x and year t;

$\varphi_1 \leftarrow$ slope of regression line to α_x for $x \in \{x_{\text{young}}, \ldots, x_{\text{old}}\}$;

foreach t **do** $\kappa_t^{(2)} \leftarrow \kappa_t^{(2)} + \varphi_1$;

foreach t **do** $\kappa_t^{(1)} \leftarrow \kappa_t^{(1)} + \alpha_{x_{\text{center}}}$;

foreach x **do**

$\alpha_x \leftarrow \Big(\sum_{t'=t_{\min}}^{t_{\max}} w_{t'} \Big)^{-1} \sum_{t=t_{\min}}^{t_{\max}} w_t \Big(\operatorname{logit}(q_{x,t}) - f_v(x,t) + \alpha_x \Big)$;

end

Algorithm 3: Parameter forecasting

Input: Parameters $\alpha, \kappa^{(1)}, \kappa^{(2)}, \kappa^{(3)}, \kappa^{(4)}, \gamma$, year $t_{\mathrm{fc}} > t_{\max}$ up to which forecasts are generated (note that $t_{\max}$ is assumed to be implicitly defined by the κ vectors), and an argument det,one-year or run-off to specify deterministic, one-year view or run-off view outputs

Output: One realization of parameter forecasts

`Forecast(`$\alpha, \kappa^{(1)}, \kappa^{(2)}, \kappa^{(3)}, \kappa^{(4)}, \gamma, t_{\mathrm{fc}}$, det/stoch`)` **begin**

- define function $f_v\colon (x,t) \mapsto \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}(x - x_{\mathrm{center}})$ $+\kappa_t^{(3)}(x_{\mathrm{young}} - x)^+ + \kappa_t^{(4)}(x - x_{\mathrm{old}})^+ + \gamma_{t-x}$;
- $\ell_t \leftarrow$ regression line to $\kappa_{t_{\min}}^{(1)}, \ldots, \kappa_t^{(1)}$ for $t = t_{\min} + 1, \ldots, t_{\max}$ with respect to weights w;
- $\sigma^{(1)} \leftarrow$ sample standard deviation of the errors $\ell_{t-1}(t) - \kappa_t^{(1)}$ for $t = t_{\min} + 2, \ldots, t_{\max}$ weighted by w^*;
- estimate parameters $a^{(\gamma)}, b^{(\gamma)}, \sigma^{(\gamma)}$ of $\gamma_{t-x} = a^{(\gamma)} + b^{(\gamma)}\gamma_{t-x-1} + \mathcal{N}(0, \sigma^{(\gamma)})$ (truncation zeros are ignored);
- $\sigma^{(2)} \leftarrow$ sample standard deviation of $\kappa_t^{(2)} - \kappa_{t-1}^{(2)}$;
- $C \leftarrow$ covariance matrix of the three innovation vectors $\kappa_t^{(2)} - \kappa_{t-1}^{(2)}, \kappa_t^{(3)} - \kappa_{t-1}^{(3)}, \kappa_t^{(4)} - \kappa_{t-1}^{(4)}$;
- $D \leftarrow \mathrm{diag}(1 + (\overline{\sigma}^{(2)})^2/\sqrt{C_{1,1}}, 1, 1)$;
- $C \leftarrow$ Cholesky decomposition of DCD;
- **if** *deterministic outputs are requested* **then**
  - **for** $t = t_{\max} + 1, \ldots, t_{\mathrm{fc}}$ **do**
    - $\kappa_t^{(1)} \leftarrow \ell_{t_{\max}}(t)$, $\kappa_t^{(2)} \leftarrow \kappa_{t-1}^{(2)}$, $\kappa_t^{(3)} \leftarrow \kappa_{t-1}^{(3)}$;
    - $\kappa_t^{(4)} \leftarrow \kappa_{t-1}^{(4)}$, $\gamma_{t-x_{\min}} \leftarrow 0$;
- **else**
  - **for** $t = t_{\max} + 1, \ldots, t_{\mathrm{fc}}$ **do**
    - **if** $t = t_{\max} + 1$ *or run-off view is requested* **then** draw $\varepsilon^{(1)}$, $\varepsilon^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)}$ i.i. from $\mathcal{N}(0,1)$ **else** $\varepsilon^{(1)}, \varepsilon^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)} \leftarrow 0$;
    - $(\varepsilon^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)}) \leftarrow (\varepsilon^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)})C$;
    - $\kappa_t^{(1)} = \ell_{t-1}(t) + \varepsilon^{(1)}(\overline{\sigma}^{(1)} + \sigma^{(1)})$;
    - $\kappa_t^{(2)} \leftarrow \kappa_{t-1}^{(2)} + \varepsilon^{(2)}$, $\kappa_t^{(3)} \leftarrow \kappa_{t-1}^{(3)} + \varepsilon^{(3)}$, $\kappa_t^{(4)} \leftarrow \kappa_{t-1}^{(4)} + \varepsilon^{(4)}$;
    - $\ell_t \leftarrow$ regression line to $\kappa_{t_{\min}}^{(1)}, \ldots, \kappa_t^{(1)}$ with respect to w;
  - **for** $t = t_{\max} + 1 - x_{\mathrm{cutoff}} + x_{\min}, \ldots, t_{\mathrm{fc}}$ **do**
    - **if** *run-off view is requested* **then** draw $\varepsilon^{(\gamma)}$ from $\mathcal{N}(0,1)$ **else** $\varepsilon^{(\gamma)} \leftarrow 0$ $\gamma_{t-x_{\min}} \leftarrow \min\{b^{(\gamma)}, 1\}\gamma_{t-1-\min} + \varepsilon^{(\gamma)}\sigma^{(\gamma)}$ (note that the γ actually are linearly faded in within $x_{\mathrm{cutoff}} - x_{\min}$ years to avoid a jump at $t_{\max}$ – also in the deterministic case – we omit the corresponding assignments for better readability);

end

Algorithm 4: Main algorithm to simultaneously compute stochastic future mortality rates of several populations

Input: Historical mortality rates $q_{x,t,p}$ of given populations $p \in P$ for $x = x_{\min}, \dots, x_{\max}, t = t_{\min}, \dots, t_{\max}$, year $t_{\text{fc}} > t_{\max}$ up to which forecasts are generated, argument view to calculate one-year view or run-off view

Output: Realizations of future mortality rates $q_{x,t,p,s}$ and future best estimate mortality rates $\hat{q}_{x,t,p}$ for $p \in P, x = x_{\min}, \dots, x_{\max}$, $t = t_{\max} + 1, \dots, t_{\text{fc}}$ and $s = 1, \dots, n_s$

`Main`$(q, t_{\text{fc}}, \text{view})$ **begin**

- set function $f_v \colon (x,t) \mapsto \text{logit}^{-1}\Big(\alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)}(x - x_{\text{center}}) + \kappa_t^{(3)}(x_{\text{young}} - x)^+ + \kappa_t^{(4)}(x - x_{\text{old}})^+ + \gamma_{t-x}\Big)$;
- $q_{\cdot,\cdot,P} \leftarrow$ combined mortality rates of all $p \in P$;
- $\hat{v}_P = (\hat{\alpha}_{\cdot,P}, \hat{\kappa}^{(1)}_{\cdot,P}, \hat{\kappa}^{(2)}_{\cdot,P}, \hat{\kappa}^{(3)}_{\cdot,P}, \hat{\kappa}^{(4)}_{\cdot,P}, \hat{\gamma}_{\cdot,P}) \leftarrow$ `Fit`$(q_{\cdot,\cdot,P})$;
- **foreach** $p \in P$ **do**
  - $\hat{v}_p = (\hat{\alpha}_{\cdot,p}, \hat{\kappa}^{(1)}_{\cdot,p}, \hat{\kappa}^{(2)}_{\cdot,p}, \hat{\kappa}^{(3)}_{\cdot,p}, \hat{\kappa}^{(4)}_{\cdot,p}, \hat{\gamma}_{\cdot,p}) \leftarrow$ `Fit`$(q_{\cdot,\cdot,p})$;
- **for** $s = 1, \dots, n_s$ **do**
  - **foreach** x *and* $t \le t_{\max}$ **do** $q_{x,t,P,s} \leftarrow q_{x,t,P}$;
  - **foreach** $p \in P, x$ *and* $t \le t_{\max}$ **do** $q_{x,t,p,s} \leftarrow q_{x,t,p}$;
  - $v_P \leftarrow$ `Forecast`$(\hat{v}_P, t_{\text{fc}}, \text{view})$;
  - $(v_p)_{p \in P} \leftarrow$ `Multi-forecast`$((\hat{v}_p)_{p \in P}, v_P, t_{\text{fc}}, \text{view})$;
  - **foreach** x, p *and* $t > t_{\max}$ **do** $q_{x,t,p,s} \leftarrow f_{v_p}(x,t)$;
- $(\hat{v}_p)_{p \in P} \leftarrow$ `Multi-forecast`$((\hat{v}_p)_{p \in P}, \hat{v}_P, t_{\text{fc}}, \text{det})$;
- **foreach** x, p *and* $t > t_{\max}$ **do** $\hat{q}_{x,t,p} \leftarrow f_{\hat{v}_p}(x,t)$;

end

Algorithm 5: Parameter forecasting for several populations

Input: Analogous to Algorithm 3 for each population p and P

Output: One realization of parameter forecasts for each population p

`Multi-forecast(`$(v_p)_{p\in P}, v_P, t_{\text{fc}}, \text{det/stoch}$`)` **begin**

- $C^{(1)} \leftarrow$ Cholesky decomposition of corr $\left(\kappa_{t,p}^{(1)} - \ell_{t-1,p}(t)\right)$, i.e. the correlation matrix of $\kappa_{t,p}^{(1)} - \ell_{t-1,p}(t)$ across populations;
- $C^{(2)} \leftarrow$ Cholesky decomposition of corr $\left(\kappa_{t,p}^{(2)} - \kappa_{t-1,p}^{(2)}\right)$;
- **foreach** $p \in P$ **do**
  - estimate parameters $a_p^{(1)}, b_p^{(1)}, \sigma_p^{(1)}$ of $\kappa_{t,p}^{(1)} = \kappa_{t,P}^{(1)} + a_p^{(1)} + b_p^{(1)}(\kappa_{t-1,p}^{(1)} - \kappa_{t-1,P}^{(1)}) + \mathcal{N}(0, \sigma_p^{(1)})$ (modified according to Section 2.5 if necessary);
  - estimate parameters $a_p^{(\gamma)}, b_p^{(\gamma)}, \sigma_p^{(\gamma)}$ of $\gamma_{t-x,p} = a_p^{(\gamma)} + b_p^{(\gamma)}\gamma_{t-1-x,p} + \mathcal{N}(0, \sigma_p^{(\gamma)})$;
  - $\sigma_p^{(2)} \leftarrow$ sample standard deviation of $\kappa_{t,p}^{(2)} - \kappa_{t-1,p}^{(2)}$;
  - $C_p \leftarrow$ covariance matrix of the three innovation vectors $\kappa_{t,p}^{(2)} - \kappa_{t-1,p}^{(2)}, \kappa_{t,p}^{(3)} - \kappa_{t-1,p}^{(3)}, \kappa_{t,p}^{(4)} - \kappa_{t-1,p}^{(4)}$;
  - $D \leftarrow \text{diag}(1 + (\overline{\sigma}_p^{(2)})^2/\sqrt{C_{p,1,1}}, 1, 1)$;
  - $C_p \leftarrow$ Cholesky decomposition of DC_pD;
- **if** *deterministic outputs are requested* **then**
  - **for** $t = t_{\max} + 1, \ldots, t_{\text{fc}}$ **do**
    - $\kappa_{t,p}^{(1)} \leftarrow \kappa_{t,P}^{(1)} + a_p^{(1)} + b_p^{(1)}(\kappa_{t-1,p}^{(1)} - \kappa_{t-1,P}^{(1)})$, $\kappa_{t,p}^{(2)} \leftarrow \kappa_{t-1,p}^{(2)}$;
    - $\kappa_{t,p}^{(3)} \leftarrow \kappa_{t-1,p}^{(3)}$, $\kappa_{t,p}^{(4)} \leftarrow \kappa_{t-1,p}^{(4)}$, $\gamma_{t-x_{\min},p} \leftarrow 0$;
- **else**
  - **for** $t = t_{\max} + 1, \ldots, t_{\text{fc}}$ **do**
    - draw $(\varepsilon_p^{(1)}, \varepsilon_p^{(2)})_{p\in P}$ i.i. $(0,1)$-normally distributed;
    - $(\varepsilon_p^{(1)})_{p\in P} \leftarrow (\varepsilon_p^{(1)})_{p\in P}C^{(1)}$, $(\varepsilon_p^{(2)})_{p\in P} \leftarrow (\varepsilon_p^{(2)})_{p\in P}C^{(2)}$, where $(\varepsilon_p^{(1)})_{p\in P}$ denotes the row vector of $\varepsilon_p^{(1)}$ by populations;
    - **foreach** $p \in P$ **do**
      - draw $\varepsilon^{(3)}, \varepsilon^{(4)}$ i.i. $(0,1)$-normally distributed;
      - $(\varepsilon_p^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)}) \leftarrow (\varepsilon_p^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)})C_p$;
      - $\kappa_{t,p}^{(1)} \leftarrow \kappa_{t,P}^{(1)} + a_p^{(1)} + b_p^{(1)}(\kappa_{t-1,p}^{(1)} - \kappa_{t-1,P}^{(1)}) + \varepsilon_p^{(1)}$
      - $\kappa_{t,p}^{(2)} \leftarrow \kappa_{t-1,p}^{(2)} + \varepsilon_p^{(2)}$, $\kappa_{t,p}^{(3)} \leftarrow \kappa_{t-1,p}^{(3)} + \varepsilon^{(3)}$,
      - $\kappa_{t,p}^{(4)} \leftarrow \kappa_{t-1,p}^{(4)} + \varepsilon^{(4)}$;
    - **for** $t = t_{\max} + 1 - x_{\text{cutoff}} + x_{\min}, \ldots, t_{\text{fc}}$ **do**
      - draw $\varepsilon^{(\gamma)}$ from $\mathcal{N}(0,1)$;
      - $\gamma_{t-x_{\min}} \leftarrow \min\{b^{(\gamma)}, 0.97\}\gamma_{t-1-x_{\min}} + \varepsilon^{(\gamma)}\sigma^{(\gamma)}$;

end