

PRICING LONGEVITY-LINKED SECURITIES IN THE PRESENCE OF MORTALITY TREND CHANGES

BY

ARNE FREIMANN

ABSTRACT

Even though the trend in mortality improvements has experienced several permanent changes in the past, the uncertainty regarding future mortality trends is often left unmodeled when pricing longevity-linked securities. In this paper, we present a stochastic modeling framework for the valuation of longevity-linked securities which explicitly considers the risk of random future changes in the long-term mortality trend. We construct a set of meaningful probability distortions which imply equivalent risk-adjusted pricing measures under which the basic model structure is preserved. Inspired by risk-based capital requirements for (re)insurers, we also establish a cost-of-capital pricing approach which then serves as the appropriate reference framework for finding a reasonable range for the market price of longevity risk. In a numerical application, we demonstrate that our model produces plausible risk loadings and show that a greater proportion of the risk loading is allocated to longer maturities when the risk of random future mortality trend changes is adequately modeled.

KEYWORDS

Longevity risk, longevity derivatives, stochastic mortality, mortality trend processes.

JEL codes: G13, G22.

1. INTRODUCTION

The risk that future mortality trends imply stronger mortality improvements than anticipated is commonly referred to as longevity risk and constitutes a major risk for life insurers and pension funds. For its modeling, measurement, and management, academics have proposed a variety of stochastic mortality models. The majority of these models, including the widely used

Astin Bulletin 51(2), 411-447. doi:[10.1017/asb.2021.5](https://doi.org/10.1017/asb.2021.5) © 2021 by Astin Bulletin. All rights reserved. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

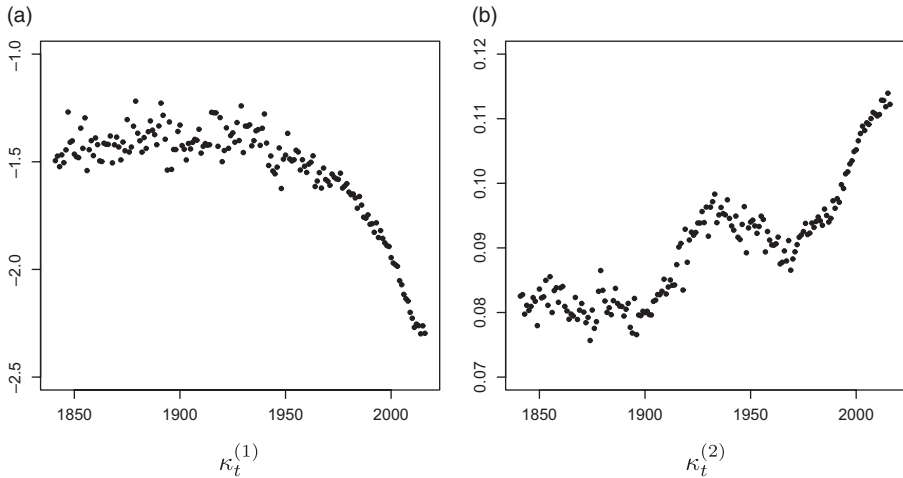


FIGURE 1: Historical period effects for English and Welsh males in the Cairns-Blake-Dowd (CBD) model.

Lee-Carter and Cairns-Blake-Dowd (CBD) model, decompose mortality across the dimensions age, period, and sometimes cohort. For stochastic mortality projections, often a random walk with drift (RWD) is used to project the time-dependent parameters (also called period effects) into the future.

However, historical mortality patterns often reveal structural breaks which are highly unlikely to be generated from a RWD. Figure 1 shows, exemplarily, the historical period effects in the CBD model for English and Welsh males (the model calibration is provided in the appendix). While the first time series clearly exhibits a significant improvement in mortality over the last century, the annual rate of improvement does not appear to be constant over time. This observation can be further substantiated by a Chow test, which provides strong statistical evidence for several break points in both time series (see Appendix A.4). Permanent trend changes in mortality have also been identified by other researchers including Li *et al.* (2011), who uses Zivot and Andrews' test, and Sweeting (2011), who also relies on the Chow test. Therefore, several authors have proposed mortality models that can detect mortality trend changes in the past including Li *et al.* (2011), O'hare and Li (2015), and Van Berkum *et al.* (2016). However, these models have in common that they solely extrapolate the most recent trend into the future without allowing it to change again at future points in time.

However, the fact that the mortality trend has changed occasionally in the past implies the risk of further trend changes in the future. Therefore, some authors have developed stochastic mortality models that explicitly capture the risk of random future mortality trend changes. In their locally linear CBD model, Liu and Li (2016) incorporate 'drift risk' into the widely used RWD setting by assuming the drifts themselves to follow another random walk. Hence,

the prevailing mortality trends, which are represented by the drift terms, gradually vary from year to year according to the realizations of this underlying random walk. They show that their approach yields wider prediction intervals in the long run compared to the original CBD model with fixed drifts reflecting the uncertainty arising from potential future mortality trend changes. Another attempt was made by Sweeting (2011) who argues that the time-dependent parameters in the CBD model for England and Wales do not follow a difference-stationary RWD but appear to be trend-stationary. Therefore, he proposes to model them as random fluctuations around piecewise linear and continuous trends with changing slopes over time. While future trend changes are also assumed to be normally distributed, they do not occur gradually from year to year but only occasionally according to a fixed trend change probability. The latter assumption is arguably more in line with the nature of detected mortality trend changes in the past, which serve as the basis for his model calibration. However, as pointed out by Börger and Schupp (2018), the assumption of normally distributed trend change magnitudes is inconsistent with a calibration approach that only identifies material trend changes in the past. Since the normal distribution has significant probability mass around zero, a rather high variance is required to achieve that simulated trend changes are on average similar in size to historical ones which leads to an overestimation of the variability in future trend changes. Hence, Börger and Schupp (2018) propose a structurally similar model by drawing the absolute trend change magnitudes from a lognormal distribution while assuming the trend change signs to be positive or negative with equal probability. Similarly to Liu and Li (2016), both Sweeting (2011) and Börger and Schupp (2018) demonstrate that their approaches yield wider predictions intervals in the long run compared to the RWD. However, as shown by Börger and Schupp (2018), using normally distributed trend change magnitudes for a trend-stationary model with changing slopes typically results in wider prediction intervals compared to a lognormal distribution for reasons discussed above.

In addition to the question of how to model longevity risk, its securitization and the recent development of the global longevity risk transfer market have attracted huge attention among both academics and practitioners. For this emerging market, a great variety of longevity-linked securities, also called longevity hedges or longevity derivatives, have been proposed in the literature, see Blake *et al.* (2019) for an overview. However, the longevity risk transfer market is still in an early state and illiquid and incomplete. Hence, longevity-linked instruments cannot be priced based on observable market prices, that is, mark-to-market, and the question arises how to put a 'reasonable' value on them. Since stochastic mortality models are normally calibrated to historical mortality data, they do not readily include a risk loading beyond best-estimate values, and further steps in the form of risk adjustments are required to arrive at reasonable and consistent (mark-to-model) prices, which are required under modern risk-based solvency guidelines, see for example Wüthrich (2016). To this end, several pricing techniques have been proposed in the literature, and

we refer to Bauer *et al.* (2010) and Leung *et al.* (2018) for an overview. In this work, we focus on two promising approaches:

- Inspired by risk-neutral valuation of financial derivatives, the *risk-adjusted pricing approach*, often also referred to as *risk-neutral approach*, is widely used, see among others Boyer and Stentoft (2013), Cairns *et al.* (2006), Chen and Cox (2009), and Leung *et al.* (2018). The basic idea is to derive an equivalent risk-adjusted measure under which prices are defined as expected values of discounted future cash flows. Compared to the objective measure, a higher probability is assigned to scenarios of stronger mortality improvements reflecting the risk adjustment made by longevity risk takers.
- Börger (2010) suggests to make inferences on prices for longevity-linked securities from capital requirements for (re)insurers under modern risk-based solvency regimes, such as Solvency II or the Swiss Solvency Test. (Re)insurance companies are required to provide adequate regulatory capital, so-called Solvency Capital Requirements (SCRs), for all risks they are exposed to including longevity. The so-called *cost-of-capital approach* postulates that a market player will only be willing to take longevity risk if the expected return exceeds the expected additional capital charges for taking the risk. Since the market is dominated by reinsurers, this approach seems to be of high practical relevance. It has recently been taken up by Levantesi and Menzietti (2017), who determine the relevant 99.5% Value at Risk (VaR) over a multi-year horizon and by Zeddouk and Devolder (2019), who establish a setup that is more in line with the 1-year view of Solvency II and compare the cost-of-capital approach with alternative approaches.

Naturally, mark-to-model valuation of longevity-linked securities relies crucially on the suitability of the underlying stochastic mortality model for the given purpose and on the adequacy of its mortality forecasts. As discussed above, the occurrence of several mortality trend changes in the past clearly calls for stochastic mortality trend modeling, which typically comes at the cost of higher model complexity. However, using a more sophisticated mortality model appears justified considering that disregarding the risk of future trend changes seems inconsistent with historical mortality patterns, might significantly understate the actual amount of risk taken, and might eventually lead to insufficient risk premiums and to inadequate prices.

In the current paper, we address the valuation of longevity-linked securities in a stochastic modeling framework which explicitly considers the risk of permanent random future mortality trend changes. At the core of our framework is a simulation model for the underlying long-term mortality trend of a reference population (typically the national population) that builds on previous works of Börger and Schupp (2018) and Schupp (2019). Furthermore, we allow for portfolio-specific mortality characteristics by considering systematic mortality differentials between selected subpopulations and the reference population as well as unsystematic fluctuations in portfolios of limited size.

The model is therefore applicable to customized as well as to standardized instruments that are linked to the mortality indices of the reference population. Within this framework, we establish the above-introduced approaches for determining adequate longevity risk premiums: first, we apply the technique of multivariate normalized exponential tilting, which is based on the widely used Wang distortion operator, to construct a set of stochastic discount factors, also called deflators, which imply equivalent risk-adjusted pricing measures under which the basic model structure is preserved. In the second step, we rely on the cost-of-capital approach to find a reasonable range for the market price of longevity risk. The associated derivation of longevity SCRs consistently with regulatory requirements under Solvency II requires to model the impact of new mortality information on best-estimate mortality assumptions over a 1-year horizon. To this end, we follow Börger *et al.* (2019b) and clearly distinguish between two mortality trends: the actual but unobservable mortality trend prevailing at a certain point in time and the trend an observer would estimate given the observed mortality patterns in previous years. Best-estimate mortality assumptions for Solvency Capital Requirement (SCR) computations are derived with the latter, whereas the former is used to generate sample paths of future mortality. We discuss, calibrate, and apply the presented approaches to price longevity swaps as examples for widely used longevity-linked instruments.

Our work contributes to the literature on longevity risk pricing in the following ways: first and foremost, we explicitly include the major risk of random permanent future mortality trend changes in the mark-to-model valuation of longevity hedges. In light of the potential threat posed by an unprecedented change in the long-term mortality trend to the solvency of longevity risk bearers, our modeling framework constitutes a valuable alternative to existing approaches. Second, we contribute to the ongoing debate of how to find reasonable market prices of longevity risk in an incomplete market by relying on regulatory capital charges for longevity risk as an appropriate reference framework. In particular, we derive (future) capital charges for longevity risk consistently with regulatory capital requirements under Solvency II and properly account for its stochastic nature. This approach has the key advantage of giving a reasonable range for the market price of longevity risk without the need for market price data, which are currently not publicly available.

The remainder of this paper is organized as follows. Section 2 introduces our stochastic modeling framework. In particular, we describe the simulation of future mortality trends and the pathwise derivation of best-estimate mortality assumptions at future valuation dates for SCR computations. The calibration of the model to the historical mortality experience of English and Welsh males is provided in the appendix. In Section 3, we construct a set of equivalent pricing measures and establish the cost-of-capital approach in the context of a regulatory capital model under Solvency II. In a numerical application in Section 4, we calibrate both pricing approaches, derive risk-adjusted

forward rates for customized and index-based longevity swaps, and compare our results with alternative modeling and pricing approaches. Finally, Section 5 concludes.

2. STOCHASTIC MORTALITY TREND MODEL

As outlined in the introduction, we now describe a stochastic modeling framework which explicitly accounts for random future changes in the long-term mortality trend. Following Börger *et al.* (2019b), we clearly distinguish between the actual (but unobservable) mortality trend (AMT) prevailing at a certain point in time and the estimated mortality trend (EMT) that an observer would estimate given the observed mortality evolution up to that point in time. Hence, our model consists of two components: a model for the stochastic development of the AMT, which is presented in Section 2.2, and a model for the derivation of the EMT, which is presented in Section 2.3. The latter is required for deriving best-estimate mortality assumptions for SCR computations, whereas the former is used to generate sample paths of future mortality. To account for portfolio-specific mortality characteristics, we extend the combined AMT/EMT setup of Börger *et al.* (2019b) to multiple populations.

2.1. Multi-population model structure

Our modeling framework builds on a multi-population extension of the CBD model structure as originally proposed by Cairns *et al.* (2006). We assume a *reference population* for which the logit of the 1-year death probabilities is parametrized as:

$$\text{logit} \left(q_{x,t}^{[R]} \right) := \log \left(\frac{q_{x,t}^{[R]}}{1 - q_{x,t}^{[R]}} \right) = \kappa_t^{(1)[R]} + (x - \bar{x}) \kappa_t^{(2)[R]},$$

where \bar{x} denotes the average age of the calibration age range. The first period effect $\kappa_t^{(1)[R]}$ captures the general level of mortality over time, while the second time-dependent parameter $\kappa_t^{(2)[R]}$ describes the slope of the mortality line in the logit plot.

Relative to the reference population, we model mortality differentials for $N_{Sub} \geq 1$ distinct subpopulations, for example of different socioeconomic status, also in logit-bilinear form, that is,

$$\text{logit} \left(q_{x,t}^{[p]} \right) - \text{logit} \left(q_{x,t}^{[R]} \right) = \kappa_t^{(1)[p]} + (x - \bar{x}) \kappa_t^{(2)[p]}, \quad p \in \{1, \dots, N_{Sub}\},$$

where \bar{x} is the same for all subpopulations and chosen analogously to the reference population. While the time-dependent parameters for the reference

population represent the common long-term mortality trend of all (sub-) populations, $\kappa_t^{(1)[p]}$ and $\kappa_t^{(2)[p]}$ account for subpopulation-specific mortality differentials. Relative modeling approaches of this kind are widely used in the field of multi-population mortality modeling, see for example Villegas *et al.* (2017).

2.2. AMT model component

For stochastic mortality projections, we clearly distinguish between the following components of longevity risk: the long-term mortality trend risk of the overall population, systematic subpopulation-specific mortality differentials, and unsystematic small sample risk in portfolios of limited size. This rigorous differentiation between the three risk drivers will later allow to assign a risk premium to each longevity risk component, which in turn allows to price index-based instruments that solely cover the risk originating from the reference population.

2.2.1. Long-term mortality trend risk

For the projection of the time-dependent parameters of the reference population, we follow Börger and Schupp (2018) and rely on a trend-stationary stochastic process with changing slopes over time. The assumption of trend-stationarity (as opposed to difference-stationarity) for the reference population is further substantiated by a Dickey–Fuller test in Appendix A.4. Hence, the observable period effects are modeled as random fluctuations around an underlying unobservable piecewise linear mortality process, that is,

$$\kappa_t^{(i)[R]} = \hat{\kappa}_t^{(i)[R]} + \epsilon_t^{(i)[R]}, \quad i = 1, 2,$$

where $\epsilon_t^{[R]} := (\epsilon_t^{(1)[R]}, \epsilon_t^{(2)[R]})'$ follows a two-dimensional normal distribution with mean zero and constant covariance matrix $\Sigma^{[R]}$. It accounts for systematic fluctuations in mortality, for example due to a mild winter or a flue wave. More sophisticated techniques to account for transitory mortality shocks (e.g., due to a pandemic) are proposed by Chen and Cox (2009) but do not seem justified for our purposes. The actual underlying mortality processes are then projected linearly as:

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

where the slope $\hat{d}_t^{(i)[R]}$, $i = 1, 2$ is interpreted as the AMT prevailing at time t . We follow Börger and Schupp (2018) and update the AMTs independently of each other as follows:

$$\hat{d}_t^{(i)[R]} = \hat{d}_{t-1}^{(i)[R]} + O_t^{(i)[R]} S_t^{(i)[R]} M_t^{(i)[R]}, \quad i = 1, 2$$

based on the following mutually independent random variables:

- The Bernoulli-distributed random variable $O_t^{(i)[R]} \in \{0, 1\}$ indicates whether a permanent trend change occurs between $t - 1$ and t or not. The trend change probability is denoted as $p^{(i)[R]} > 0$.
- Given a trend change occurs, $S_t^{(i)[R]} \in \{-1, 1\}$ determines its sign, where we assume positive or negative trend changes to be equally likely under the objective measure.
- Finally, $M_t^{(i)[R]} > 0$ denotes the absolute trend change magnitude, which is modeled by a lognormal distribution with parameters $\mu_M^{(i)[R]}$ and $\sigma_M^{(i)[R]2}$.

Once a trend change occurs, the AMT changes permanently according to the realized trend change sign and the lognormally distributed trend change magnitude. As argued by Börger and Schupp (2018), the above decomposition of the trend change intensities offers some desirable properties. Analogously to the models of Liu and Li (2016) and Sweeting (2011), a symmetric distribution for the trend change intensities assures that the prevailing AMT (even though unobservable) always represents the best-estimate trend for any future point in time. However, using a heavy-tailed lognormal distribution (instead of a normal distribution with significant mass around zero) produces rather significant trend changes, which is arguably in line with the nature of material trend changes in the past.

Since data on historical mortality trend changes are sparse, the calibration of stochastic mortality trend processes typically involves a considerable amount of parameter uncertainty. Unlike Liu and Li (2016) and Sweeting (2011), we explicitly account for parameter uncertainty in the starting values of the simulation as well as in the trend change parameters by sampling them at the start of each simulation from suitable distributions, see Appendix A.3 for details.

2.2.2. Subpopulation-specific mortality differentials

In the field of multi-population mortality modeling, vector autoregressive integrated moving average (VARIMA) models are commonly used to project mortality differentials between closely related populations. In Appendix A.5, we find that the first-order integrated VARIMA model is suitable for our data set and that a multivariate RWD in particular provides the best compromise between a satisfactory fit and a parsimonious model structure in terms of the Bayesian information criterion (BIC). Hence, we jointly model the vectors of subpopulation-specific period effects $\kappa_t^{[p]} := (\kappa_t^{(1)[p]}, \kappa_t^{(2)[p]})'$, $p \in \{1, \dots, N_{Sub}\}$ as:

$$\kappa_{t+1}^{[p]} = \mu^{[p]} + \kappa_t^{[p]} + \epsilon_t^{[p]}, \quad p \in \{1, \dots, N_{Sub}\},$$

where $\mu^{[p]} \in \mathbb{R}^2$ denotes the drift and $\epsilon_t^{[p]} \in \mathbb{R}^2$ denotes the annual random innovations for subpopulation $p \in \{1, \dots, N_{Sub}\}$. The joint vector $\epsilon_t^{[Sub]} := (\epsilon_t^{[1]}, \dots, \epsilon_t^{[N_{Sub}]})' \in \mathbb{R}^{2N_{Sub}}$ of annual innovations follows a multivariate normal

distribution with mean zero and covariance matrix $\Sigma^{[Sub]} \in \mathbb{R}^{(2N_{Sub}) \times (2N_{Sub})}$. We follow Villegas *et al.* (2017) and model these subpopulation-specific mortality differentials independently of the stochastic trend processes for the reference population. The latter serve as the long-term mortality trend for all subpopulations, which, by construction, implies a reasonable correlation between the subpopulations. Note that a multivariate RWD has already been used in previous studies for modeling socioeconomic mortality differentials for England, see for example Villegas and Haberman (2014).

We would like to stress that the RWD can easily be replaced by alternative time series processes if desired for other data sets or model purposes. For instance, often mean-reverting processes (such as the first-order vector autoregressive (VAR) process) are used to enforce ‘coherent’, that is, non-diverging, mortality rates between closely related populations in the long run, see for example Villegas *et al.* (2017). For a comparison of different modeling assumptions and a critical discussion on the concept of ‘coherence’, we refer to Li *et al.* (2017).

2.2.3. *Unsystematic small sample risk*

Finally, idiosyncratic risk in a portfolio of limited size is accounted for by sampling survivors over time from a Binomial distribution. In each model path, we draw realizations for survivors aged $x + 1$ at time $t + 1$ coming from subpopulation p from the following Binomial distribution:

$$B_{x+1,t+1}^{[p]} \sim \text{Binom} \left(B_{x,t}^{[p]}, 1 - q_{x,t+1}^{[p]} \right), \quad p \in \{1, \dots, N_{Sub}\},$$

given $B_{x,t}^{[p]}$ survivors aged x still alive at time t .

2.3. **EMT model component**

For deriving longevity SCRs consistently with regulatory requirements under Solvency II, the framework needs to model the impact of new mortality information on best-estimate mortality assumptions at future points in time (see Section 3.2 for details). As argued by Börger *et al.* (2019b), an observer at some future point in time T is generally not able to clearly distinguish between a recent change in the AMT and a ‘normal’ random fluctuation around it. For the best-estimate mortality projection starting at time T , an actuary would calibrate a mortality model to the most recent available mortality data. For simplicity, we assume recalibration of the same CBD model structure to sufficiently large populations, so that unsystematic variations in the data are negligible. When an additional year of data becomes available, recalibrating the multi-population CBD model to the extended data set would not affect the period effects in previous years. This is also known as ‘new-data-invariant property’ (cf. Tan *et al.*, 2014) and allows to directly derive the updated best-estimate mortality assumptions based on the realized time processes up to time

T , which makes the modeling framework highly efficient in practical applications. Note that for many other models, for instance for the Lee–Carter model, a full computationally expensive model recalibration would be required.

In our multi-population setting, this pathwise derivation of best-estimate mortality consists of two steps: estimating the AMT for the reference population followed by an adjustment for potentially differing mortality levels and trends for the subpopulations.

2.3.1. Reference population

For the best-estimate mortality projection for the reference population beyond time T , we follow a deterministic central path of mortality under the CBD model, that is,

$$\begin{aligned} \text{logit} \left(\tilde{q}_{x,t}^{[R]}(T) \right) &= \tilde{\kappa}_T^{(1)[R]} + (t - T) \tilde{d}_T^{(1)[R]} \\ &\quad + (x - \bar{x}) \left(\tilde{\kappa}_T^{(2)[R]} + (t - T) \tilde{d}_T^{(2)[R]} \right), \quad t > T, \end{aligned}$$

where $\tilde{\kappa}_T^{(i)[R]}$, $i = 1, 2$ denotes the prevailing mortality level and $\tilde{d}_T^{(i)[R]}$, $i = 1, 2$ denotes the current EMT. Using a deterministic projection of mortality rates emulates common actuarial practice and is frequently used as a pragmatic alternative to deriving the conditional mean (which would require a computationally expensive additional layer of inner simulations), see for example Cairns *et al.* (2014) or Börger *et al.* (2019b). The parameters $\tilde{\kappa}_T^{(i)[R]}$ and $\tilde{d}_T^{(i)[R]}$, $i = 1, 2$ need to be estimated based on the observed mortality patterns in previous years. To this end, we follow Börger *et al.* (2019b) by applying a weighted linear regression on the most recent data points based on the following exponentially decaying weights:

$$w^{(i)}(t, T) := \frac{1}{(1 + 1/\psi^{(i)})^{T-t}}, \quad i = 1, 2, \quad t \leq T.$$

The weighting parameters $\psi^{(i)} \geq 0$, $i = 1, 2$ need to be specified to find a reasonable trade-off between avoiding an overreaction of the EMT to ‘normal’ random fluctuations around the AMT on the one hand and ensuring a prompt detection of a recent mortality trend change on the other hand. The first objective can be achieved by including enough data points, whereas the second objective requires sufficient weight on the most recent data points, which naturally contain the most relevant information. We numerically determine the optimal weighting by minimizing the mean squared errors between the unobservable AMT and the derived EMT, see Appendix A.6 for details. We would like to stress that exponential weighting guarantees that the updated EMT will remain unchanged if the new data point realizes exactly as expected under the previous EMT.

2.3.2. *Subpopulations*

Given best-estimate mortality for the reference population, subpopulation-specific rates are projected beyond time T by means of the following adjusted central path of mortality:

$$\tilde{q}_{x,t}^{[p]}(T) := \text{logit}^{-1} \left(\text{logit} \left(\tilde{q}_{x,t}^{[R]}(T) \right) + \tilde{\kappa}_T^{(1)[p]} + (t - T) \tilde{v}_T^{(1)[p]} \right. \\ \left. + (x - \bar{x}) \left(\tilde{\kappa}_T^{(2)[p]} + (t - T) \tilde{v}_T^{(2)[p]} \right) \right)$$

for each subpopulation $p \in \{1, \dots, N_{Sub}\}$, where the subpopulation-specific adjustment terms $\tilde{\kappa}_T^{(i)[p]}, i = 1, 2$ account for differing mortality levels and $\tilde{v}_T^{(i)[p]}, i = 1, 2$ capture differing mortality trends relative to the reference population. For the sake of consistency, we derive these quantities by applying the same weighted linear regression (including the same weights) as for the reference population. Due to the properties of the weighted linear regression, this ensures ‘order-invariance’ in the sense that the above two-step estimation approach produces the same best-estimate mortality projection as directly applying a weighted linear regression to the overall subpopulation-specific period effects $\kappa_t^{(i)[R]} + \kappa_t^{(i)[p]}, i = 1, 2$.

3. PRICING APPROACHES

We consider a market participant, referred to as the *risk taker*, who is willing to take on longevity risk by issuing longevity-linked securities at time zero. To focus on longevity risk, we do not deal with operational risk or counterparty credit risk and assume independence between mortality and interest rates. The time-zero random present value of all future instrument cash flows to the counterparty, also referred to as the *hedger*, is denoted as:

$$H(0) := \sum_{t>0}^{\tau} B(t)^{-1} CF_H(t),$$

where τ denotes the contract maturity, $(B(t))_{t \geq 0}$ denotes the value process of a risk-free bank account with $B(0) := 1$, and $CF_H(t), 0 < t \leq \tau$ denotes the cash flow at time t , where a positive cash flow represents a payment from the risk taker to the hedger. In exchange, the risk taker demands the instrument’s objective best-estimate value plus a risk premium. Depending on the instrument’s payout structure, either a single contract premium is charged at inception or the risk premium is directly included in the payoff structure by means of risk-adjusted forward rates, so that no payment exchanges hands at inception. In either case, the risk premium demanded on top of the objective best-estimate value needs to be determined.

We now establish two pricing methods for deriving reasonable risk premiums in the previously introduced modeling framework: the risk-adjusted

pricing approach based on a pricing measure in Section 3.1 and the cost-of-capital approach in Section 3.2. In Section 3.3, we briefly discuss and qualitatively compare both approaches. By way of illustration, we will price longevity swaps in Section 4 assuming constant deterministic interest rates, that is, $B(t) := (1+r)^t$, $t \geq 1$ for some constant interest rate r , although we avoid any loss of generality in the approaches presented in this section.

3.1. Risk-adjusted measure

A widely used approach for pricing longevity-linked instruments is to derive a risk-adjusted measure \mathbb{Q} , which is equivalent to the objective (or real-world) measure \mathbb{P} , under which prices are defined as expected values of discounted future cash flows. For forward-type instruments, the forward rates are typically determined at time zero, so that no payment exchanges hands at inception, that is, $\mathbb{E}^{\mathbb{Q}}(H(0)) = 0$. Unlike many financial assets, the underlying longevity risk drivers are not continuously tradeable. Hence, longevity-linked payments cannot be replicated and a unique risk-adjusted measure cannot be deduced based on the principle of no-arbitrage. Instead, we rely on the concept of *deflators*, also called *state price densities*, to construct an equivalent risk-adjusted measure. In our incomplete market setting in discrete time, prices can be represented in the following equivalent form:

$$\mathbb{E}^{\mathbb{Q}} \left(\sum_{t>0}^{\tau} CF_H(t) B(t)^{-1} \right) = \mathbb{E}^{\mathbb{P}} \left(\sum_{t>0}^{\tau} CF_H(t) \prod_{s=1}^t \varphi_s \right),$$

where the random variables $\varphi_s \gg 0$, $s \geq 1$, commonly referred to as *span-deflators*, play the role of stochastic discount factors that assign a value at time $s-1$ to random cash flows that occur at time s . Hence, there is a one-to-one correspondence between risk-adjusted valuation and stochastic discounting and the task of constructing an appropriate pricing measure is equivalent to specifying reasonable span-deflators. For the theoretical foundations of market-consistent actuarial valuation, we refer to Section 2 in Wüthrich (2016).

In our stochastic mortality trend model, longevity-linked payments depend on multiple risk drivers simultaneously, namely the most prominent risks of unpredictable occurrence (O), sign (S), and magnitude (M) of future trend changes, as well as annual fluctuations (ϵ) around the AMT, and subpopulation-specific mortality differentials (Sub). Given stochastically independent risk drivers, we define the following multiplicative span-deflators:

$$\varphi_t := \varphi_t^r \varphi_t^O \varphi_t^S \varphi_t^M \varphi_t^\epsilon \varphi_t^{Sub}, \quad t \geq 1,$$

where φ_t^r , $t \geq 1$ plays the role of interest rate discounting, for example $\varphi_t^r := (1+r)^{-1}$ in a constant interest rate environment, and φ_t^{Risk} , $t \geq 1$ denotes the span-deflators for the longevity risk driver $Risk \in \{O, S, M, \epsilon, Sub\}$, which need to be carefully specified, cf. Section 2.6.2 in Wüthrich (2016). Loosely speaking,

meaningful span-deflators for longevity risk pricing should take high values in states that are less favorable for longevity risk takers and low values in states of higher mortality. Therefore, they imply an equivalent pricing measure that assigns, compared to the objective measure, higher probability mass to scenarios of stronger mortality improvements, which typically trigger higher (or more) payouts of longevity-linked instruments. In line with the meaning of the period effects in the CBD model, the risk adjustment should aim at shifting the distributions of both time processes downward, particularly also the second period effect since longevity at high ages generally presents the greatest risk. Thereby, a risk loading beyond best-estimate values is induced for pricing purposes. To avoid ambiguity, the span-deflators should satisfy $\mathbb{E}(\varphi_t^{Risk}) = 1$ for all $t \geq 1$ and each longevity risk driver $Risk \in \{O, S, M, \epsilon, Sub\}$.

To construct a class of meaningful span-deflators (or equivalent pricing measures, respectively), we follow Chen and Cox (2009) and Boyer and Stentoft (2013) and apply the technique of multivariate normalized exponential tilting, which offers several appealing features (cf. Wang, 2007): first, given independent risks, it is equivalent to applying the widely used (multivariate) Wang transform to each longevity risk driver individually, which in turn ensures consistency between the risk drivers, and we refer to Wang (2002, 2007) for details. At this point, we would like to stress that our distortions do *not* act directly on the distribution of the instrument one seeks to price, which would be in the original spirit of Wang (2002), but on the distributions of the underlying longevity risk drivers. As pointed out by Chen and Cox (2009), this has the advantage of giving a risk-adjusted measure that properly accounts for correlations of mortality over time and is applicable to a wide range of instruments. Second, the (multivariate) Wang transform preserves the Bernoulli, normal, and lognormal distribution. Hence, the basic model structure is preserved under \mathbb{Q} which allows to directly simulate the risk-adjusted distribution, which is of particular importance for practical applications. In detail, we apply the (multivariate) Wang transform based on the distortion parameters $\lambda_{Risk}^{(i)} > 0, i = 1, 2$ to each risk driver $Risk \in \{O, S, M, \epsilon, Sub\}$ individually and obtain the following span-deflators and implied risk-adjusted dynamics for each year $t \geq 1$:

• **Trend change occurrence**

We define $\varphi_t^O := \prod_{i=1}^2 \varphi_t^{(i)O}$ with span-deflators for time process $i = 1, 2$ of

$$\varphi_t^{(i)O} := \begin{cases} \frac{\Phi(\Phi^{-1}(p^{(i)[R]}) + \lambda_O^{(i)})}{p^{(i)[R]}} \geq 1, & \text{if } O_t^{(i)[R]} = 1 \text{ (trend change occurred)} \\ \frac{1 - \Phi(\Phi^{-1}(p^{(i)[R]}) + \lambda_O^{(i)})}{1 - p^{(i)[R]}} \leq 1, & \text{if } O_t^{(i)[R]} = 0 \text{ (no trend change occurred),} \end{cases}$$

where Φ denotes the standard normal cumulative distribution function (CDF). Under the risk-adjusted measure, the random variables $O_t^{(i)[R]}, i = 1, 2$ indicating the occurrence of a trend change between two consecutive years are again Bernoulli-distributed with risk-adjusted trend change

probabilities $\Phi(\Phi^{-1}(p^{(i)[R]}) + \lambda_O^{(i)}) \geq p^{(i)[R]}, i = 1, 2$ (see Definition 2.3 in Wang, 2007). Hence, the risk-adjusted measure assigns a higher probability to scenarios with more trend changes.

• **Trend change sign**

Regarding the signs of future trend changes, we define $\varphi_i^S := \prod_{i=1}^2 \varphi_i^{(i)S}$, where the span-deflators for time process $i = 1, 2$ are given by:

$$\varphi_i^{(i)S} := \begin{cases} \frac{\Phi(\Phi^{-1}(0.5) + \lambda_S^{(i)})}{0.5} \geq 1, & \text{if } O_t^{(i)[R]} = 1, S_t^{(i)[R]} = -1 \text{ (negative trend change)} \\ \frac{1 - \Phi(\Phi^{-1}(0.5) + \lambda_S^{(i)})}{1 - 0.5} \leq 1, & \text{if } O_t^{(i)[R]} = 1, S_t^{(i)[R]} = +1 \text{ (positive trend change)} \\ 1, & \text{else (no trend change occurred).} \end{cases}$$

Hence, the deflators take low values in scenarios of positive trend changes in favor of paths with negative trend changes. Under \mathbb{Q} , negative trend changes then have a probability of $\mathbb{Q}(S_t^{(i)[R]} = -1) = \Phi(\Phi^{-1}(0.5) + \lambda_S^{(i)}) \geq 0.5, i = 1, 2$.

• **Trend change magnitude**

For the stochastic trend change magnitudes $M_t^{(i)[R]}$, we choose $\varphi_i^M := \prod_{i=1}^2 \varphi_i^{(i)M}$ based on the following span-deflators for time process $i = 1, 2$:

$$\varphi_i^{(i)M} := \begin{cases} \exp\left(\lambda_M^{(i)} \frac{\log(M_t^{(i)[R]}) - \mu_M^{(i)[R]}}{\sigma_M^{(i)[R]}} - \frac{1}{2} \lambda_M^{(i)2}\right), & \text{if } O_t^{(i)[R]} = 1 \text{ (trend change occurred)} \\ 1, & \text{else (no trend change occurred).} \end{cases}$$

Under \mathbb{Q} , the trend change magnitudes $M_t^{(i)[R]}, i = 1, 2$ are again lognormally distributed with risk-adjusted parameters $\mu_M^{(i)[R]} + \lambda_M^{(i)} \sigma_M^{(i)[R]} \geq \mu_M^{(i)[R]}, i = 1, 2$ and same volatility as under \mathbb{P} (see Definition 2.3 in Wang, 2007).

• **Annual fluctuations around the AMT**

Considering the normally distributed fluctuations $\epsilon_t^{[R]}$ around the AMT as risk driver, we obtain span-deflators of the form:

$$\varphi_t^\epsilon := \prod_{i=1}^2 \exp\left(-\lambda_\epsilon^{(i)} \frac{\epsilon_t^{(i)[R]} - 0}{\sqrt{\Sigma_{i,i}^{[R]}}} - \frac{1}{2} \lambda_\epsilon^{(i)2}\right) \frac{1}{c_\epsilon} = \exp\left(-\Lambda_\epsilon' D_{\Sigma^{[R]}}^{-\frac{1}{2}} \epsilon_t^{[R]} - \frac{1}{2} \Lambda_\epsilon' \Lambda_\epsilon\right) \frac{1}{c_\epsilon},$$

where $D_{\Sigma^{[R]}}$ denotes a 2-by-2 matrix with entries $\Sigma_{i,i}^{[R]}, i = 1, 2$ on the diagonal and $\Lambda_\epsilon := (\lambda_\epsilon^{(1)}, \lambda_\epsilon^{(2)})'$ is the relevant market price of risk vector. The normalizing coefficient $c_\epsilon := \exp\left(-\frac{1}{2} \Lambda_\epsilon' \Lambda_\epsilon + \frac{1}{2} \Lambda_\epsilon' \Sigma_\rho^{[R]} \Lambda_\epsilon\right)$ with $\Sigma_\rho^{[R]}$ being the correlation matrix of $\epsilon_t^{[R]}$ ensures that $\mathbb{E}(\varphi_t^\epsilon) = 1$ by the properties of the multivariate lognormal distribution. Under \mathbb{Q} , annual fluctuations around the AMT are again normally distributed with risk-adjusted mean $-D_{\Sigma^{[R]}}^{-\frac{1}{2}} \Sigma_\rho^{[R]} \Lambda_\epsilon$ and same covariance matrix $\Sigma^{[R]}$ (see Theorem 5.1 in Wang, 2007).

TABLE 1
OBJECTIVE AND RISK-ADJUSTED MODEL DYNAMICS.

Risk driver	Distribution under \mathbb{P}	Distribution under \mathbb{Q}
Trend change occurrence	Bernoulli($p^{(i R)}$)	Bernoulli($\Phi(\Phi^{-1}(p^{(i R)}) + \lambda_O^{(i)})$)
Trend change sign	$\mathbb{P}(S_t^{(i R)} = -1) = 0.5$	$\mathbb{Q}(S_t^{(i R)} = -1) = \Phi(\Phi^{-1}(0.5) + \lambda_S^{(i)})$
Trend change magnitude	$\mathcal{LN}(\mu_M^{(i R)}, \sigma_M^{(i R)^2})$	$\mathcal{LN}(\mu_M^{(i R)} + \lambda_M^{(i)} \sigma_M^{(i R)}, \sigma_M^{(i R)^2})$
Fluctuations around AMT	$\mathcal{N}(0, \Sigma^{[R]})$	$\mathcal{N}(-D_{\Sigma^{[R]}}^{\frac{1}{2}} \Sigma_{\rho}^{[R]} \Lambda_{\epsilon}, \Sigma^{[R]})$
Mortality differentials	$\mathcal{N}(0, \Sigma^{[Sub]})$	$\mathcal{N}(-D_{\Sigma^{[Sub]}}^{\frac{1}{2}} \Sigma_{\rho}^{[Sub]} \Lambda_{Sub}, \Sigma^{[Sub]})$

• **Subpopulation-specific mortality differentials**

Analogously, we obtain for the Gaussian subpopulation-specific mortality differentials:

$$\varphi_t^{Sub} := \exp\left(-\Lambda'_{Sub} D_{\Sigma^{[Sub]}}^{-\frac{1}{2}} \epsilon_t^{[Sub]} - \frac{1}{2} \Lambda'_{Sub} \Lambda_{Sub}\right) \frac{1}{c_{Sub}},$$

where $D_{\Sigma^{[Sub]}}$ denotes a diagonal matrix with entries $\Sigma_{i,i}^{[Sub]}$, $i = 1, \dots, 2N_{Sub}$. For the market price of risk vector $\Lambda_{Sub} := \frac{1}{N_{Sub}}(\lambda_{Sub}^{(1)}, \lambda_{Sub}^{(2)}, \dots, \lambda_{Sub}^{(1)}, \lambda_{Sub}^{(2)})' \in \mathbb{R}^{2N_{Sub}}$, we assume a common market price of risk for subpopulation-specific mortality differentials $\lambda_{Sub}^{(i)}$, $i = 1, 2$ which is evenly distributed among the subpopulations. Of course, this assumption might be relaxed when fitting the model to actual market prices, which is at present not possible due to the lack of publicly available pricing information (see Section 3.3 for a discussion). Again, a normalization coefficient of the form $c_{Sub} := \exp(-\frac{1}{2} \Lambda'_{Sub} \Lambda_{Sub} + \frac{1}{2} \Lambda'_{Sub} \Sigma_{\rho}^{[Sub]} \Lambda_{Sub})$ is required with $\Sigma_{\rho}^{[Sub]}$ representing the correlation matrix. Under \mathbb{Q} , the Gaussian model structure is preserved and subpopulation-specific mortality differentials are again driven by a multivariate normal distribution with risk-adjusted mean $-D_{\Sigma^{[Sub]}}^{\frac{1}{2}} \Sigma_{\rho}^{[Sub]} \Lambda_{Sub}$ and same covariance matrix as under \mathbb{P} (see Theorem 5.1 in Wang, 2007).

Besides, we omit a further adjustment for diversifiable small sample risk and we follow most of the existing literature and refrain from pricing parameter uncertainty to avoid over-complexity. After sampling the model parameters at the start of each simulation path from suitable distributions as outlined in Appendix A.3, the change of measure for pricing purposes is conducted conditional on the drawn parameter set. As summarized in Table 1, the basic model structure is retained under this change of measure.

However, the question of how to calibrate the market prices $\lambda_{Risk}^{(i)} > 0$, $i = 1, 2$ assigned to the risk drivers $Risk \in \{O, S, M, \epsilon, Sub\}$, that is, which span-deflators to choose from the derived set of candidates, remains open. Also,

a discussion on whether all risk drivers should be included in the deflator is warranted. We address these issues in Section 3.3.

3.2. Cost-of-capital approach

Under modern risk-based solvency regimes, insurance companies are required to back their business with adequate economic capital. From the perspective of a longevity risk taker, the expected return needs to compensate at least for the additional expected costs of providing economic capital for assuming the risk. Hence, the expected cost of capital for taking on the risk can be interpreted as a minimum risk loading on top of the objective best-estimate value.

Under Solvency II, the SCR is defined as the 99.5% VaR of the basic own funds over a 1-year horizon, where the basic own funds correspond to the difference between the market value of assets and the market value of liabilities. Loosely speaking, the SCR corresponds to the capital required to cover all losses which may occur over the following year at a confidence level of at least 99.5%. Alternatively, the Solvency II directive allows companies to apply a standard formula for longevity risk based on a simplified one-off shock approach. However, we focus on a risk-based (partial) internal model, since the standard formula's adequacy for pricing purposes is questionable due to structural shortcomings, see Börger (2010). In what follows, we exclusively focus on the longevity SCR and neglect any further potential diversification effects with an existing business mix (see Section 3.3 for a discussion).

From the risk taker's perspective, longevity risk over a 1-year horizon consists of two components (cf. Börger, 2010):

- First, the risk that next year's realized mortality will be lower than anticipated, for example due to a mild winter, triggering higher (or more) payments than anticipated.
- Second, the risk that the company will have to depreciate its position at the end of the year due to revised long-term mortality assumptions for the time beyond.

Clearly, both components might have an impact on the value of a longevity-linked security. Following Börger (2010) and Börger *et al.* (2019b), we assume that the remaining assets do not contribute to the longevity SCR and that there is no loss-absorbing capacity of technical provisions. Under these assumptions, the SCR in year T for issuing a longevity-linked instrument H is defined as the 99.5% quantile of

$$P(T, T+1) (\tilde{H}(T+1) + CF_H(T+1)) - \tilde{H}(T),$$

where

- $\tilde{H}(T)$ denotes the time T best-estimate value of all future cash flows,
- $CF_H(T+1)$ denotes the random cash flows that might occur over the year,

- $\tilde{H}(T+1)$ denotes the best-estimate value of all future cash flows at time $T+1$ according to potentially revised mortality assumptions, and
- $P(T, T+1)$ represents the time T price of a zero coupon bond with maturity $T+1$.

The instrument's valuation at time T (or $T+1$, respectively) is conducted using the best-estimate central path of mortality derived with the EMT model component at that point in time (see Section 2.3) and the prevailing risk-free interest rate term structure.

The time-zero random present value of the cost of capital is then given by:

$$CoC := r_{CoC} \sum_{t \geq 0} P(0, t+1) SCR(t),$$

where $r_{CoC} \geq 0$ denotes the minimum rate of return in excess of the risk-free rate that investors expect for providing capital. Obviously, the SCR at a future point in time depends on realized mortality up to that point in time and hence is a random variable. Consequently, the accumulated cost of capital is also a random variable and its exact computation requires the determination of each year's SCR conditional on realized mortality. We deal with this additional layer of complexity within a two-level nested Monte Carlo simulation: given an outer simulation path for realized mortality, the SCR at a certain point in time is computed by simulating mortality over a 1-year horizon, reevaluating the position using the EMT model component, and empirically deriving the 99.5% quantile. Note that the use of a pragmatic central path of mortality in the EMT model component avoids an additional layer of nested simulations in practical applications and constitutes a valid approach for symmetric payout structures since it provides a reasonable approximation for the mean. Nevertheless, the valuation of option-type contracts over a multi-year horizon would call for a third level of nested simulations (or for a more tractable stochastic mortality model). Overall, the entire distribution of the time-zero random present value of future capital charges can be derived.

The risk taker demands at least the instrument's objective best-estimate value plus the expected cost of capital giving a lower bound for the risk premium of $\mathbb{E}^{\mathbb{P}}(CoC)$. Assuming a cost-of-capital rate of $r_{CoC} = 6\%$, it corresponds to the *risk margin* under Solvency II, which has to be reserved in addition to the best-estimate liabilities. In practice, the risk margin is usually approximated by assuming a future mortality evolution according to its current best estimate, see for example Zeddouk and Devolder (2019). In contrast, we empirically derive an entire distribution for future capital charges from which the risk margin can be derived.

3.3. Discussion and qualitative comparison

In this section, we briefly compare and comment on the presented pricing approaches with regard to their economic justification, practical applicability, and calibration.

3.3.1. *Economic justification*

As argued by Cairns *et al.* (2006), pricing longevity-linked securities under a risk-adjusted measure would only be fully justified in a complete market in which all longevity-linked payments can be replicated. Hence, given the current stage of the market, the risk-adjusted measure is not unique and different market players might rely on different choices. However, in light of the potential size of the longevity risk transfer market, further market development toward higher liquidity can be expected (cf. Blake *et al.*, 2019). Once a deep and liquid market develops, observable market prices might ultimately imply a unique risk-adjusted measure. Even if risk-adjusted valuation under a pricing measure is not yet fully economically justified for longevity risk pricing, it provides at least a theoretically sound approach for deriving prices which might eventually prevail in a fully developed market.

The cost-of-capital approach on the other side is inspired by risk-based capital requirements for (re)insurers. It is in line with the interpretation of the risk margin under Solvency II as a loading for non-hedgeable risk that other (re)insurance companies would be expected to require to take over the obligations. Of course, this argument only applies to (re)insurers that are regulated by a solvency regime. Nevertheless, since the global longevity risk transfer market is at present dominated by reinsurance companies, this approach seems to be of high practical relevance and economically justified.

3.3.2. *Practical applicability*

From a practical point of view, risk-adjusted valuation under a pricing measure offers several desirable features. As previously outlined, the change of measure preserves the basic model structure which allows to directly simulate the risk-adjusted distribution and to efficiently price various types of longevity-linked securities, including option-type contracts, in a Monte Carlo simulation. In particular, a set of forward rates can be derived for a class of small building blocks, such as q-forwards or s-forwards. Instruments which can be written as a linear combination of these smaller building blocks can then be priced based on these forward rates due to the linearity of the pricing operator. However, the calibration of multiple market prices of risk for practical applications is not straightforward as we address in the next section.

By contrast, the cost-of-capital approach only requires specification of a single cost-of-capital rate for which a legally prescribed reference point of 6% is available under Solvency II. On the other side, it involves computationally complex SCR derivations over the entire term of the hedge contract. Furthermore, it merely offers a valid starting point for a reasonable risk premium. In fact, prices might also depend on potential diversification benefits with the risk taker's existing business mix, which we do not explicitly deal with. In practice, SCRs are typically derived for each submodule separately and then aggregated by explicitly taking into account potential diversification benefits. The cost of capital for issuing a longevity hedge net of diversification effects

could be quantified by means of a reallocation principle (such as the Euler Principle). Consequently, different companies might have different appetites for longevity risk, and the pricing operator is therefore generally not linear.

3.3.3. *Calibration*

First, we address the question of whether all risk drivers for which we proposed a suitable distortion in Section 3.1 should be included in the deflator. In principle, investors should only be compensated for taking undiversifiable risks. Hence, unsystematic fluctuations due to a small portfolio size are not considered in the deflator. In contrast, the systematic risks of random occurrence, sign, and magnitude of future trend changes should clearly be included. Also, annual fluctuations around the AMT which account for systematic transitory effects on mortality, for example, less people than usual are dying from flu, cannot readily be diversified by increasing the contract volume. In particular in the short run, such transitory effects may trigger higher hedge payments than anticipated which justifies consideration in the deflator. However, risk takers might to some extent mitigate their exposure to these transitory effects by issuing contracts with different maturities. Hence, some investors might arguably be willing to accept a lower risk premium for transitory fluctuations around the AMT. Similarly, the exposure to systematic mortality differentials cannot be diversified with increasing volume but, to some extent, by pooling portfolios with different socioeconomic characteristics. However, the volume of annuity business is typically biased toward the healthier and more affluent part of the overall population. Hence, it seems at least questionable whether longevity risk takers are able to maintain a well-balanced portfolio with regard to socioeconomic characteristics. Since these diversification opportunities depend on the risk taker's existing portfolio or his ability to write business in different market segments, respectively, consideration in the deflator appears justified. To demonstrate the full capabilities of our model in the numerical part, we include both annual fluctuations around the AMT and subpopulation-specific mortality differentials in the deflator. Nevertheless, we would like to stress that the market prices for individual risk drivers can easily be switched off if desired.

Regarding the calibration of the market price of longevity risk, there is an ongoing debate on how to find reasonable values in an incomplete market. If a rich set of price data was publicly available, the optimal distortion parameters could be numerically calibrated to fit market prices in terms of a predefined target function, see for example Chen and Cox (2009). However, this approach would require estimation of up to 10 distortion parameters in our setup. Unfortunately, sufficient pricing information on long-term longevity-linked securities is currently not publicly available. As argued by Wang (2007) and Chen and Cox (2009), multivariate normalized exponential tilting ensures a consistent interpretation of the tilting parameters among different risks via percentile mapping to standard normal variables. This justifies the widely used

assumption of equal market prices for all risk drivers and allows to interpret the common distortion parameter, denoted by $\lambda > 0$, as market price of longevity risk, see for example Cairns *et al.* (2006), Chen and Cox (2009), or Leung *et al.* (2018). Under this restriction, the calibration narrows down to the specification of a single parameter. Nevertheless, even though several ideas have been proposed on how to find a reasonable market price of longevity risk in the absence of reliable market data, no consensus has been reached. For instance, Lin and Cox (2005) rely on market annuity quotes, Loeys *et al.* (2007) attempt to deduce a reasonable market price from other asset classes like stocks, or Cairns *et al.* (2006) and Leung *et al.* (2018) rely on the price of the announced but never issued EIB/BNP bond.

Naturally, there is also a large uncertainty regarding the ‘correct’ cost-of-capital rate. In practice, the rate of return at which shareholders are willing to provide equity might differ among different (re)insurers depending on their solvency and credit rating. For the risk margin computation under Solvency II, the cost-of-capital rate is currently set to 6% which offers a valid starting point for finding a reasonable market price of longevity risk: it can be calibrated to the capital charges for a given longevity-linked instrument so that the risk-adjusted measure reproduces this risk loading (cf. Börger, 2010; Zeddouk and Devolder, 2019). Given the interpretation of the risk margin under Solvency II, this is also in the spirit of Wüthrich and Merz (2013) who argue that risk-adjusted valuation based on probability distortions should provide values at which ‘*liabilities could be transferred between two knowledgeable and willing parties in an arm’s length transaction*’. In light of the current market environment, we believe that the presented cost-of-capital approach provides an appropriate reference framework for finding a reasonable market price of longevity risk. Obviously, the results depend on the considered instrument and vary between companies. While (re)insurers will typically target a return on equity above the regulatory cost-of-capital rate on the one hand, they may, on the other hand, also accept lower risk premiums due to diversification benefits. Overall, we believe that a calibration based on a cost-of-capital rate of 6% is appropriate. We implement this approach in the next section.

4. NUMERICAL ILLUSTRATIONS

For our numerical illustrations, we calibrate the mortality model to the historical mortality experience of the male population of England and Wales. We consider the national population as the reference population and we rely on $N_{Sub} = 5$ socioeconomic subpopulations (ordered from the most to the least deprived areas) based on the Index of Multiple Deprivation (IMD) for England. We extrapolate mortality up to a limiting age of 130 years. For more details on the model calibration, including the underlying data set, we refer to the appendix. Furthermore, we assume a risk-free interest rate of $r = 2\%$ and a cost-of-capital rate of $r_{CoC} = 6\%$. We perform a two-level nested Monte Carlo

simulation with 10,000 outer sample paths and additional 10,000 inner 1-year scenarios for the SCR computations.

In this section, we apply the proposed pricing methods in a simplified setting. By way of illustration, we consider a hedger wishing to hedge the longevity risk arising from a portfolio of immediate (or deferred) life annuities. Starting at the retirement age of $x_R = 65$ years, the beneficiaries receive a lifelong payment of one unit of currency paid at the beginning of each year. The exemplary portfolio consists of $N_{Book} = 10,000$ policyholders belonging to the same cohort, where we consider different starting ages $x_0 \in \{50, 65, 80\}$ at time zero.

Following Haberman *et al.* (2014), we capture portfolio-specific mortality characteristics with a so-called ‘characterization approach’. Assuming that the majority of portfolio-specific mortality characteristics are captured by the considered socioeconomic subpopulations, we interpret the hedger’s book population as a subset of the reference population with a specific socioeconomic structure. For illustrative purposes, we use an exemplary initial socioeconomic book composition of $\eta = (0\%, 0\%, 30\%, 30\%, 40\%)$. The number of initial policyholders coming from subpopulation $p \in \{1, \dots, N_{Sub}\}$ is then given by $B_{x_0,0}^{[p]} := \eta_p N_{Book} \in \mathbb{N}$. As pointed out by Haberman *et al.* (2014), this approach offers several advantages. Most importantly, the model is calibrated to a credible and sufficiently rich data set rather than to smaller and therefore often less reliable portfolio-specific data. Hence, it can be applied to portfolios of any size without mistaking unsystematic variations in the data due to a small sample size for systematic mortality differentials. Also, the evolution of the socioeconomic composition over time is adequately captured. Of course, portfolio characteristics might differ from our exemplary values, for example, when considering pension funds for specific industries. In this case, they can easily be altered (for instance by assigning more weight to the less affluent subpopulations) without having to recalibrate the model, which offers great flexibility. As demonstrated by Villegas *et al.* (2017), a suitable weighting across five socioeconomic subpopulations usually allows to reliably capture the key mortality characteristics of typical pension schemes.

In Section 4.1, we introduce longevity swaps as suitable hedging instruments for this setting. As outlined in the previous section, we then derive a range of reasonable market prices of longevity risk in Section 4.2 based on the cost-of-capital approach and discuss the corresponding risk-adjusted forward rates in Section 4.3. Finally, we compare our results with alternative modeling and pricing approaches in Section 4.4.

4.1. Longevity swaps

At present, the longevity risk transfer market is dominated by longevity swaps (cf. Blake *et al.*, 2019). These contracts can either be customized, that is, tailored to the characteristics of the hedger’s book population, or index-based and linked to the reference population.

4.1.1. Customized longevity swaps

In a customized longevity swap, the risk taker pays a floating series of cash flows corresponding to the realized annuity payments in exchange for a set of pre-agreed payments resulting in annual net payments of

$$CF_H(t) := 1_{\{x_0+t \geq x_R\}} (S_{x_0+t,t} - F_{x_0+t,t}), \quad 0 \leq t \leq \tau,$$

where the floating lag $S_{x_0+t,t} := \sum_{p=1}^{N_{Sub}} B_{x_0+t,t}^{[p]}$ corresponds to the sum of all annuity payments in year t for which the annuity provider is liable. The fixed forward rates $F_{x_0+t,t}$ need to be specified at inception through the pricing approach. By construction, an unlimited ($\tau = \infty$) customized longevity swap provides perfect protection against longevity risk.

4.1.2. Index-based longevity swaps

Alternatively, the instrument can be linked to the mortality indices of the reference population resulting in annual payments of

$$CF_H(t) := 1_{\{x_0+t \geq x_R\}} (S_{x_0+t,t}^{[R]} - F_{x_0+t,t}^{[R]}), \quad 0 \leq t \leq \tau,$$

where the survivor index $S_{x_0+t,t}^{[R]} := \sum_{p=1}^{N_{Sub}} B_{x_0,0}^{[p]} \prod_{s=0}^{t-1} (1 - q_{x_0+s,s+1}^{[R]})$, $t \geq 0$ is based on the ex-post survival probability for an individual from the reference population aged x_0 at time zero and $F_{x_0+t,t}^{[R]}$ denotes the index-based forward rates. From the hedger's perspective, this design gives rise to population basis risk, since the evolution of the book population might deviate from that captured by the reference population mainly for two reasons (cf. Haberman *et al.*, 2014 or Villegas *et al.*, 2017): first, due to systematic mortality differentials, which is modeled via the subpopulations, or second, due to unsystematic fluctuations linked to smaller diversification effects, which is modeled by sampling survivors from a Binomial distribution. On the other side, an index-based design is expected to be more appealing to investors since it does not require any knowledge on the characteristics of the insured population.

4.2. Capital charges and the market price of longevity risk

We start by analyzing capital charges for issuing unlimited customized longevity swaps. Figure 2 shows histograms for the cost of capital and the corresponding risk margins (as the expected capital charges $\mathbb{E}^{\mathbb{P}}(CoC)$) for the three considered starting ages.

As anticipated, the expected capital charges are higher for lower starting ages, since the company has to provide regulatory capital over a longer contract duration. Moreover, the company's exposure to longevity risk is more pronounced for lower starting ages since longevity risk has more time to accumulate, which requires higher SCRs. Interestingly, the histograms reveal a considerable uncertainty regarding future capital charges. In fact, their standard deviations, which are listed in Table 2, result in considerable coefficients

TABLE 2

MEAN AND STANDARD DEVIATION OF THE COST OF CAPITAL FOR ISSUING UNLIMITED CUSTOMIZED LONGEVITY SWAPS FOR DIFFERENT STARTING AGES AND CORRESPONDING MARKET PRICES OF LONGEVITY RISK.

	$x_0 = 50$	$x_0 = 65$	$x_0 = 80$
$\mathbb{E}^{\mathbb{P}}(CoC)$	8330	4137	1313
$SD^{\mathbb{P}}(CoC)$	1920	750	197
Corresponding λ	0.316	0.297	0.225

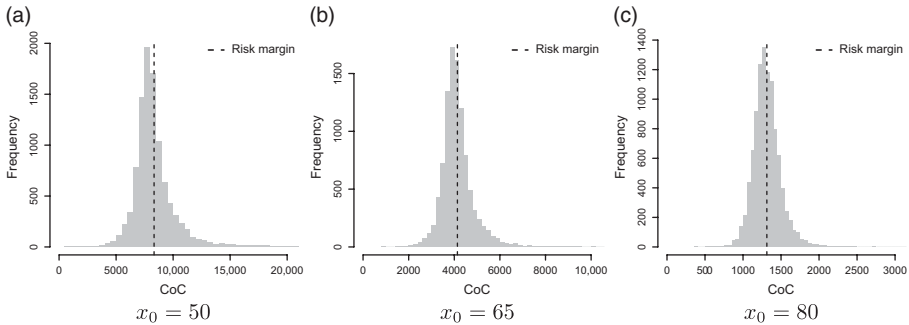


FIGURE 2: Histograms for the random present value of the cost of capital and corresponding risk margins for issuing unlimited customized longevity swaps for different starting ages.

of variation between 15% and 23%. We conclude that regulatory capital charges differ significantly between the simulation paths. The reason for this is twofold. First, future SCR_ts naturally depend on realized mortality, in particular on the occurrence of unanticipated trend changes. While negative trend changes lead to higher SCR_ts in the following years, positive trend changes imply a capital relief. Obviously, scenarios of strong mortality improvements require higher SCR_ts against longevity risk over a longer time horizon compared to scenarios of higher mortality. Second, even before a trend change occurs, the overall level of SCR_ts generally differs between the simulation paths since we account for parameter uncertainty in the trend change parameters as outlined in Appendix A.3.

In the second step, we numerically derive the corresponding market prices of longevity risk λ so that the risk-adjusted pricing approach produces the same risk loading as the cost-of-capital approach. Table 2 shows the risk margins and the corresponding market prices of longevity risk for the considered starting ages. As expected, the calibrated market prices of risk are positive since the demand for de-risking is coming from an annuity provider. As pointed out by Cairns *et al.* (2006), λ might be negative if the market was dominated by hedgers of mortality risk. Furthermore, we observe that for lower starting ages, higher values of λ are required to reproduce the risk margin. This is due to the fact that long-term longevity exposures are rather capital-intensive under risk-based internal models.

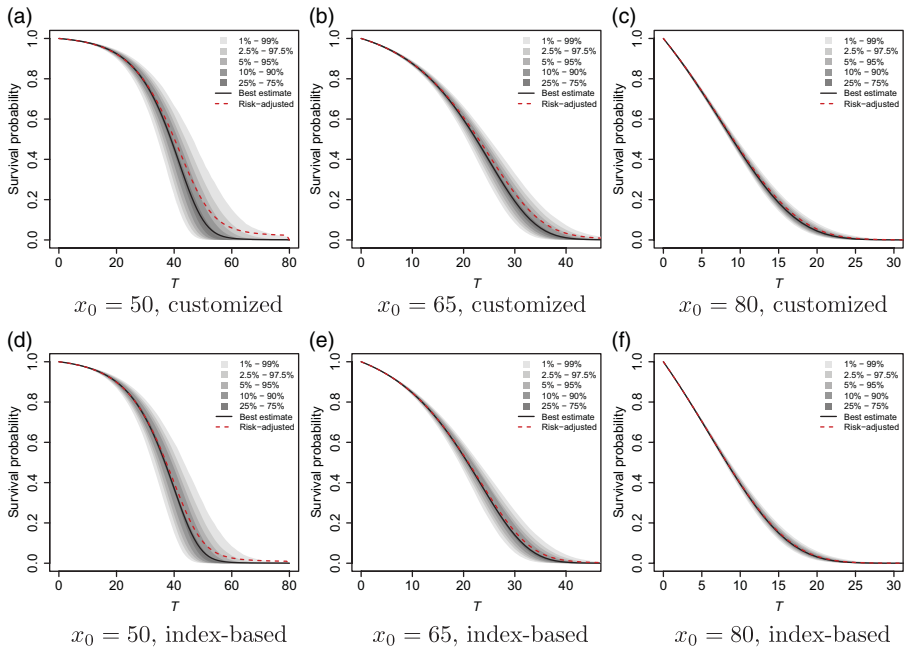


FIGURE 3: Quantile plots, best-estimate, and risk-adjusted survival probabilities for customized (upper row) and index-based (lower row) longevity swaps for different starting ages.

For comparison, values between 0.15 and 0.4 are commonly reported for the market price of longevity risk. For instance, Leung *et al.* (2018) report values between 0.15 and 0.35 for the risk-adjusted pricing approach based on the CBD model and Levantesi and Menzietti (2017) arrive at slightly higher values between 0.35 and 0.41 when matching the risk-adjusted CBD model with the cost-of-capital approach. However, it should be kept in mind that reported market prices of longevity risk are typically model-specific and not directly comparable due to differences in the intrinsic model setups. Nevertheless, our calibrated market prices of longevity risk seem to be of reasonable magnitude.

4.3. Risk-adjusted forward rates

Relying on the specified set of market prices of longevity risk, we apply the risk-adjusted pricing approach to derive forward rates for longevity swaps. Figure 3 shows quantile plots for the T -year survival probabilities along with their best-estimate values (black solid lines) and their risk-adjusted counterparts (red dashed lines) for the considered starting ages, where the upper row shows the customized and the lower row the index-based survival curves.

We first look at the customized designs. When comparing the prediction bands for different starting ages, we observe that the uncertainty in future survival probabilities is more pronounced for lower starting ages. This is due to

the fact that there is more time for potential future trend changes to occur and to unfold. Furthermore, we observe that the prediction bands are rather narrow in the short term and widen over time reflecting the nature of longevity risk as a typically slowly accumulating risk. With regard to this increasing degree of uncertainty over time, the forward rates show a reasonable pattern: they typically lie above their best-estimate counterparts and the distances between both curves, which indicate the implied risk loading, increase over time in line with the growing uncertainty regarding future survival rates. Obviously, most of the risk premium is allocated to longer maturities. We further discuss the structure of the implied risk loading in Section 4.4.2.

When comparing the index-based (lower row) with the customized survival curves (upper row), we observe that the index-based prediction bands are slightly narrower. This is because they are not affected by socioeconomic mortality differentials and small sample risk. Since socioeconomic mortality differentials are included in the deflator, the implied risk loading is notably higher for the customized designs than for the index-based variants. However, besides being narrower, the prediction bands are similar in shape to those for the customized designs. Hence, it can be concluded that most of the uncertainty in future mortality originates from the long-term mortality trend risk of the reference population. Overall, the derived risk-adjusted forward rates appear plausible in terms of intrinsic risk loading.

4.4. Comparison with alternative modeling and pricing approaches

In this section, we compare our (risk-adjusted) trend process and the resulting swap rates with alternative modeling and pricing approaches. Since the index-based designs do not provide any further insights, we focus on the customized designs in this section.

4.4.1. Comparison with time series with constant trend drift

We now consider the following alternative modeling approaches for the reference population that do not allow for trend changes:

- Switching off the risk of future trend changes by setting $p^{(i)[R]} := 0$ for $i = 1, 2$ to illustrate the implications of incorporating trend changes into the mortality model.
- Using the widely adopted RWD. Following Cairns *et al.* (2006), we calibrate the drift and covariance matrix to the most recent 20 years of data and draw them at the start of each model path from a suitable normal-inverse-Wishart distribution to account for parameter uncertainty (we refer to their paper for technical details).

We start with a comparison of the resulting 99% prediction intervals for the period effects, which are shown in Figure 4. The effect of stochastic trend changes becomes evident when comparing the prediction intervals for

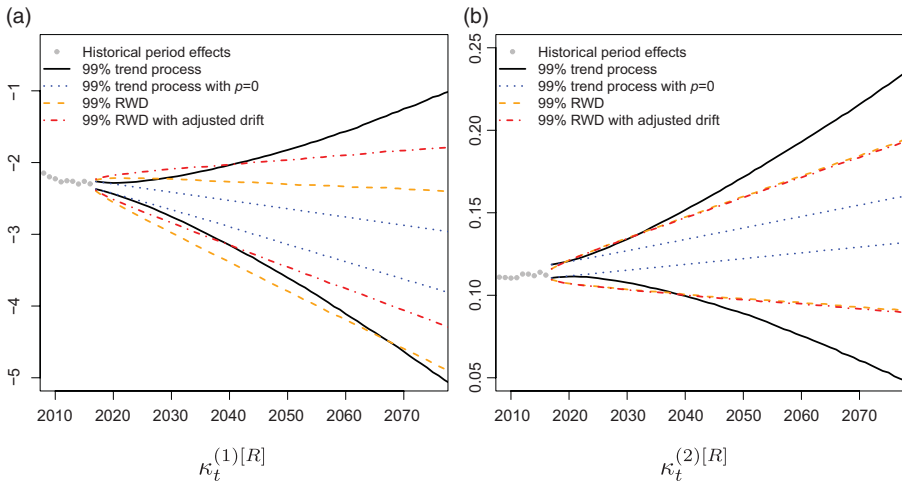


FIGURE 4: Comparison with alternative modeling approaches: 99% prediction intervals for the CBD period effects for the trend process (black), its counterpart with constant trend (blue), and the RWD (orange with adjusted drift).

the stochastic trend process (solid black lines) with those for its counterpart with constant trend (blue dotted lines). While they practically coincide in the first years, the former widen considerably over time reflecting the growing uncertainty arising from potential future trend changes. The blue prediction intervals, on the other side, remain unrealistically narrow in the long run. This is due to the fact that the only sources of uncertainty are the stochastic starting trends and the random annual fluctuations around it. Clearly, a trend-stationary process with constant trend substantially underestimates longevity risk in the long run.

Comparing the trend process with the RWD (orange dashed lines), we first observe that the prediction intervals differ considerably in terms of position for $\kappa_t^{(1)[R]}$. This is because the drift is estimated as the average change in the period effects over the whole calibration period of 20 years, whereas the starting values for the trend process depend on the most recent mortality patterns since the last detected trend change. Hence, the latter has the advantage of picking up potential changes in the long-term mortality trend much faster than the RWD.

For comparability, the red dot-dashed lines show the prediction intervals for the RWD when setting the drift estimate equal to the expected starting trends of the trend process. Compared to the trend process, the prediction intervals are substantially wider in the short run. However, they widen rather slowly over time and remain rather narrow in the long run due to the fixed drift, which has been criticized by several authors including Börger and Schupp (2018) and Liu and Li (2016). In contrast, the prediction intervals of the trend process widen at faster rates in line with the growing uncertainty in future

TABLE 3

MEAN AND STANDARD DEVIATION OF THE COST OF CAPITAL FOR ISSUING UNLIMITED CUSTOMIZED LONGEVITY SWAPS FOR DIFFERENT STARTING AGES AND CORRESPONDING MARKET PRICES OF LONGEVITY RISK; COMPARISON BETWEEN TREND PROCESS AND ALTERNATIVE APPROACHES WITHOUT TREND CHANGES.

Model	Trend			Trend $p^{(t R)} = 0$			RWD		
	50	65	80	50	65	80	50	65	80
$\mathbb{E}^{\mathbb{P}}(CoC)$	8330	4137	1313	1764	1030	458	2130	1101	442
$SD^{\mathbb{P}}(CoC)$	1920	750	197	284	174	74	1008	566	229
Corresponding λ	0.316	0.297	0.225	0.205	0.145	0.100	0.057	0.039	0.032

mortality trends and consequently overtake those generated by the RWD over a longer projection horizon beyond 30 years. Overall, we conclude that the trend process provides a plausible and significantly different provision of uncertainty compared to approaches with constant mortality trends/drifts, especially in the long run. In what follows, we rely on the adjusted drifts to ensure that our analyses do not get blurred simply because of different initial trend assumptions.

Next, we carry out the risk margin computation from Section 4.2 under both alternative modeling approaches. For consistency, best-estimate mortality for SCR computations is also derived by applying a weighted linear regression to the most recent data points, where the optimal weighting is derived by means of the same optimization criterion (see Appendix A.6). Table 3 shows the resulting risk margins. As anticipated, they are substantially smaller for all considered starting ages when the trend change probabilities are set to zero. For instance for a starting age of $x_0 = 65$ years, the expected cost of capital is more than four times higher when random trend changes are allowed for. This is due to the fact that the variability in best-estimate mortality assumptions over a 1-year horizon, which essentially determines the SCRs, is significantly higher when the mortality trend is not only sampled at the start of the simulation but can also randomly change at future points in time. The same argument also applies to the RWD, where the drifts are sampled at the start of each model path but remain fixed over the projection horizon. Consequently, the risk margins are less than half the size of those based on the trend process for all considered starting ages.

Finally, we calibrate the corresponding market prices of longevity risk based on the derived risk margins. For the RWD, the change to a risk-adjusted measure is performed by applying suitable distortions to the normally distributed annual innovations as originally proposed by Cairns *et al.* (2006). The resulting market prices of longevity risk are included in Table 3. In line with our previous findings, they lie significantly below the values which we obtain with our trend process for all considered starting ages. We infer that the allowance for unexpected mortality trend changes yields more conservative estimates for the market price of longevity risk and higher prices for longevity

hedges when relying on a cost-of-capital reference framework. We therefore conclude that our model constitutes a valuable alternative to the widely used (risk-adjusted) RWD.

4.4.2. Comparison of risk-adjusted forward rates with other pricing approaches

In Section 3.1, we have constructed our risk-adjusted trend process by applying the Wang distortion to each longevity risk driver individually. Now, we compare the resulting forward rates with those based on the following alternative pricing approaches:

- applying the Wang transform to best-estimate survival rates $\mathbb{E}^{\mathbb{P}}(S_{x_0+t,t})$, $t > 0$ in the spirit of Lin and Cox (2005);
- applying the Wang transform to the empirical cumulative distribution functions (CDFs) of $S_{x_0+t,t}$, $t > 0$;
- using the risk-adjusted RWD of Cairns *et al.* (2006) from the previous subsection.

For each approach, we assume a constant time-invariant distortion parameter which needs to be specified first. To ensure comparability, we assume for a moment that the cumulative risk loading for the unlimited customized longevity swap with starting age of 65 years from Section 4.2 is given (e.g., observable in the market) and use it as a common reference point for the calibration of these pricing approaches. While all approaches then produce, by construction, the same cumulative risk premium for the given unlimited longevity swap, we can compare the forward rates with regard to the allocation of the risk premium over the term of the hedge. To this end, we follow Cairns *et al.* (2006) and convert the risk loading for the time t hedge cash flow $CF_H(t)$ into an average rate of return per annum for the investor of

$$\delta(t) := \frac{1}{t} \log \left(\frac{F_{x_0+t,t}}{\mathbb{E}^{\mathbb{P}}(S_{x_0+t,t})} \right)$$

with $F_{x_0+t,t}$ representing the respective forward rates. The results are shown in Figure 5.

Regarding the two alternative approaches based on the Wang transform, the most striking observation is that much of the risk premium is allocated to shorter maturities when applying the Wang transform to best-estimate survival rates (green dashed lines). As pointed out by Bauer *et al.* (2010), this proceeding does not properly capture the uncertainty in future survival rates and yields risk premiums which are therefore not proportional to the aggregate risk. In contrast, applying the Wang distortion to the empirical CDFs of future survival rates (orange dash-dotted lines) seems to provide more adequate forward rates, given that a greater proportion of the risk premium is allocated to longer maturities in line with the growing uncertainty in future survival rates. Yet, this approach also allocates a greater proportion of the risk

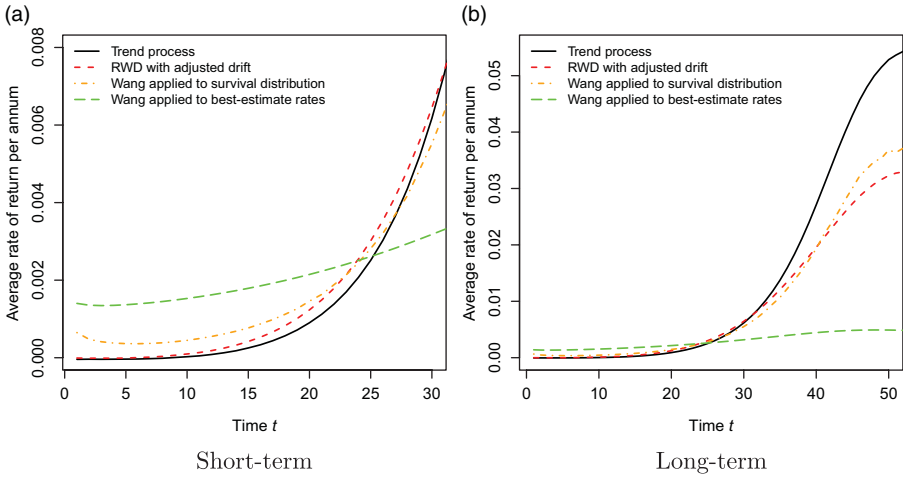


FIGURE 5: Average rate of return per annum $\delta(t)$ for the time t cash flow of the customized longevity swap with starting age $x_0 = 65$ years based on different pricing approaches.

premium to shorter maturities compared to our risk-adjusted trend process (black solid lines) and also yields a significantly different risk loading structure for several reasons. First, this approach does not differentiate between randomness that originates from undiversifiable systematic changes in mortality and that arising from diversifiable fluctuations in portfolios of limited size for which the investor should not obtain a risk premium. Second, applying the Wang transform to the final distributions of the hedge cash flows does not properly account for correlations of mortality over time. Moreover, it has the major drawback of depending on the specific payout structure of the instrument under consideration and might therefore require different distortions as soon as multiple instruments are priced. By applying the Wang distortion directly to the underlying longevity risk drivers, the risk-adjusted trend process overcomes these issues and produces risk loadings that appear plausible.

Finally, when comparing the forward rates based on our risk-adjusted trend process with those based on the risk-adjusted RWD (red dashed lines), we observe that the former assigns a lower risk premium to short maturities. For instance for $t = 15$, the average annual risk premium corresponds to 2.6 basis points for the former compared to 4.3 basis points for the latter. While the risk premium naturally increases in both models over time in line with the growing uncertainty regarding future survival rates, it picks up at a much faster rate for the trend process, ending up at a higher level for long-term cash flows. We conclude that the risk-adjusted trend process allocates a greater proportion of the risk loading to longer maturities than the commonly used RWD. This is in line with the prediction intervals in Section 4.4.1 and appears plausible given the much higher exposure of long-term longevity-linked cash flows to the risk of unexpected changes in the long-term mortality trend.

5. CONCLUSION

Pricing longevity-linked securities in the absence of a liquid and complete market has recently been addressed by many authors. Most of the proposed pricing approaches are based on stochastic mortality models in which one or more time-dependent parameters are projected into the future. As argued by several authors in the field of longevity trend modeling, long-term mortality projections should account for the risk of future mortality trend changes. Otherwise, the risks involved in a longevity transaction will be systematically understated which might lead to insufficient risk premiums and to inadequate prices.

We contribute to the literature on longevity risk pricing by addressing the determination of adequate longevity risk premiums within a stochastic modeling framework which explicitly models the risk of random future changes in the long-term mortality trend. We construct a set of meaningful stochastic discount factors based on suitable probability distortions which imply equivalent risk-adjusted pricing measures under which the basic model structure is preserved. These deflators help to understand how the risk-adjusted measure relates to its objective counterpart, under which the model is calibrated. To find a reasonable range for the market price of longevity risk in the current incomplete market environment, we follow Börger (2010) and rely on a cost-of-capital approach in the context of a regulatory capital model under Solvency II. Unlike previous studies, we explicitly derive entire distributions for future capital charges for longevity risk from which adequate risk premiums can be deduced. Moreover, we also allow mortality in selected subpopulations to differ from the underlying trend of the overall population. Hence, our framework is applicable to customized as well as to index-based instruments. We discuss the presented approaches with regard to their economic justification, practical applicability, and calibration.

In a numerical illustration, we apply the proposed methods to price longevity swaps. For different contract designs, we first derive the required risk margin under the cost-of-capital approach and then construct the risk-adjusted measure, so that both approaches imply the same risk premium. We show that this results in a reasonable range of market prices for longevity risk. Furthermore, we demonstrate that the risk-adjusted pricing approach produces highly plausible forward rates in terms of intrinsic risk loading. We find that in the presence of mortality trend changes, most of the risk premium is allocated to longer maturities. This seems reasonable given the longer time horizon for longevity trend changes to unfold. In light of the potentially devastating impact of an unanticipated mortality trend change on the solvency of longevity risk bearers, our modeling framework constitutes a valuable alternative to existing approaches for longevity risk pricing.

Interestingly, we find that future capital charges for longevity risk are subject to a considerable degree of uncertainty. Since commonly used approximations for the risk margin only provide a point estimate for future capital charges rather than an entire distribution, they might systematically underestimate

longevity risk. In particular in scenarios of strong mortality improvements, annuity providers might not only be liable for more payments to policyholders, but they also need to account for higher capital charges. We plan to address this issue and the potential implications for longevity risk management in future works.

ACKNOWLEDGMENTS

I would like to thank the Editor and two anonymous referees for many thoughtful comments and suggestions that helped to significantly improve earlier versions of this manuscript. I am also very grateful to Matthias Börger, Jochen Ruß, Johannes Schupp, and Hans-Joachim Zwiesler for inspiring discussions, valuable comments, and support.

REFERENCES

- BAUER, D., BÖRGER, M. and RUSS, J. (2010) On the pricing of longevity-linked securities. *Insurance: Mathematics and Economics*, **46**(1), 139–149.
- BLAKE, D., CAIRNS, A.J.G., DOWD, K. and KESSLER, A.R. (2019) Still living with mortality: The longevity risk transfer market after one decade. *British Actuarial Journal*, **24**, e1, 1–80.
- BÖRGER, M. (2010) Deterministic shock vs. stochastic value-at-risk – an analysis of the Solvency II standard model approach to longevity risk. *Blätter der DGVFM*, **31**(2), 225–259.
- BÖRGER, M., SCHÖNFELD, J. and SCHUPP, J. (2019a) *Calibrating mortality processes with trend changes to multi-population data*. Ulm University working paper.
- BÖRGER, M. and SCHUPP, J. (2018) Modeling trend processes in parametric mortality models. *Insurance: Mathematics and Economics*, **78**, 369–380.
- BÖRGER, M., SCHUPP, J. and RUSS, J. (2019b) *It takes two: Why mortality trend modeling is more than modeling one mortality trend*. Ulm University working paper.
- BOYER, M.M. and STENTOFT, L. (2013) If we can simulate it, we can insure it: An application to longevity risk management. *Insurance: Mathematics and Economics*, **52**(1), 35–45.
- CAIRNS, A.J.G., BLAKE, D. and DOWD, K. (2006) A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *The Journal of Risk and Insurance*, **73**(4), 687–718.
- CAIRNS, A.J.G., DOWD, K., BLAKE, D. and COUGHLAN, G.D. (2014) Longevity hedge effectiveness: A decomposition. *Quantitative Finance*, **14**(2), 217–235.
- CHEN, H. and COX, S.H. (2009) Modeling mortality with jumps: Applications to mortality securitization. *The Journal of Risk and Insurance*, **76**(3), 727–751.
- CHOW, G.C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, **28**(3), 591–605.
- DICKEY, D.A. and FULLER, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, **74**(366a), 427–431.
- HABERMAN, S., KAISHEV, V., MILLOSIOVICH, P., VILLEGAS, A.M., BAXTER, S., GACHES, A., GUNNLAUGSSON, S. and SISON, M. (2014) Longevity basis risk: A methodology for assessing basis risk. Institute and Faculty of Actuaries, Sessional Research Paper.
- KWIATKOWSKI, D., PHILLIPS, P.C., SCHMIDT, P. and SHIN, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**(1–3), 159–178.
- LEUNG, M., FUNG, M.C. and O'HARE, C. (2018) A comparative study of pricing approaches for longevity instruments. *Insurance: Mathematics and Economics*, **82**, 95–116.
- LEVANTESI, S. and MENZIETTI, M. (2017) Maximum market price of longevity risk under solvency regimes: The case of Solvency II. *Risks*, **5**(2), 1–21.

- LI, J.S.-H., CHAN, W.-S. and CHEUNG, S.-H. (2011) Structural changes in the Lee-Carter mortality indexes: detection and implications. *North American Actuarial Journal*, **15**(1), 13–31.
- LI, J.S.-H., CHAN, W.-S. and ZHOU, R. (2017) Semicohherent multipopulation mortality modeling: The impact on longevity risk securitization. *The Journal of Risk and Insurance*, **84**(3), 1025–1065.
- LIN, Y. and COX, S.H. (2005) Securitization of mortality risks in life annuities. *The Journal of Risk and Insurance*, **72**(2), 227–252.
- LIU, Y. and LI, J.S.-H. (2016) The locally linear Cairns-Blake-Dowd model: A note on delta-nuga hedging of longevity risk. *ASTIN Bulletin*, **47**(1), 79–151.
- LOEYS, J., PANIGIRTZOGLU, N. and RIBEIRO, R.M. (2007) *Longevity: A market in the making*. JPMorgan Global Market Strategy.
- O'HARE, C. and LI, Y. (2015) Identifying structural breaks in stochastic mortality models. *Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, **1**(2), 1–14. <https://doi.org/10.1115/1.4029740>.
- SCHUPP, J. (2019) *On the modeling of variable mortality trend processes*. Ulm University working paper.
- SWEETING, P. (2011) A trend-change extension of the Cairns-Blake-Dowd model. *Annals of Actuarial Science*, **5**(2), 143–162.
- TAN, C.I., LI, J., LI, J.S.-H. and BALASOORIYA, U. (2014) Parametric mortality indexes: From index construction to hedging strategies. *Insurance: Mathematics and Economics*, **59**, 285–299.
- VAN BERKUM, F., ANTONIO, K. and VELLEKOOP, M. (2016) The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal*, **2016**(7), 581–603.
- VILLEGAS, A.M. and HABERMAN, S. (2014) On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in England. *North American Actuarial Journal*, **18**(1), 168–193.
- VILLEGAS, A.M., HABERMAN, S., KAISHEV, V.K. and MILLOSovich, P. (2017) A comparative study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin*, **47**(3), 631–679.
- WANG, S. (2002) A universal framework for pricing financial and insurance risks. *ASTIN Bulletin*, **32**(2), 213–234.
- WANG, S. (2007) Normalized exponential tilting. *North American Actuarial Journal*, **11**(3), 89–99.
- WÜTHRICH, M.V. (2016) *Market-Consistent Actuarial Valuation*. New York: Springer.
- WÜTHRICH, M.V. and MERZ, M. (2013) *Financial Modeling, Actuarial Valuation and Solvency in Insurance*. New York: Springer.
- ZEDDOUK, F. and DEVOLDER, P. (2019) Pricing of longevity derivatives and cost of capital. *Risks*, **7**(2), 1–29.

ARNE FREIMANN (Corresponding author)
Institute for Finance and Actuarial Sciences (ifa)
Lise-Meitner-Str. 14, 89081 Ulm, Germany
Institute of Insurance Science, University of Ulm
Helmholtzstr. 20, 89069 Ulm, Germany
E-Mail: a.freimann@ifa-ulm.de

APPENDIX

A. MODEL CALIBRATION

A.1. Data

For the reference population, we use data of the male population of England and Wales for the years 1841–2016 over the age range of 60–109 years, which is available in the Human

Mortality Database (data downloaded on 01 July 2018 from: <http://www.mortality.org>). Regarding the subpopulations, we rely on mortality data of English males sorted by quintiles of the Index of Multiple Deprivation (IMD) for the years 2001–2016 over the age range of 60–89 years, which we obtained from the Office for National Statistics. Note that the relative modeling setting can deal with a shorter data history and a different age range for the subpopulations than for the reference population. As the official measure for relative deprivation of small areas in England, the IMD measures deprivation in 32,844 areas with an average of 1500 residents each. Based on their IMD score, these areas are grouped into quintiles (ordered from the most to the least deprived) to obtain five homogeneous subpopulations of the male English population. We calibrate the multi-population CBD model structure via a standard two-stage maximum likelihood estimation (MLE) approach assuming binomially distributed deaths, see Villegas *et al.* (2017).

A.2. Extrapolation for high ages

When calibrating the CBD model to different socioeconomic subpopulations, one needs to be aware that the fitted mortality lines in the logit plot typically cross over somewhere in the highest age range. The reason for this lies in the structure of socioeconomic mortality differentials, which are usually more distinct for younger ages and diminish for high ages, see for example Villegas *et al.* (2017). To avoid an unrealistic reversal of the ranking among the subpopulations for highest ages in our simulations, we set the mortality rates equal to those of the reference population as soon as its mortality curve would be crossed, that is,

$$q_{x,t}^{[p]} := \begin{cases} \max \{ q_{x,t}^{[R]}, q_{x,t}^{[p]} \}, & \text{for } p = 1, 2 \\ \min \{ q_{x,t}^{[R]}, q_{x,t}^{[p]} \}, & \text{for } p = 4, 5 \end{cases}$$

and leave the middle subpopulation, which might lie (depending on the mortality scenario) above or below the reference population, unadjusted.

A.3. Calibration of stochastic trend process

For the calibration of the AMT process, we apply an iterative pseudo-MLE approach as proposed by Schupp (2019) and refer to this paper for technical details. Since data on historical mortality trend changes are sparse, we account for parameter uncertainty in the starting values of the simulation as well as in the trend change parameters.

Figure A.1 shows the historical period effects $\kappa_t^{(i)[R]}$, $i = 1, 2$ and the best possible realizations for the underlying trend processes $\hat{\kappa}_t^{(i)[R]}$, $i = 1, 2$ for all relevant potential numbers (k) of historical trend changes, which are identified by the calibration algorithm. The identified trend change years are given in Table A.1 along with the corresponding parameter estimates and assigned weights. We account for parameter uncertainty by sampling the starting values from the respective empirical distribution.

Moreover, we obtain the following trend change parameters:

$$\begin{aligned} (\hat{p}^{(1)[R]}, \hat{\mu}_M^{(1)[R]}, \hat{\sigma}_M^{(1)[R]}) &= (0.0223, -4.5453, 0.4105), \\ (\hat{p}^{(2)[R]}, \hat{\mu}_M^{(2)[R]}, \hat{\sigma}_M^{(2)[R]}) &= (0.0246, -7.4134, 0.2027) \end{aligned}$$

TABLE A.1
EMPIRICAL DISTRIBUTIONS FOR THE AMT STARTING VALUES.

k	$\hat{\kappa}_{t=0}^{(1)[R]}$	$\hat{d}_{t=0}^{(1)[R]}$	$\mathbb{P}(\%)$	Trend change years
2	-2.3099	-2.10×10^{-2}	6.24	1931, 1985
3	-2.3374	-2.43×10^{-2}	0.54	1931, 1980, 1997
4	-2.3020	-1.15×10^{-2}	93.22	1929, 1978, 1999, 2011
k	$\hat{\kappa}_{t=0}^{(2)[R]}$	$\hat{d}_{t=0}^{(2)[R]}$	$\mathbb{P}(\%)$	Trend change years
3	0.1144	5.85×10^{-4}	43.14	1902, 1930, 1972
4	0.1156	6.96×10^{-4}	3.89	1901, 1932, 1969, 1987
5	0.1143	3.55×10^{-4}	34.67	1900, 1932, 1969, 1988, 2009
6	0.1131	3.49×10^{-4}	18.30	1890, 1906, 1930, 1967, 1988, 2004

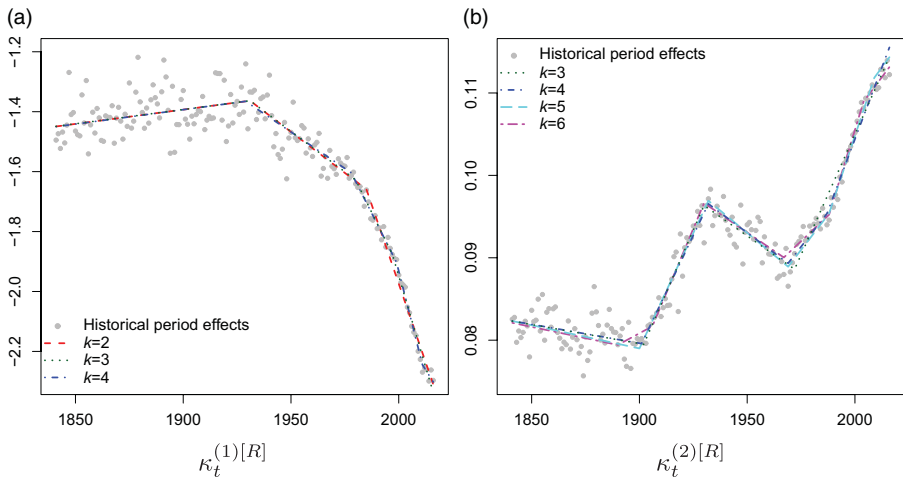


FIGURE A.1: Period effects for English and Welsh males (dotted), best possible realizations for the actual trend processes given different numbers of trend changes k (colored dashed lines).

with corresponding covariance matrices of standard errors of

$$SE^{(1)} = \begin{pmatrix} 1.353 \times 10^{-4} & 3.535 \times 10^{-5} & -1.754 \times 10^{-5} \\ 3.535 \times 10^{-5} & 4.616 \times 10^{-2} & 3.331 \times 10^{-4} \\ -1.754 \times 10^{-5} & 3.331 \times 10^{-4} & 2.322 \times 10^{-2} \end{pmatrix},$$

$$SE^{(2)} = \begin{pmatrix} 1.860 \times 10^{-4} & -1.211 \times 10^{-3} & 8.060 \times 10^{-4} \\ -1.211 \times 10^{-3} & 4.285 \times 10^{-2} & -2.153 \times 10^{-2} \\ 8.060 \times 10^{-4} & -2.153 \times 10^{-2} & 2.124 \times 10^{-2} \end{pmatrix}.$$

TABLE A.2

CHOW TEST STATISTICS FOR THE NULL HYPOTHESIS THAT THERE IS NO STRUCTURAL BREAK AT THE IDENTIFIED TREND CHANGE YEARS AND DICKEY-FULLER (DF) TEST STATISTICS FOR THE NULL HYPOTHESIS THAT THE TIME PROCESSES BETWEEN TWO TREND CHANGE YEARS HAVE A UNIT ROOT AGAINST THE ALTERNATIVE OF TREND-STATIONARITY; SIGNIFICANCE LEVELS: ***1%, ** 5%, *10%.

Period	Chow statistic	DF statistic	Period	Chow statistic	DF statistic
1841–1929		−8.160835***	1841–1902		−6.065415***
1929–1978	4.660071***	−6.187494***	1902–1930	116.0099***	−6.986161***
1978–1999	1.521318***	−6.799186***	1930–1972	143.1342***	−6.865163***
1999–2011	6.860339***	−2.818677	1972–2016	331.7755***	−2.514045
2011–2016	2.478597***	−3.428520*			

(a) $\kappa_t^{(1)[R]}$; trend change years 1929, 1978, 1999, 2011

(b) $\kappa_t^{(2)[R]}$; trend change years 1902, 1930, 1972

To account for parameter uncertainty, we follow Börger *et al.* (2019a): for each simulation path, we generate a multivariate normal random vector with mean equal to the estimated trend change parameters and covariance matrix $SE^{(i)}$. Subsequently, the first component of the vector is transformed to a beta distribution with same mean and variance to obtain trend change probabilities between zero and one. Similarly, the third component is transformed to a gamma distribution with unchanged mean and variance to ensure positivity.

Finally, the covariance matrix for random fluctuations around the AMT is estimated as:

$$\Sigma^{[R]} = \begin{pmatrix} 3.865 \times 10^{-4} & 1.720 \times 10^{-5} \\ 1.720 \times 10^{-5} & 2.036 \times 10^{-6} \end{pmatrix},$$

for which we neglect parameter uncertainty since it is typically not material.

A.4. Statistical evidence for a piecewise linear trend-stationary process

Having identified potential structural break points in the historical CBD time series with the algorithm of Schupp (2019), we evaluate their statistical significance. Following Sweeting (2011), we partition the historical time series into consecutive sections separated by the identified break points and apply the test of Chow (1960). Under the assumption of normally distributed error terms with constant variance, we test the null hypothesis that the coefficients in two linear regressions fitted to each consecutive pair of sections are equal. The results for both CBD time processes are given in Table A.2. For the sake of brevity, we only show the results for the break points with highest likelihood according to Table A.1. The null hypothesis of no structural break is rejected for all identified break points at a 1% level of significance which provides strong statistical evidence for the presence of several mortality trend changes.

Furthermore, we apply the test designed by Dickey and Fuller (1979) to each of these consecutive sections of period effects to test the null hypothesis of a unit root against the alternative of trend-stationarity. The results are included in Table A.2. The Dickey–Fuller (DF) test suggests that the historical period effects are trend-stationary rather than difference-stationary in four out of five periods for $\kappa_t^{(1)[R]}$ and in three out of four periods for $\kappa_t^{(2)[R]}$, respectively, with the level of significance being 1% in most cases. Overall, our findings are in line with those of Sweeting (2011) and provide strong statistical evidence

TABLE A.3

KPSS TEST STATISTICS FOR THE NULL HYPOTHESIS THAT THE TIME PROCESSES OF SOCIOECONOMIC MORTALITY DIFFERENTIALS ARE LEVEL-STATIONARY AGAINST THE ALTERNATIVE OF A UNIT ROOT; SIGNIFICANCE LEVELS: ***1%, ** 5%, *10%.

	IMD-1	IMD-2	IMD-3	IMD-4	IMD-5
$\kappa_t^{(1)[p]}$	0.62344710**	0.55980637**	0.09315834	0.61351390**	0.60069293**
$\kappa_t^{(2)[p]}$	0.29389570	0.28458448	0.08543594	0.37001952*	0.32292157

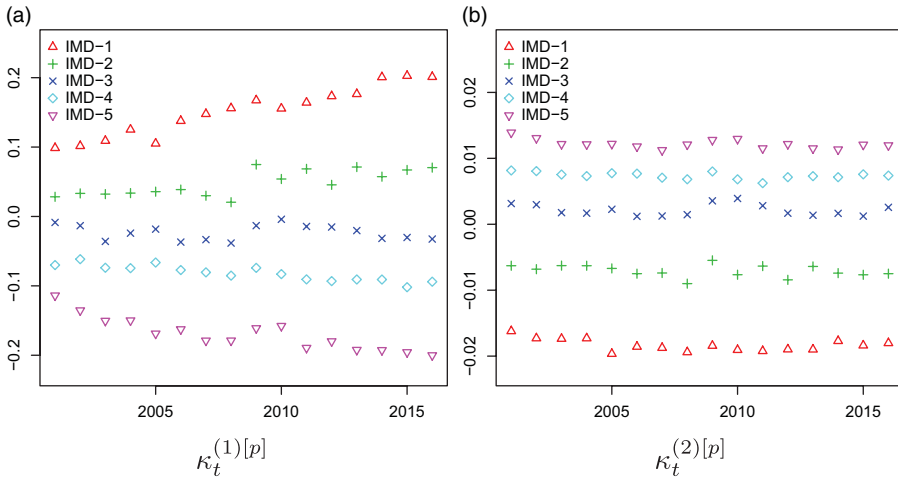


FIGURE A.2: Historical time series of socioeconomic mortality differentials for IMD quintiles.

for using a trend-stationary process around piecewise linear trends with changing slopes over time.

A.5. Calibration of time processes for socioeconomic mortality differentials

Figure A.2 shows the historical time series of subpopulation-specific mortality differentials $\kappa_t^{(1)[p]}$ (in the left panel) and $\kappa_t^{(2)[p]}$ (in the right panel), respectively. Both period effects reveal a clear and intuitive ranking among the subpopulations in terms of socioeconomic status.

To identify a suitable model of the class of VARIMA(p, d, q) models, we apply a Box–Jenkins approach: first, we observe a widening of the gap between the least and the most deprived subpopulations over the calibration period in the left panel of Figure A.2, which indicates non-stationarity for $\kappa_t^{(1)[p]}$. This is substantiated by the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (cf. Kwiatkowski *et al.*, 1992) in Table A.3 which rejects the null hypothesis of (weak) level-stationarity for $\kappa_t^{(1)[p]}$ for all but the middle subpopulation at a 5% level of significance. After applying a differencing approach, the integrated time series appears to be (weakly) stationary; in particular repeated application of the KPSS test does not reject the null hypothesis of first-order difference-stationary for any subpopulation at a

5% level of significance. Since we could also neither find any significant signs of seasonality nor any indications for the existence of $q > 0$ moving-average terms in the autocorrelation plots, we restrict ourselves to the class of VARIMA($p, 1, 0$) models. Relying on the BIC, we finally find that a VARIMA(0, 1, 0) model, that is, a multivariate RWD, provides the best compromise between a satisfactory fit and a parsimonious model structure with a BIC of -1440.77 . For comparison, its counterpart without drift provides a slightly worse value of -1414.23 and we obtain values between -1318.59 and -1338.23 for VARIMA(1, 1, 0) models by considering various restrictions for the autoregressive matrix. The RWD is calibrated via a standard MLE approach using all 16 years of available data.

A.6. Calibration of optimal EMT weights

As suggested by Börger *et al.* (2019b), we determine the optimal EMT weighting by minimizing the mean squared errors between the unobservable AMT (or drift, respectively) and the derived EMT in a Monte Carlo simulation and obtain $(\psi^{(1)}, \psi^{(2)}) = (2.225, 2.752)$ for the trend process, $(7.364, 10.212)$ when setting the trend change probabilities to zero, and $(39.828, 89.818)$ for the RWD. For the latter, we restrict the regression to the most recent 20 data points in line with the calibration period. Higher weights for the variants with constant trends/drifts indicate that less weight should be assigned to the most recent data points.