# Basis Risk in Index Based Longevity Hedges: A Guide For Longevity Hedgers

Andrew J.G. Cairns[1], [2] Ghali El Boukfaoui[3] [4]

**Abstract**

This paper considers the assessment of longevity basis risk in the context of a general index-based hedge. We develop a detailed framework for measuring the impact of a hedge on regulatory or economic capital that takes population basis risk explicitly into account. The framework is set up in a way that accommodates a variety of regulatory regimes such as Solvency II as well as local actuarial practice, attempting, therefore, to bridge the gap between academia and practice.

This is followed by a detailed analysis of the capital relief resulting from a hedge that uses a call spread as the hedging instrument. We find that the impact of population basis risk on capital relief (expressed in terms of a 'haircut' relative to the case with no population basis risk) depends strongly on the exhaustion point of the hedge instrument. In particular, in a Solvency II setting, if the exhaustion point lies well below the 99.5% Value-at-Risk, population basis risk has negligible impact and the haircut is zero.

# 1 Introduction

This paper concerns the assessment of index-based longevity hedges, and seeks to establish a common understanding between regulators and insurers of how to assess the impact of population basis risk on capital charges and capital relief.

When considering the various forms of longevity risk transfer chain, one can reasonably assume that the first link of the chain is key in the ability to accomplish an effective risk transfer along the chain. In particular, if the form of the risk is opaque or if its understanding requires an insider knowledge of the population for which longevity risk is to be transferred, the chain will likely be limited in both depth and width. For example, the chain could end up being limited to one level of transfer (primary writers to first risk

---

[1]Maxwell Institute for Mathematical Sciences, Edinburgh; and Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS, UK. Email: A.J.G.Cairns@hw.ac.uk

[2]Actuarial Research Centre, Institute and Faculty of Actuaries

[3]Financial Institutions Group, Societe Generale Corporate and Investment Bank, 245 Park Avenue, NY 10167 New York, USA.

[4]This paper has been commissioned and funded by Societe Generale, and is co-authored by Andrew Cairns and Ghali El Boukfaoui. The views and opinions expressed in this article are those of the authors and do not necessarily reflect those of Societe Generale or any of its affiliates.

taker only, without further dissemination of the risk to other market players) with only a handful of risk takers who are willing to take on this risk. These risk takers can be, for example, insurers or re-insurers who might have some appetite for this risk due to their business mix at a specific point in time. Typically, the cedent is a pension plan, and the risk exchanged is directly linked to the cedent's own mortality experience rather than a public mortality index.

While this type of chain has proved to be very successful (see, for example, Kessler et al., 2015), the capacity of the limited number of risk takers will ultimately limit the potential for further longevity risk transfer. Further innovation is required to achieve a wide allocation of risk amongst a diversified set of risk takers. This would allow the impact of a materialized longevity risk to be spread across a more diversified set of financial institutions with much greater capacity to absorb longevity risk, and reduce the systemic impact of longevity-linked events. From the perspective of the hedger, expansion of the market to a more diverse set of risk takers has a further advantage: the combination of financial innovation and a wider range of risk takers should make the market in longevity risk more competitive and reduce the risk transfer premiums paid by hedgers.

In order to create a more effective risk transfer chain, the risk should be defined in a way that allows a broader spectrum of market participants (i.e. those with the capacity to influence pricing in a competitive market) to quantify and analyze the longevity risk without knowledge of the specific population underpinning the longevity risk to which the primary cedent is exposed. Some of these non-traditional market participants will value the uncorrelated nature of longevity risk to other more traditional asset classes, a feature that adds to the attractiveness of the longevity risk investment and puts additional pressure on prices to the benefit of the primary risk hedger.

The market expansion ideas expressed above build on the work of Blake et al. (2006) and Cairns et al. (2008). Specifically, Section 6.4 in Cairns et al. (2008) outlines a simplified risk-transfer chain. In the first link of the chain, pension plans enter into a customised longevity risk transfer with, typically, a reinsurer. In the second link of the chain, the reinsurer, having acquired a diversified set of longevity exposures, enters into an index-based longevity hedge with the wider capital markets. A key outcome is that the reinsurer then has increased capacity to engage with primary hedgers (pension plans). This paper focuses on the second link in this chain, and, in particular on the perspective of an insurer or reinsurer ceding longevity risk using an index-based hedge. Cairns et al. (2008) envisaged a standard, index-based longevity bond for this second link, Coughlan et al. (2007) a $q$-forward, while Michaelson and Mulholland (2015) outlined the concept of a call spread option which we develop in detail in this paper. Blake et al. (2013) discuss how longevity satisfies key conditions for development of a successful market, particularly with the use of standardised indices and security design.

The efficiency of the longevity risk transfer chain depends largely on the form of the initial layer of risk transfer. Based on this principle, we are considering an index based longevity hedge which allows for a competitive dynamic. This index based longevity hedge will be defined in Section 3 and is designed to facilitate hedging of longevity trend risk through the use of a customized index based on publicly available general population data, that leaves the hedger (typically an insurer or pension fund) exposed to some level of basis risk.

The objective of this paper is to provide a framework for modelling population basis risk and quantifying its impact on the effectiveness of a longevity hedge, and provides a concrete implementation of the framework proposed by Cairns et al. (2014). The process as a whole might seem (and necessarily is) complex, but we break it down into a series of small, manageable, understandable and easily documentable steps.[5] This should allow market participants (risk hedgers, regulators etc) to understand better the allocation between their exposure to the general population longevity trend risk and their specific population longevity risk. In turn, this allows them to make an informed decision with respect to the cost of transferring each component of longevity risk.

The proposed modelling and assessment framework will enable hedgers to make a fair comparison of the different longevity risk management options. Given a decision to reduce their exposure to longevity risk, some longevity hedgers might find it attractive to retain this basis risk (for example, they have more expertise assessing their own idiosyncratic risks, and some appetite for retaining it) and transfer the more general population risk to a large market at an attractive price. The risk measure considered in this paper is captured in the context of an economic capital model which can be considered similar to that of an insurance entity subject to Solvency II, although we will avoid any loss of generality in the approach presented.

The structure of the paper is as follows. In Section 2 we outline the liability to be hedged, with the index-based hedge instrument to be used in Section 3. Section 4 introduces the regulatory context. Section 5 outlines the framework for modelling population basis risk and quantifying its impact. This forms the core of the paper and is broken down into a series of 22 easy steps. Section 6, by way of example, outlines a two population model that can be used as part of this framework, and Section 7 provides some numerical results for a typical portfolio of deferred and immediate annuitants. Section 9 concludes.

# 2 The Liability

## 2.1 The underlying liability

Let's assume that the liability holding the longevity risk can be represented by a set of model points, each of which has well defined characteristics. These characteristics include (but are not limited to) the following:

[5] Good documentation reflects good internal risk management practice, but also helps facilitate discussions with regulators.

- age of the primary annuitant;
- gender of the primary annuitant;
- age of the secondary annuitant if any;
- gender of the secondary annuitant if any;
- annual Benefit of the main and secondary annuitants;
- age at which benefits start.

These elements all together are generally called the Exposure.

The liability being considered is the present value (or accumulated value, depending on what time the liability is being evaluated at) of the annuity obligations. For valuation purposes, a point estimate at time 0 is required (e.g. under Solvency II this would be a best estimate plus the risk margin). In many countries this is achieved pragmatically by following a central path for general population mortality, and then applying experience ratios to adjust for differences in the level of mortality in the two populations (more detail later in this section). These experience ratios are typically sex and age dependent and capture the ratio between the specific population mortality and the general population mortality. The experience ratios can also be time dependent if the hedger considers that the relationship between the central paths of specific and general population mortality are expected to change over time.

The hedge index (also called the *hedge underlying*) will be defined in the next section and intends to replicate as closely as possible the future value of the liabilities in a way that minimizes basis risk, while only using general population mortality. Some basis risk will remain as the actual realization of the general population mortality, even adjusted with experience ratios, will not follow exactly the realization of the specific population mortality. Also, the future liability reserving calculation, seen from today, will likely change due to the future projected relationship between the hedger's book and the general population. All these are potential sources of basis risk and will be captured in the modeling described in later sections.

We now introduce the following definitions and notations:

- $L =$ time zero present value (PV) of the liability to be hedged: a random variable that will only be known at the time of the last death.
- $L(0)$ is a point estimate of the liability $L$ at time 0, and is defined to be identical to the reserve put aside by the insurer with respect to their future obligations. (For example, this could be defined as the technical reserve under Solvency II.)
- $L(T) =$ is a point estimate of the liability $L$ at time $T$, given the mortality information up to time $T$. We express this as the accumulated value of actual liabilities

up to $T$ plus a point estimate the present value at $T$ of liabilities falling after $T$, although this can then be discounted to time zero for more direct comparison with $L(0)$. $L(T)$ is a random variable that will not be known until time $T$.

There are two populations: the hedger's own population, P, and the general (or reference) population, G, each with its own mortality experience.

We assume the liabilities within the hedger's population are evaluated with reference to *model points* $i = 1, \ldots, N$, as discussed previously. In the discussion below, each model point has its own associated sequence of benefit payments and cohort mortality rates. Thus, for example, mortality rates are defined with reference to the model point and the year, rather than the more conventional age and year.

- $q_1^P(i,t) = Pr($individual in $i$ dead by time $t$ given alive at time $t-1) =$ retrospective *underlying* population P mortality rate corresponding to model point $i$ primary lives.[6]

- $S_1^P(i,t) =$ proportion of population P model point $i$ primary lives to survive up to time $t$ based on (simulated) *underlying* mortality rates: that is, given the retrospective underlying mortality rates, $S_1^P(i,t)$ is the retrospective probability of surviving to time $t$. Thus,
$$S_1^P(i,t) = \prod_{y=1}^{t} \left(1 - q_1^P(i,y)\right).$$

For spouses/partners, we have corresponding mortality rates, $q_2^P(i,t)$ and survival probabilities $S_2^P(i,t)$. In probabilistic terms, the lifetimes of individuals and their spouses/partners are, here, assumed to be independent random variables. Consequently, $(1 - S_1^P(i,t))S_2^P(i,t) =$ the proportion of model point $i$ at $t$ under which the primary life has died, but the spouse/partner has survived to $t$.[7]

Model point $i$ benefits are:

- $B_1(i,t) =$ benefit payable to primary lives at time $t$ if the primary lives of model point $i$ are still alive at time $t$.[8]

- $B_2(i,t) =$ benefit payable to spouse/partner at time $t$ if the primary lives of model point $i$ are dead at $t$ but the corresponding spouse/partners are still alive.

[6] Note that we are not including the idiosyncratic risk associated with the individual random lifetimes given the simulated mortality rates in each stochastic scenario. In proportional terms, the additional idiosyncratic risk will become smaller, the larger the underlying number of lives (assuming no concentration of risk by amounts).

[7] Independence is assumed here for simplicity within the context of the larger framework. The broader framework can, of course, accommodate dependence between primary and secondary lives.

[8] We assume here for simplicity of exposition that payments are made at the end of each year. The framework can, of course, be adapted to account for monthly or other payment patterns.

For the general population, we have corresponding mortality rates. For example, for general population lives corresponding to the model point $i$ primary lives, the *underlying* mortality rate is $q_1^G(i,t) = Pr(i$ dead by time $t$ given they were alive at $t-1)$, while $q_2^G(i,t) =$ general population *underlying* mortality rate corresponding to the spouse/partner of model point $i$.

Knowledge of historical mortality rates for both populations allows us to calculate experience ratios that specify the anticipated population P mortality rate as a proportion of the population G mortality rate. Thus

- $\varepsilon_1^0(i,t) =$ experience ratio calibrated at time 0, for valuation purposes at time 0, linking the general population to the hedger's mortality for model point $i$ in year $t$, discussed at the beginning of this section, and would normally depend on age, gender and time. These typically reflect the historical relationship between mortality rates in the two populations by age and gender. In some cases they might simply reflect the ratio of mortality rates in the two populations in the most recent year or years. Alternatively they might be calibrated using median projections of a multi-population stochastic mortality model.

- $\varepsilon_2^0(i,t) =$ experience ratio calibrated at time 0, for valuation purposes at time 0, linking the general population to the hedger's mortality for the spouse/partner of $i$. These will normally depend on age, gender and time.

- Experience ratios are recalibrated from time to time to reflect the most recent experience. Notationally, therefore, $\varepsilon_1^T(i,t)$ and $\varepsilon_2^T(i,t)$ represent the future experience ratios (recalibrated at time $T$) when using the information available up to time $T$ which includes the realized mortality of the specific population P, and general population G. These experience ratios are used for reserving at time $T$ for benefits payable after $T$, in the same way as $\varepsilon_{1/2}^0(i,t)$ are used for reserving at time 0 of benefits payable after time 0.

Experience ratios are also required as part of the hedge contract and require specification at time 0. Potentially, these might be different from those used in the time zero valuation of liabilities (taking into account potential lapses for instance). Those specified in the hedge contract will be denoted by $\varepsilon_j^{H0}(i,t)$ to allow for this differentiation. In practice, though, these are likely to be the same: that is, $\varepsilon_j^{H0}(i,t) = \varepsilon_j^0(i,t)$ for all $i$, $j$ and $t$.

The liability reserve survivorship functions theoretically use the specific population P mortality, which then requires point estimates for valuation.

$$S_1^P(i,t) = \prod_{y=1}^{t} \left(1 - q_1^P(i,y)\right) \tag{1}$$

$$S_2^P(i,t) = \prod_{y=1}^{t} \left(1 - q_2^P(i,y)\right). \tag{2}$$

For valuations, point estimates of these quantities are required. Theoretically, this might be achieved through the taking of expectations, but in practice, in many countries, valuations are carried out deterministically using, for example, the median path of the distribution for general population mortality (resulting, then, in estimated mortality rates $\hat{q}_1^{LG}(i,y)$ and $\hat{q}_2^{LG}(i,y)$).[9] Valuation experience ratios are then applied to convert to specific population mortality. The present value of the obligations (also called liability reserve or technical reserve) is defined, therefore, by the following expression:

$$L(0) = \sum_{i=1}^{N}\sum_{t=1}^{M} P(0,t)\left(\hat{S}_1^P(i,t)B_1(i,t) + (1-\hat{S}_1^P(i,t))\hat{S}_2^P(i,t)B_2(i,t)\right), \tag{3}$$

where $\hat{S}_j^P(i,t) = \prod_{y=1}^{t}\left(1-\varepsilon_1^0(i,y)\hat{q}_1^{LG}(i,y)\right)$ represents the estimate of $S_j^P(i,t)$, for $j = 1,2$, based on the $\hat{q}_j^{LG}(i,y)$, $M$ is the latest possible payment date across all model points, and the general discount function $P(t,T)$ represents the zero-coupon bond price at time $t$ for 1 unit of currency payable at $T > t$.

If benefits are payable other than at the year end, survival probabilities need to be adapted to incorporate fractions of a year using clearly state assumptions about the pattern of mortality within a year. Similarly, a monthly or continuous discount curve needs to be used to take account correctly for the timing of each payment. The simplifying assumptions (annual payments in both the liabilities and, in the next section, the synthetic hedge underlying) should not, therefore, invalidate the main conclusions of this paper.

## 2.2 Liability valuation at time $T$

Evaluation of hedge effectiveness requires knowledge of the distribution of the liability at time $T$, $L(T)$.

For consistency with the time 0 valuation, this is computed as follows:

$$\begin{aligned} L(T) \quad = \quad & \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{P(0,t)}{P(0,T)}\left(S_1^P(i,t)B_1(i,t) + (1-S_1^P(i,t))S_2^P(i,t)B_2(i,t)\right) \\ & + \sum_{i=1}^{N}\sum_{t=T+1}^{M}\frac{P(0,t)}{P(0,T)}\left(\tilde{S}_1^P(i,t)B_1(i,t) + (1-\tilde{S}_1^P(i,t))\tilde{S}_2^P(i,t)B_2(i,t)\right), \end{aligned} \tag{4}$$

where the reserve survivorship functions use (a) actual hedge population mortality, $q_j^P(i,y)$ up to time $T$, and (b) *recalibrated* experience ratios in combination with projected general

[9] The 'L' in the superscript 'LG' is included to emphasize that the projection is for liability valuation.

population mortality.[10] Thus, for $j \in \{1,2\}$:

$$\begin{aligned} \text{for } t \leq T,\ S_j^P(i,t) &= \prod_{y=1}^{t} \left(1 - q_j^P(i,y)\right) \ \text{(defined previously)} \\ \text{and for } t > T,\ \tilde{S}_j^P(i,t) &= S_j^P(i,T) \prod_{y=T+1}^{t} \left(1 - \varepsilon_j^T(i,y)\bar{q}_j^{LG}(i,y)\right). \end{aligned} \quad (5)$$

where $\bar{q}_j^{LG}$ represents the general population projected mortality for the primary ($j = 1$) or secondary life ($j = 2$) using a mortality model calibrated at time $T$ using all of the general population information available up to time $T$.[11] The model to be used in making this projection and its calibration do not need to be the same as those used in the calculating the hedge instrument payoff. Additionally, the experience ratios, $\varepsilon_j^T(i,y)$ are recalibrated at time $T$ using the experience data from both populations up to time $T$, using a method that is consistent with that used at time 0. From the time 0 perspective, this model and the method of calibration should represent our best estimate of what will be used by the valuation actuary at time $T$.

# 3 The synthetic longevity hedge

The hedge underlying will be constructed in a manner that intends to replicate the liability valuation at time $T$ as closely as possible while keeping the general population mortality as the only source of uncertainty. Because the hedging instrument has a defined maturity $T$, and in order to capture the potential future annuity payments that are scheduled beyond that maturity, we will also introduce a concept of future liability valuation which would capture both accumulated cash flows up to $T$ and the present value of cash flows after $T$.

## 3.1 The hedging instrument

We start with the following definitions:

- $T$ = payoff date (maturity of the hedge)
- $X(T)$ = hedge underlying (still to be defined)
- $H(T) = u\, h(X(T))$ = hedge instrument payoff at time $T$ defined as a call spread:

$$h(x) = \max\left\{0,\ \min\left\{\frac{x - AP}{EP - AP},\ 1\right\}\right\}$$

[10] For consistency with the standard formula approach in Solvency II, the discount curve used here is $P(0,t)/P(0,T)$ rather than $P(T,t)$. This provides a capital requirement for longevity risk on its own, which is then separately blended in with the capital requirements for other risks, including interest-rate risk.

[11] And, again, the LG superscript indicating that this projection is for liability valuation purposes.

where $AP$ refers to the Attachment Point, and $EP$ the Exhaustion Point

- $u =$ notional quantity.

When the hedge is in place, the liability is netted with the hedge payoff at time $T$ and the hedged liability can be defined by the net portfolio value at time $T$

$$P = L(T) - H(T).$$

## 3.2 The hedge underlying

For a general index-based hedge, the underlying $X(T)$ needs to derive its value from observed mortality rates in the general population only, up to time $T$. We now propose a specific form that aims to match the liability dollar for dollar, replacing $X(T)$ by the synthetic liability, $\tilde{L}(T)$:

$$\begin{aligned} \tilde{L}(T) &= \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{P(0,t)}{P(0,T)}\left(S_1^{HG}(i,t)B_1(i,t) + (1-S_1^{HG}(i,t))S_2^{HG}(i,t)B_2(i,t)\right) \\ &\quad + \sum_{i=1}^{N}\sum_{t=T+1}^{M}\frac{P(0,t)}{P(0,T)}\left(\bar{S}_1^{HG}(i,t)B_1(i,t) + (1-\bar{S}_1^{HG}(i,t))\bar{S}_2^{HG}(i,t)B_2(i,t)\right). \end{aligned} \tag{6}$$

In (6), for $j \in \{1,2\}, t \leq T$:

$$S_j^{HG}(i,t) = \prod_{y=1}^{t}\left(1 - \varepsilon_j^{H0}(i,y)q_j^G(i,y)\right). \tag{7}$$

$S_j^{HG}(i,t)$ uses actual mortality of the general population in each of years 1 to $T$, adjusted by the experience ratios specified in the contract at time 0, $\varepsilon_j^{H0}(i,y)$. It is worth noting that, by construction, the formulation of the hedge underlying is very similar to the liability reserve formula defined earlier (equation (4)).

For $j \in \{1,2\}, t > T$

$$\bar{S}_j^{HG}(i,t) = S_j^{HG}(i,T)\prod_{y=T+1}^{t}\left(1 - \varepsilon_j^{H0}(i,y)\bar{q}_j^{HG}(i,y)\right). \tag{8}$$

In (8), $\bar{q}_j^{HG}$ represents the general population projected mortality for the primary ($j = 1$) or secondary life ($j = 2$) using a mortality model calibrated at time $T$ using data for the general population only, and using all information available up to time $T$. The model to be used in making this projection and its calibration are part of the hedge contract at time 0. For the hedge to be most effective, this specification should represent the best estimate at time 0 of the reserving model that will be used at time $T$. This will normally be the reserving model used at time 0 (for example, a combination of general population model calibration and projection combined with the time 0 experience ratios).

The notional quantity of the option is chosen to be $u = EP - AP$. This is a natural choice that reflects the careful formulation of $\tilde{L}(T)$, which should approximately match uncertainty in the true liability, $L(T)$, dollar for dollar. (For a given mathematical objective, it might be possible to optimise over $u$, but we then lose the intuition that underpins the choice of $u = EP - AP$.)

# 4 Context of a capital model

The risk of under reserving with respect to longevity is generally captured through an economic capital model. This model will rely on the distribution of $L(T)$ (alternatively $L$ or $L(1)$) or be based on a formulaic shock that would be prescribed by a regulator. Generally speaking, the capital charge can be described as:

$$C = VaR(L(T), \alpha) - E(L(T)) \tag{9}$$

where $\alpha$ is the confidence level at which the Value-at-Risk (VaR) of the capital model is calculated.[12]

## 4.1 Proxy for the Standard Formula

In Solvency II regulated jurisdictions, there is a non-negligible number of companies who are using the Standard Formula for the determination of the capital charge. The Standard Formula uses a deterministic shock to the central path of mortality to determine the stressed amount of the liability and therefore set the capital charge. In the numerical implementation section of this paper, we will detail a proxy approach that allows for the estimation of basis risk while assuming the use of the Standard Formula for the capital charge. The approach relies on creating a proxy distribution of liabilities that would imply a capital charge that is identical to the one produced following the Standard Formula. That distribution is not meant to represent an actual internal model but rather to be only used as a proxy for the purpose of the basis risk analysis.

The proxy approach allows the basis risk quantification approach to be described assuming use of an internal model.

## 4.2 Capital Relief

When the hedge $H$ is in place, the liability distribution is impacted by the hedge overlay. The capital charge with a hedge in place can be described with the following formula:

$$\tilde{C} = VaR(L(T) - H(T), \alpha) - E(L(T) - H(T)). \tag{10}$$

[12]The capital charge might, alternatively, and depending on the view of regulators, refer to the median rather than the mean of $L(T)$.

The capital relief is then defined as:

$$\begin{aligned} R &= C - \tilde{C} \qquad (11) \\ &= \left(VaR(L(T),\alpha) - E(L(T))\right) - \left(VaR(L(T)-H(T),\alpha) - E(L(T)-H(T))\right). \end{aligned}$$

In the calculation of the "Capital Relief", the distribution of $L(T)$ can be either:

1. Dependent on the variations of general population mortality only, this would implicitly assume that the realization of the specific population mortality would follow the experience ratios defined at time zero: that is, we assume that $q_j^P(i,t) = \varepsilon_j^{H0}(i,t) q_j^G(i,t)$, meaning $L(T) = \tilde{L}(T)$. The capital charge (and therefore the capital relief) would, therefore, be calculated in the absence of basis risk. We will denote it by $C_1$ where $C_1 = f(q^G)$, and $q^G$ is shorthand for information about general population mortality up to $T$.

2. Dependent on both the variations of general and specific population mortality, the time 0 experience ratios, $\varepsilon_j^{H0}(i,t)$, are still used for calculating the hedge payoff but the liability realization up to the maturity of the Hedge T will be represented directly through the specific population mortality. The liability best estimate will then be represented through the recalibrated central path of mortality at maturity as well as the recalibrated experience ratios at maturity, $\varepsilon_1^T(i,t)$ and $\varepsilon_2^T(i,t)$. The capital charge (and therefore the capital relief) would then be inclusive of basis risk. We will denote it by $C_2$ where $C_2 = \tilde{f}(q^G, q^P)$, and $q^P$ represents additional information about the specific population mortality up to $T$.

In summary, the hedge underlying is purely dependent on the general population mortality. The liability, on the other hand, can be either defined with the use of deterministic experience ratios or through the actual realization of the specific population mortality.

# 5 Framework for the calculation of capital relief

In this section we outline a general framework that can be used in the calculation of the capital charge with and without a hedge in place, and develops in more detail an earlier framework proposed by Cairns et al. (2014).

The framework is outlined in Figure 1 and has been broken down into a set of 22 steps, with each step representing a specific set of calculations or actions. Broadly speaking our objective is as follows:

- Collect historical data that is relevant in the calculation of the capital charge.
- Put a hedge instrument in place with payoff at time $T$.
- Use simulated data up to time $T$ and further projections based on those simulations from time $T$ to produce simulated values for the liability and the hedge instrument payoff at time $T$.

- Use the results to calculate the hedge effectiveness and therefore and appropriate amount of capital relief taking into account basis risk (through a haircut).

Full details of each step, including the potential for variants, are given in Appendix A, but we summarize here the higher-level elements.

- Steps 1 and 2: Gather relevant historical data for the general and specific populations, G and P, up to time 0. Use this data to calculate the experience ratios that will be embedded at time 0 in the calculation of the hedge instrument payoff (or use other experience ratios that may best reflect the hedger's view on future specific population mortality compared to general population mortality).

- Steps 3, 4 and 5: **The simulation model:** Specify the two-population mortality model and the underlying time series processes for projecting death rates. Estimate the age and period effects and time series parameters for forecasting period effects.

- Steps 6, 7, 14A, 15, 17: use the simulation model to simulate death rates from time 0 to time $T$. Extrapolate the rates to high and low ages, if required, to calculate the time 0 to $T$ components of the actual liability values and the synthetic liability values underlying the hedge-instrument payoff at time $T$.

- Steps 8, 9, 10, 11 and 14B: **The hedge payoff model:** The hedge-instrument payoff requires projected death rates for the general population using a model specified at time 0, a recalibration method also specified at time 0, and the recalibration itself to be carried out at time $T$ using the historical and simulated death rates up to time $T$. If required, the model specifies how death rates should be extrapolated to high and low ages. Use the hedge payoff model to project the general population death rates beyond time $T$ and to high and low ages.

- Steps 12, 13, 16 and 18: **The liability valuation model:** Identified at time 0, the model that is most likely to be used at time $T$ to value the specific-population liabilities at time $T$, including a recalibration methodology. This model might be different from the time 0 liability valuation model, the time 0 simulation model and the time $T$ hedge payoff model. Use the scenario realization up to time $T$ to fit the model, calibrate parameters and calculate projected death rates beyond $T$ for the specific population with extrapolation to high and low ages.

- Steps 19, 20 and 21: Use experience ratios embedded in the hedge instrument contract at time 0, along with simulated (up to time $T$) and projected death rates (beyond time $T$) for the general population, to calculate the synthetic liability, $\tilde{L}(T)$ that underlies the hedge-instrument payoff. Calculate the hedge-instrument payoff, $H(T)$, that is derived from the underlying $\tilde{L}(T)$.

- Step 22: Use the simulated (up to time $T$) and projected death rates (beyond time $T$) for the hedger's population, P, to calculate the actual liability, $L(T)$.

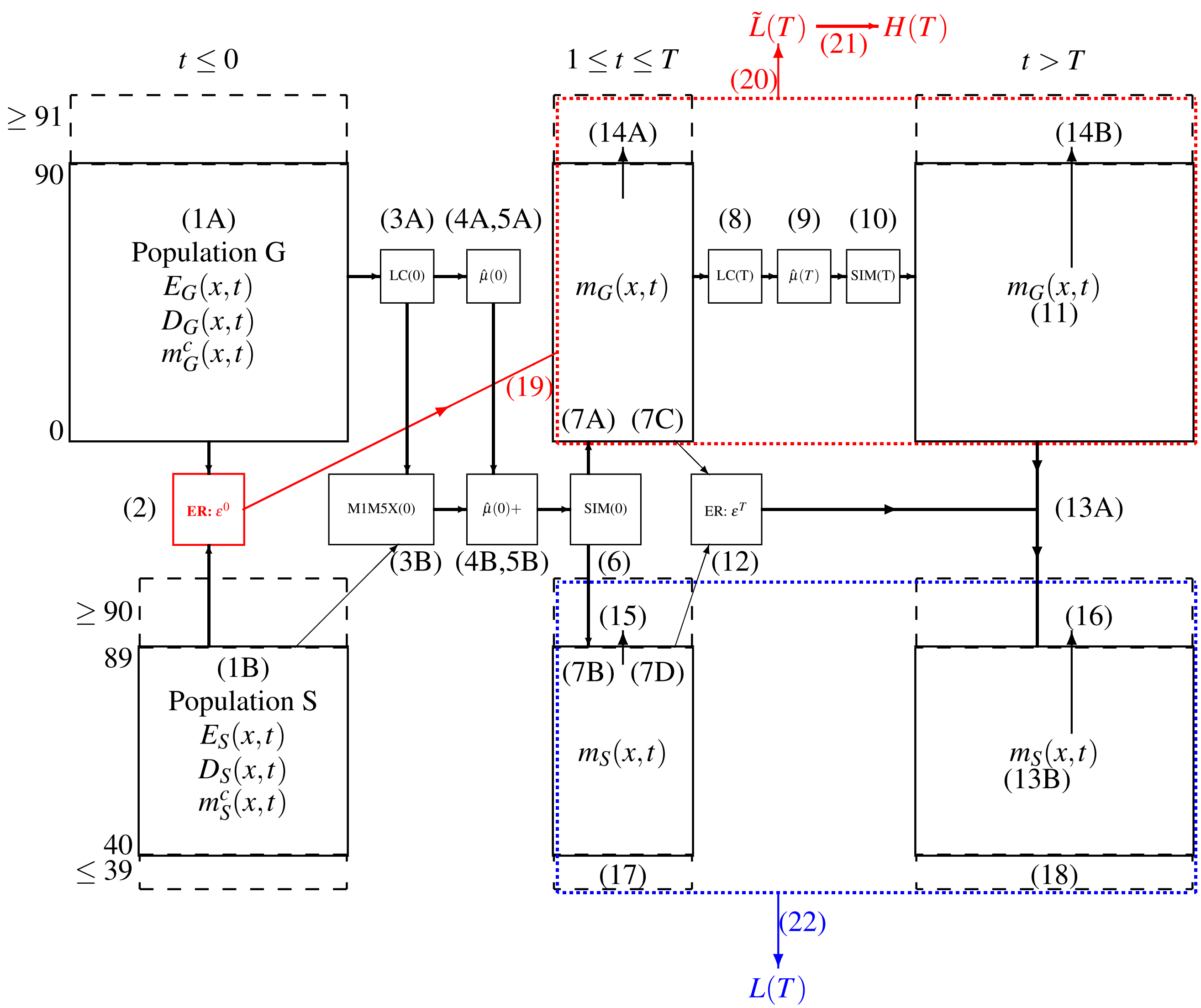



Figure 1: Steps required in the calculation of longevity capital requirements. For a detailed description of steps 1 to 22, see Appendix A.

# 6 A two-population simulation model

Analyzing the effectiveness of a hedge and its impact on capital charges requires the use of a full two-population mortality model. This model needs to cover all relevant ages that have an impact on liability cashflows and the calculation of the hedge instrument payoff.

In what follows, $m^G(x,t)$ and $m^P(x,t)$ represent the age-specific death rates in the general and specific populations respectively for age $x$ last birthday at the date of death in calendar year $t$.[13]

By way of illustration, we have used the following model (that we label as M1-M5X) that models the general population using the Lee and Carter (1992) model (M1), and the spread between the general and specific population using a generalisation of the CBD model (M5X).[14] [15] Specifically:

$$\begin{aligned} \ln m^G(x,t) &= A(x)+B(x)K(t) && (12)\\ \ln m^P(x,t) &= \ln m^G(x,t)+\ln m^{adj}(x,t) && (13)\\ \text{where the spread } \ln m^{adj}(x,t) &= a(x)+k_1(t)+(x-\bar{x})k_2(t) && (14) \end{aligned}$$

and $\bar{x}=65$. Mortality rates are then derived using the relationship $q^G(x,t)=1-\exp(-m^G(x,t))$ and $q^P(x,t)=1-\exp(-m^P(x,t))$. These are then transposed into cohort tables that can be applied to the individual model points

Although use of the alternative Li and Lee (2005) model enjoys some prominence in some countries (see, for example, KAG, 2014; Antonio et al., 2015), the two-population model suggested here (by way of example) is used in preference, for the following reasons:

- M1-M5X contains ideas in the popular model of Li and Lee (2005) but has a more general and flexible term structure of correlation for mortality at different ages that is more in the spirit of Cairns et al. (2016). It also allows for different relative rates of improvement at different ages in the two populations.
- Certain calibrations of the Li and Lee model can result in implausibly high correlations between mortality rates in the two groups at different ages. If correlations between key building blocks are unrealistic, then analyses of hedge effectiveness (which depend heavily on correlation) cannot be considered to be reliable.

[13]Note that the indexation here to $x$ and $t$ contrasts with the earlier model-point oriented mortality rates such as $q_1^G(i,y)$ which are cohort specific.

[14]In live situations, hedgers should take responsibility for establishing what stochastic model is appropriate for their specific population as well as their chosen general population. As a recent example, Cairns et al. (2016) develop a more general multipopulation model for 10 distinct subgroups of the Danish males population, with the underlying data serving as a testbed for longevity hedge innovators.

[15]For a review of multi-population models for the assessment of basis risk, see Haberman et al. (2014).

The age and period effects are estimated using Poisson maximum likelihood in two steps. First, estimate the $A(x)$, $B(x)$ and $K(t)$ age and period effects in the Lee-Carter model for the general population. Then, given the $A(x)$, $B(x)$ and $K(t)$, estimate the $a(x)$, $k_1(t)$ and $k_2(t)$ age and period effects for the specific population.

We then model the period effects $K(t)$, $k_1(t)$ and $k_2(t)$ as time series processes: $K(t)$ is modelled as a random walk, while $k_1(t)$ and $k_2(t)$ are modelled as AR(1) time series processes that mean revert to 0. Random innovations for all three are correlated. Thus:

$$K(t) = K(t-1) + \nu + \sigma_0 Z_0(t) \tag{15}$$
$$k_1(t) = a_1 k_1(t-1) + \sigma_1 Z_1(t) \tag{16}$$
$$k_2(t) = a_2 k_1(t-1) + \sigma_2 Z_2(t), \tag{17}$$

where the vector of innovations, $Z(t) = (Z_0(t), Z_1(t), Z_2(t))^T$ is assumed to be multivariate standard normal with mean 0, individual variances of 1 and correlation matrix $\rho_Z$. In the numerical examples that follow, point estimates for the time series parameters $\nu$, $a_1$, $a_2$, $sigma_0$, $\sigma_1$, $\sigma_2$ and $\rho_Z$ have been used. However, allowance for parameter uncertainty could be included, although this would necessitate further adjustment to the proxy model to continue to ensure consistency between the stochastic approach and the standard formula stress test.

# 7 Hedge example and results

In this section we illustrate the application of the framework and two-population stochastic mortality model described in the previous sections. *All present values in the text that follows are quoted as present values at time 0 to ensure consistency with the Solvency II technical reserve and stress test.*

## 7.1 Data

For the general population we have used male national population mortality for the Netherlands (published by the Central Bureau of Statistics) from 1970 to 2013, ages 0 to 90, and for the specific population we have used Netherlands insurance data from 2002 to 2013, ages 40 to 89, made available by the CVS (Dutch Centre for insurance statistics).

## 7.2 Synthetic annuity portfolio

For clarity of exposition and to focus on application of the framework, we consider a portfolio of male lives holding single-life annuities (that is, with no spouse's benefit), with benefits payable annually at the end of each year, given survivorship to the end of that year.[16]

For each cohort considered, we have a level annuity payable annually in arrears from time 0, or from age 65 for cohorts aged less than 65. The weights attached to the cohorts, $w_x$, are as follows:

- $w_x = x - 25$ for $x = 40, \ldots, 49$
- $w_x = 25$ for $x = 50, \ldots, 64$
- $w_x = 90 - x$ for $x = 65, \ldots, 89$

These result in a weighted average age of 60.2.

## 7.3 Results

In the results that follow, we have used $10,000$ scenarios for an acceptable level of precision at the 99.5% level of probability, with a hedge maturity of $T = 10$ years.

Figure 2 shows various quantities for the deferred annuities and immediate annuities for single ages: the deterministic best estimate; the best estimate with the 20% stress; and the distribution of the present value at time 0 of $\tilde{L}(T)$ (fans showing a 90% confidence interval, and the 99.5% quantile). The volatility of the stochastic simulation model has been calibrated so that *for the portfolio described in Section 7.2* the 20% stress equates to the 99.5% quantile of the PV at time zero of the time-$T$ synthetic liability, $\tilde{L}(T)$. This effectively creates the proxy model for the Standard Formula approach. From Figure 2 (solid line for the stress and dashed line for the 99.5% VaR) we can observe the well known fact that the 20% stress test underestimates the required capital at young ages using a stochastic internal model and overestimates the capital requirement at high ages.

Figure 3 shows a scatterplot of the present value of the portfolio of annuities and deferred annuities versus the hedge payoff (PV(0) of $L(T)$ versus $\tilde{L}(T)$). For the given calibration of the model, the means of $\tilde{L}(T)$ and $L(T)$ discounted to time 0 are 10679 and 10689 respectively, with standard deviations of 350 and 395 respectively. The correlation between the two is 0.978. The imperfect correlation reflects population basis risk between times 0 and $T$, and in the consequent recalibration of the experience ratios at time $T$ that affects valuation of liabilities after $T$. The fact that the correlation is quite high reflects, primarily, the common dependence in both $\tilde{L}(T)$ and $L(T)$ on the recalibrated drift parameter $\nu$ (equation (15)). The higher standard deviation for $L(T)$ reflects the additional randomness arising from the recalibrated experience ratios at time $T$.

[16] In practice, benefits will be payable on a more regular basis and, therefore, more evenly spread through the year. The contractual structure of the hedge underlying, $\tilde{L}(T)$, reflects this simple annual payment assumption.

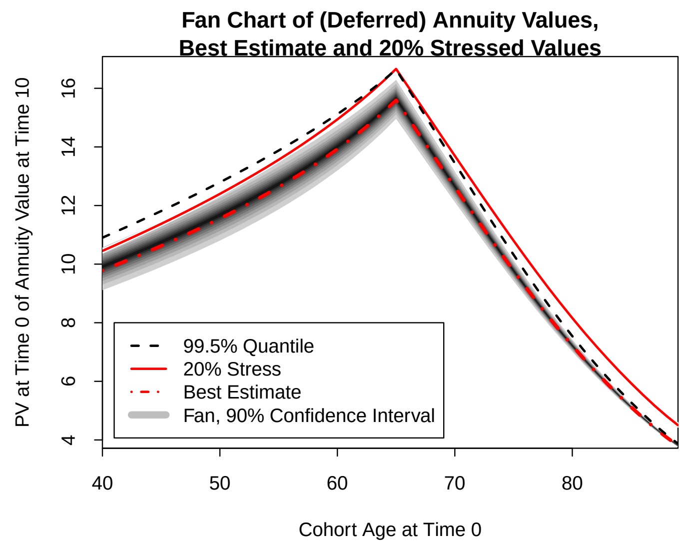
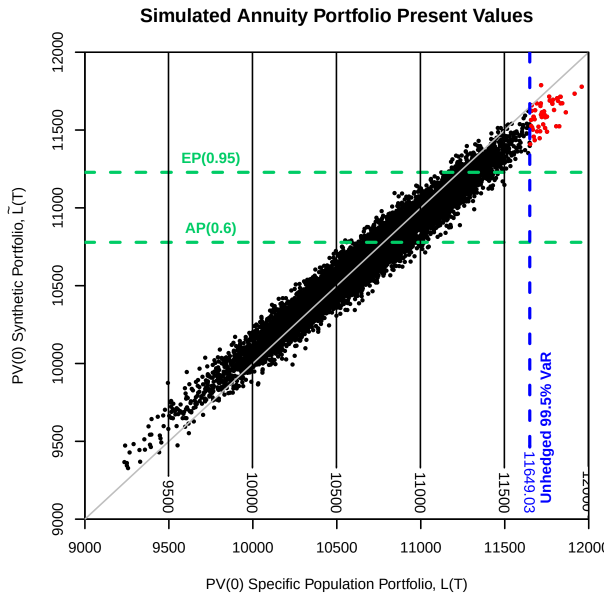



Figure 2: Annuity values (deferred annuities below age 65) for individual ages. Fan and dashed lines show uncertainty in the synthetic liability, $\tilde{L}(T)$, discounted to time 0, for $T = 10$ years. The (red) dot-dashed line shows the best estimate of the liability (using general population mortality forecasts multiplied by the experience ratios, $\varepsilon^0(x)$). The solid (red) line shows equivalent values under the 20% stress test.

| | $T = 10$ | $T = 15$ | $T = 20$ |
|---|---|---|---|
| $\sigma_G$ (*) | 4.775 | 4.587 | 4.739 |
| $L(0)$ | | 10626 | |
| $SCR_{20\%stress}$ (*) | | 840 | |
| $E[\tilde{L}(T)]$ | 10679 | 10687 | 10674 |
| $E[L(T)]$ | 10689 | 10703 | 10697 |
| $SD[\tilde{L}(T)]$ | 350 | 354 | 361 |
| $SD[L(T)]$ | 395 | 396 | 396 |
| $AP(0.6)$ | 10779 | 10788 | 10778 |
| $EP(0.95)$ | 11228 | 11247 | 10697 |
| $SCR_{10}$ (*) | 840 | 840 | 840 |
| $SCR_{11}$ | 478 | 470 | 461 |
| $SCR_{20}$ | 960 | 970 | 923 |
| $SCR_{21}$ | 598 | 600 | 544 |

Table 1: Key statistics for the liabilities $L(T)$ and $\tilde{L}(T)$ for maturities $T = 10, 15, 20$. $L(0)$ represents the time 0 liability. $AP(0.6)$ and $EP(0.95)$: 60% and 95% quantiles for the attachment and exhaustion points for each $T$. $SCR_{10}$ =SCR for $\tilde{L}(T)$ with no hedging. $SCR_{11}$ =SCR for $\tilde{L}(T) - H(T)$. $SCR_{20}$ =SCR for $L(T)$. $SCR_{21}$ =SCR for $L(T) - H(T)$. (*) $\sigma_G$ calibrated to ensure that $SCR_{10} =$ equals the SCR under the 20% stress test.



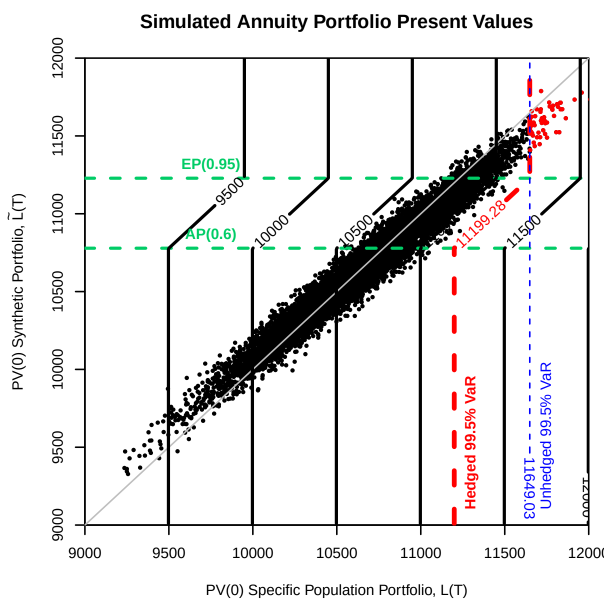




Figure 3: Scatterplot of the simulated portfolio present values, $L(T)$ (discounted to time 0), for the specific population (population 2) versus the synthetic population, $\tilde{L}(T)$ (population 1 mortality multiplied by the experience ratios). Attachment and exhaustion points at the 60% and 95% quantiles of $\tilde{L}(T)$ respectively (green horizontal dashed lines). Top: solid vertical lines show contours of the unhedged portfolio value and (blue dashed line) the 99.5% VaR threshold; red dots highlight the 0.5% highest liabilities above the 99.5% VaR. Bottom: solid kinked lines show contours of the hedged portfolio value and (red dashed line) the 99.5% VaR threshold; red dots highlight the 0.5% highest liabilities above the 99.5% VaR.

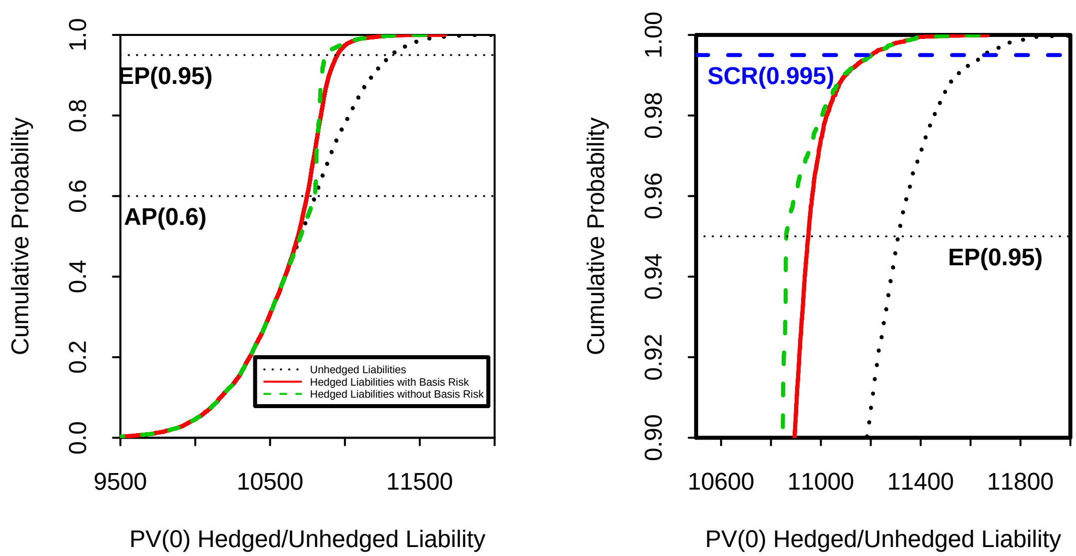



Figure 4: Cumulative distribution of the unhedged ($L(T)$) and hedged liabilities ($L(T) = H(T)$) over the full range of probabilities (left) and the top 10% (right). Distributions for the hedged liability are shown for the cases with (red solid line) and without (green dashed line) population basis risk.

The best estimate and 20% stressed values for the unhedged portfolio are 10626 and 11466 resulting in an SCR of 840 or 7.9% ($SCR_{20\%stress}$ in Table 1).[17] The corresponding difference between the 99.5% quantile of $\tilde{L}(T)$ and its mean is also 840 ($SCR_{10}$ in Table 1) due to the recalibration of the simulation model volatility (the proxy model). Using the proxy model with population basis risk the SCR increases to 960 ($T = 10$; $SCR_{20}$ in Table 1) reflecting the higher standard deviation.

We now introduce the hedge.

As one example, suppose that the attachment and exhaustion points are set at the 60% and 95% quantiles of the hedge underlying, $\tilde{L}(T)$: 10779 and 11228 respectively (see Table 1); a difference of 449.5 (which we set as the notional $u$).

Under this scenario the SCR falls from $SCR_{20} = 960$ to $SCR_{21} = 598$ (Table 1; see Figures 3 and 6 for the corresponding VaR figures). In the alternative scenario with no population basis risk, the SCR changes from $SCR_{10} = 840$ to $SCR_{11} = 478$ (Table 1).[18]

[17] The best estimate differs from the mean of $\tilde{L}(T)$ for two reasons. First, the distribution of $\tilde{L}(T)$ might not be symmetrical. Second, and more importantly, the sliding window for recalibration at time $T$ results in a distribution for the central trend after time $T$ ($\nu$ is equation (15)) that is not centred around the trend estimated at time 0.

[18] Note that $VaR(\tilde{L}(T) - H(T), 0.995)$ and $VaR(L(T) - H(T), 0.995)$ are, therefore, both $EP - AP$ lower than their unhedged counterparts. From the hedger's perspective, this represents the true release of capital. The reductions in the corresponding SCR's are, in contrast, lower as there is a matching drop of $E[H(T)] = 87$ in the benchmark mean.

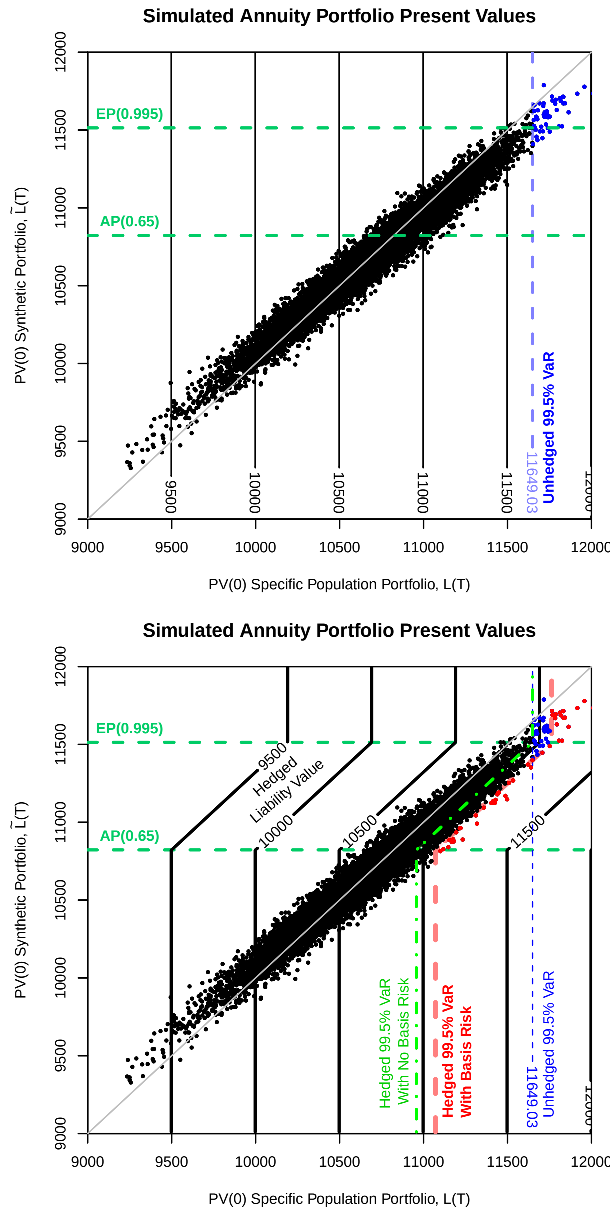



Figure 5: Scatterplot of the simulated portfolio present values, $\tilde{L}(T)$ (discounted to time 0), for the specific population (population 2) versus the synthetic population, $L(T)$ (population 1 mortality multiplied by the experience ratios). Attachment and exhaustion points at the 65% and 99.5% quantiles of $\tilde{L}(T)$ respectively (green horizontal dashed lines). Top: solid vertical lines show contours of the unhedged portfolio value and (blue dashed line) the 99.5% VaR threshold; blue dots highlight the 0.5% highest liabilities above the 99.5% VaR. Bottom: solid kinked lines show contours of the hedged portfolio value and (red dashed line) the 99.5% VaR threshold; red dots highlight the 0.5% highest liabilities above the 99.5% VaR. Blue dots and dashed line remain from the top plot for comparison. Green dot-dashed line: 99.5% VaR threshold contour if there was no basis risk.

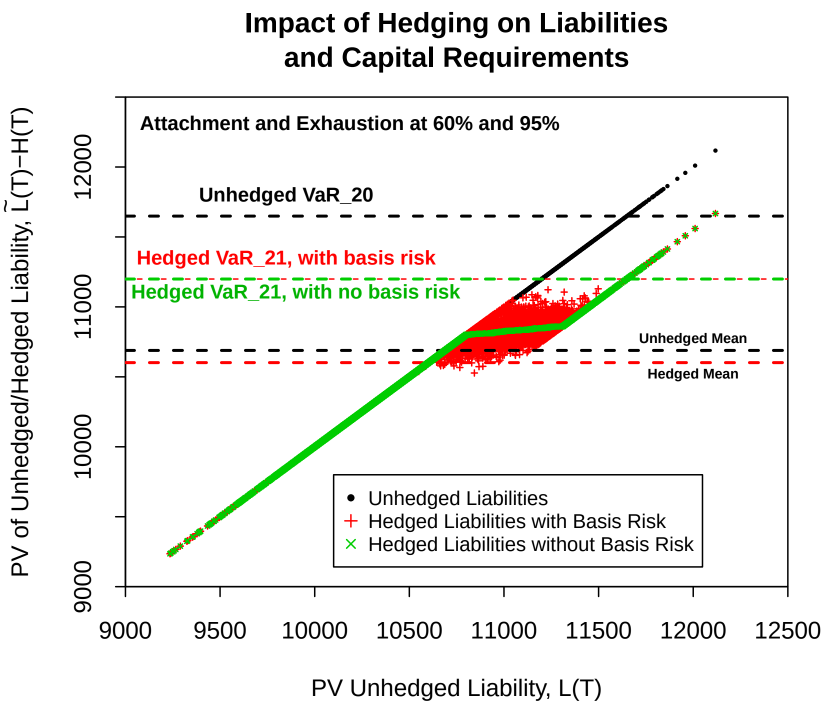



Figure 6: Scatterplot of simulated liabilities before and after hedging when the attachment and exhaustion points are at the 60% and 95% quantiles of the underlying. Horizontal lines: best estimate liability, and additional SCRs for the hedged (with and without basis risk) and unhedged liabilities.

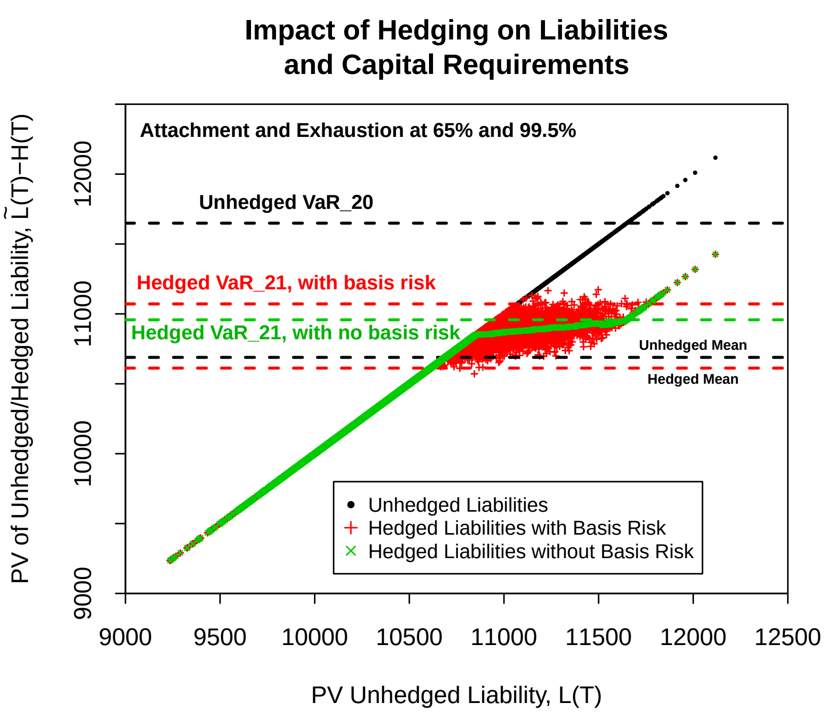



Figure 7: As Figure 6 except the attachment and exhaustion points are at the 65% and 99.5% quantiles.

The haircut is defined as
$$HC = 1 - \frac{SCR_{20} - SCR_{21}}{SCR_{10} - SCR_{11}}$$
which measures how much the capital relief should be reduced as a result of the inclusion of population basis risk. With no population basis risk the capital relief would be $SCR_{10} - SCR_{11}$ compared to $SCR_{20} - SCR_{21}$ with population basis risk. In this numerical example, the haircut turns out to be zero for reasons that we discuss below.

Figure 3 also contains contours for the portfolio value without (top) and with (bottom) the hedge in place. In both cases the 50 worst outcomes out of the 10,000 scenarios are highlighted in red. In the lower plot, the contours have two kinks which correspond to the attachment and exhaustion points (AP and EP). For this particular choice of EP and AP, the kinked section of the 99.5% contour (red dashed line) lies below the cloud of 10,000 dots, and so the 50 worst outcomes are the same as in the unhedged simulations. In these 50 worst scenarios, $H(T) = EP - AP$ always pays out in full, reflecting the high correlation between $L(T)$and $\tilde{L}(T)$.

In Figure 4 shows the cumulative distributions for the unhedged and hedged liabilities, with the right hand plot zoomed in on the top 10% to allow us to see better the 99.5% quantile. For the hedge without population basis risk, the $N = 10,000$ values of $L(T)$ and $\tilde{L}(T)$ were ordered and paired off to remove the impact of population basis risk, but retain the different marginal distributions. Without population basis risk, therefore, we see that the green dashed curve coincides with the dotted black curve for the unhedged liability up to the 0.6 level of probability. And above 0.95, there is a constant horizontal shift from the green dashed curve to the dotted black curve equal to $EP - AP$. Now compare the hedged distributions with and without population basis risk. Visible differences occur only between the 45% and 99% probability levels (approximately). This reiterates the earlier observation that, at the 99.5% level of probability the hedging instrument will pay off in full when EP is set at the 95% quantile of $\tilde{L}(T)$.

Figure 6 illustrates these points in a slightly different way, showing how the hedge has an impact on the present value for this specific combination of attachment and exhaustion points using a different type of scatterplot. The red crosses (+'s) show individual outcomes for the unhedged versus the hedged liability. The left-hand straight boundary correponds to outcomes where $H(T) = 0$, and the right-hand straight boundary corresponds to outcomes where $H(T) = EP - AP$. The more fuzzy region in the middle corresponds to outcomes where $0 < H(T) < EP - AP$.

For the 0.5% of worst outcomes where $L(T) > 11649$, $H(T)$ always equals its maximum value, so that the empirical 99.5% quantile of $L(T) - H(T)$ is simply the empirical 99.5% quantile of $L(T)$ minus $(EP - AP)$. For similar reasons, the empirical 99.5% quantile of $\tilde{L}(T) - H(T)$ is also the empirical 99.5% quantile of $\tilde{L}(T)$ minus $(EP - AP)$. So the haircut is zero.[19]

[19]Figure 6 also shows what the outcomes would look like in the absence of population basis risk (green x's). In this case, population basis risk is removed by ordering the simulated $L(T)$and $\tilde{L}(T)$, and then pairing

The haircut will typically be greater than zero if, in the context of Figure 6, we were to observe some red crosses in the fuzzy middle region that fell above the mean $+ SCR_{21}$ dashed red line (see, for example, Figure 7).

This observation is critically dependent on the value of the exhaustion point and the correlation between $L(T)$ and $\tilde{L}(T)$. For example, Figures 5 and 7 show corresponding results when the quantiles of the attachment and exhaustion points increased to 65% and 99.5% respectively. In the lower plot in Figure 5 we see that the 50 worst scenarios with hedging (red dots) are now different from the worst 50 without hedging (blue dots). The thin red dashed line also shows what the 99.5% VaR threshold would be if there was no population basis risk ($cor(L(T), \tilde{L}(T)) = 1$), and we can see that this would be significantly lower. This translates into a significant haircut. Figure 7 shows graphically the impact on the SCR calculations. The high exhaustion point means that a proportion of the cloud of red crosses now crosses over the 99.5% VaR, pushing up this quantity relative to the situation where there is no population basis risk (green crosses).[20]

In Figure 8 we show how the haircut depends on the underlying quantiles for the attachment and exhaustion points. In line with our discussion above, we observe that the size of the haircut depends mainly on the exhaustion point and very little on the value of the attachment point. More importantly, the haircut is empirically zero or close to zero for exhaustion points up to about the 97% quantile of $\tilde{L}(T)$ (for this level of correlation, and VaR confidence level). Above this, the haircut increases, reflecting the discussion above.

In Figure 9 we show how $SCR_{21} = VaR(L(T) - H(T), 0.995) - E[L(T) - H(T)]$ (expressed as a percentage of the unhedged $SCR_{20}$) depends on the underlying quantiles for the attachment and exhaustion points. For higher values of the exhaustion point, a haircut has been applied to account for non-negligible impact of population basis risk.

## 7.4 Sensitivity of results to the hedge maturity

The main analysis above has focused on a 10-year-maturity hedge, with some discussion of how sensitive results are to the choice of attachment and exhaustion points.

We have also investigated how sensitive results are to the maturity of the hedge.

As an example, we looked at a $T = 20$ year hedge (Table 1; also $T = 15$). Results are very similar to those for $T = 10$:

- The correlation between $L(T)$ and $\tilde{L}(T)$ is very similar (about 0.980).
- If we use the same volatility calibration as the $T = 10$ year maturity (figures not quoted in Table 1), the standard deviations of $L(T)$ and $\tilde{L}(T)$ are only very slightly

off the ordered samples so that $L(T)$and $\tilde{L}(T)$are comonotonic.

[20]Note that in the middle kinked section of the line of green crosses is sloping slightly upwards. This is the result of $\tilde{L}(T)$ having a higher variance than $L(T)$.

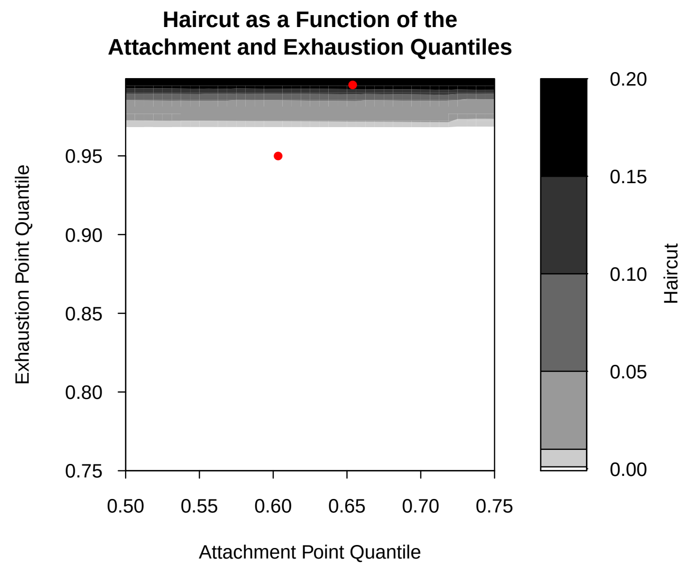



Figure 8: Shaded contour plot (heat map) showing how the percentage haircut depends on the probabilities that the hedging instrument underlying falls short of the attachment and exhaustion points. in the lower, white region all values are very close to zero. Red dots show the AP/EP = 60/95% and 65/99.5% examples.

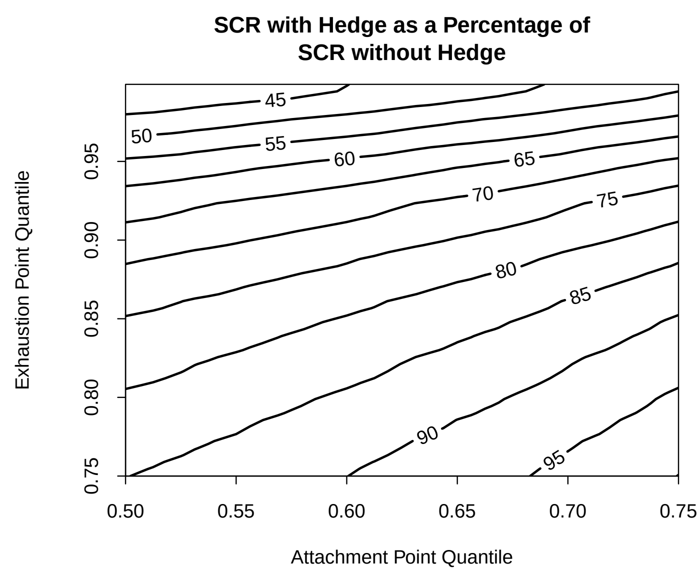



Figure 9: Contour plot showing how the SCR, with the hedge in place, with basis risk included and accounted for, depends on the probabilities that the hedging instrument underlying falls short of the attachment and exhaustion points. The SCR is expressed as a percentage of the original SCR with no hedge in place.

higher. Further investigation shows that in individual stochastic scenarios, a high proportion of the uncertainty in the ultimate PV of the liability $L$ crystallises in the first 10 years.

- With a modest further adjustment to the model volatility (to match the 20% stress) the $T = 20$ equivalents to Figures 3 to 9 look very similar as do the numerical results in Table 1.
- Figure 6 changes slightly more, with more uncertainty in the younger age deferred annuity present values.

# 8 Involvement of the capital markets

So far we have focused on the hedge itself and its impact on the hedger's capital requirement. Often the issuer of the call spread option will wish to pass the risk on to the capital markets. This might take the form of a catastrophe bond, possibly issued through a special purpose vehicle. A possible cashflow profile for this would be as follows:

- Bondholders pay the initial bond price of $B(0)$ at time 0.
- The bond pays a floating rate of LIBOR+$x$% per 100 notional (e.g. quarterly or half-yearly), which is guaranteed, and not linked longevity risk.
- The repayment of the bond notional is at time $T$. Payment is in full (100% of notional) if the hedge underlying, $\tilde{L}(T)$, falls below the attachment point $AP$. Above that the repayment of principal is reduced. More specifically, the final repayment of principal is

$$\begin{array}{ll} N & \text{if } \tilde{L}(T) < AP \\ N - \left(\tilde{L}(T) - AP\right) & \text{if } AP \leq \tilde{L}(T) < EP \\ N - (EP - AP) & \text{if } EP \leq \tilde{L}(T). \end{array}$$

If $N = EP - AP$ then the bond recovery rate will be zero if $\tilde{L}(T) \geq EP$. The bond spread of $x$% would depend on the market's perception of the risk of a reduction in the bond principal, and the market price of this risk, as well as an illiquidity premium.

In general, therefore, a liability of $L(0)$ can be hedged through the issuance of a catastrophe bond with a notional of $N$. In the numerical example with the 60% and 95% attachment and exhaustion points, we would have $N = EP - AP = 449.5$ versus $L(0) = 10626$: i.e. bond notional that is 4.2% of the liability.

# 9 Conclusions

This paper has focused on the use of index-based financial contracts as a means of hedging an insurer's exposure to longevity risk. Index-based longevity hedges give rise to basis risk of various types including population basis risk due to the differing nature of the reference population and the hedger's population in the short and long term.

We have focused on the development of a rigorous approach to allow clarity in the assessment of the impact of population basis risk. The proposed approach breaks down the calculations into a sequence of 22 steps that can be used as a template for clear documentation for both internal use and for discussions with regulators. Additionally, the template should provide academic researchers with a clearer view of how longevity risk is assessed in practice in contrast with academic ideals. In doing so, we hope that the paper will facilitate further dialogue between academics and practitioners by breaking down perceived technical and language barriers.

Using a call spread option contract as an important example, we have investigated the impact of such a contract on an insurer's capital requirements within a Solvency II setting, using a haircut measure to express the loss of hedge effectiveness due to population basis risk. Detailed analysis found that the haircut was quite sensitive to the choice of exhaustion point of the option. In the chosen example, an exhaustion point close to the Solvency II 99.5% quantile could give rise to relatively large haircuts. For lower exhaustion points the haircut was found to be very close to zero. In the latter cases, therefore, population basis risk has a negligible impact. In these cases, structural basis risk predominates (i.e. the non-linear payoff profile of the option). But if structural basis risk predominates, then, arguably, there is no difference, from a Solvency II perspective, between an index-based hedge as described here and a formal reinsurance contract.

# Appendices

## A Detailed commentary on Figure 1

1. A,B: collect the historical exposures and deaths data for populations G and P. Calculate historical death rates.

2. $ER : \varepsilon^{H0}$: Convert death rates to mortality rates.

   Calculate single age experience ratios using the most recent 10 years.

   Smooth the exposure ratios across ages (e.g. smooth using linear regression, a kernel smoother or splines). Extrapolate smoothed experience ratios to high and low ages using a clearly documented method.

   The age specific experience ratios, $\varepsilon^{H0}(x)$ become embedded in the hedge instrument contract at time 0. These values might also be used in the calculation of the Solvency II technical reserve.

3. Identify a suitable two-population stochastic mortality model and fit this to the historical deaths and exposures to provide estimates of the underlying age, period and cohort effects. Here:

   A: $LC(0)$: Fit the Lee Carter model (M1) to the most recent 44 years of data from population G up to time 0.

   B: $M1M5X(0)$: Fit the M1M5X model to Population P data. Use the most recent 12 years of population P data up to time 0. Use the output from $LC(0)$ as input to this step.

4. Identify a suitable time series model for forecasting period (and cohort) effects for the two populations. Here:

   A: Random walk with drift for the population 1 period effect in M1.

   B: AR(1) time series processes for the population 2 period effects.

5. Estimate the time series process parameters for the models in step 4. Here:

   A: Estimate its process parameters using the 44 years of data output from $LC(0)$. Parameter estimates are represented by $\hat{\mu}(0)$.

   B: Estimate process parameters using the most recent 12 years of data output from $M1M5X(0)$. Parameter estimates are represented by $\hat{\mu}(0)$.

   $\hat{\mu}(0)+$: Adjust parameter estimates: for example

   - Use the population G drift estimated in Stage 4A (consistent with local practice).
   - Scale the population G volatility for $K(t)$.

- Scale the population P volatilities for $k_1(t)$ and $k_2(t)$.

Volatilities might be recalibrated to ensure that the stochastic 99.5% capital charge matches the 20% stress test capital charge, that is to create a proxy distribution for the standard formula (20% corresponds to the standard formula shock under Solvency II).

Time series parameter estimation and any subsequent adjustments must be explicitly documented.

6. $SIM(0)$: Use the simulation model to generate a stochastic scenario from time 0 up to the hedge instrument payoff date at time $T$.

7. A, B: Store simulation model output: underlying death rates for populations G and P for years $1,\ldots,T$.

   C, D: (Optional) Input exposures for the relevant age range and years $1,\ldots,T$.

   Use simulated underlying death rates plus exposures to simulate death counts using the Poisson assumption. (Poisson risk can be switched off, and deaths set equal to the expected value.)

   Calculate crude death rates for years $1,\ldots,T$.

8. $LC(T)$: Recalibrate the LC model (M1) to the most recent 44 years of population G deaths and exposures data up to time $T$.

   (Here we use a rolling 44-year window consistent with local practice.)

   Note that different methods of recalibration might be used for the hedge instrument and for the valuation of the liabilities.

9. $\hat{\mu}(T)$: recalibrate the time series process parameters of the LC model.[21] Different methods of recalibration might result in different estimates.

10. $SIM(T)$: Specify fully and run the simulation model for the single population G.

    Here, we assume a single scenario, using the median projection.

    Different methods of recalibration might result in different projections, which we label as $\bar{q}^{LG}(x,t)$ for the liabilities and $\bar{q}^{HG}(x,t)$ for the hedge instrument.

11. Store simulation model output: underlying death rates for population G for years $T+1, T+2, \ldots$.

12. This step and the following step 13: specify the liability valuation model; calibrate this model using populations G and P mortality; use this model to project population P mortality given the experience up to time $T$. Here:

[21] Comment: The results of steps 8 and 9 will be used to calculate the hedge instrument payoff, and calculate liability values at $T$.

$ER : \varepsilon^T$: Calculate single age exposure ratios, $\varepsilon^T(x)$, using the most recent 10 years of mortality rates up to time $T$. Smooth the exposure ratios across ages.

13. A,B: Use the outputs from Stages 11 and 12 to calculate projected mortality rates for population P at the core ages.
14. A, B: Extrapolate the population G mortality curves beyond age 90 in each of years $1,\dots,T;T+1,T+2,\dots$.

    Here, we use the Kannisto method.[22]
15. Extrapolate population P mortality beyond age 89 in each of years $1,\dots,T$.

    Here, we use the Kannisto method.
16. Extrapolate population P mortality beyond age 89 in each of years $T+1,T+2,\dots$.

    Here, we use the Kannisto method.
17. Extrapolate population P mortality below age 40 in each of years $1,\dots,T$.

    Here: extrapolate $a(x) = a(40)$ for $x < 40$ in the M1M5X model. Then, this stage becomes part of stages 4, 6 and 7B.[23]
18. Extrapolate population P mortality below age 40 in each of years $T+1,T+2,\dots$.

    Here: extrapolate $a(x) = a(40)$ for $x < 40$ in the M1M5X. Then, this stage becomes part of stages 4, 6 and 7B.
19. Extract the death rates $m^G(x,t)$ for $x = 0,\dots,89$ and $t = 1,\dots,T$ and extrapolated death rates for higher ages and later years: that is, the red dashed box in Figure 1.

    Calculate the population G mortality rates, $q^G(x,t) = 1-\exp(-m^G(x,t))$ for $x = 0,\dots,120$ and $t = 1,2,\dots$.

    Use the experience ratios calculated in step 2, to calculate the synthetic mortality rates $q^G(x,t)\varepsilon^{H0}(x)$ for $x = 0,\dots,120$ and $t = 1,2,\dots,T$, and $\bar{q}^{HG}(x,t)\varepsilon^{H0}(x)$ for $x = 0,\dots,120$ and $t = T+1,\dots$.
20. Use the synthetic mortality rates in stage 19 to calculate synthetic cohort survival rates, and the synthetic portfolio present value at time $T$, $\tilde{L}(T)$.
21. Use $\tilde{L}(T)$ to calculate the hedge instrument payoff, $H(T)$.
22. Extract the population P death rates $m^P(x,t)$ for $x = 0/40,\dots,89$ and $t = 1,\dots,T$ and extrapolated death rates for higher/lower ages and later years: that is, the blue dashed box in Figure 1.

---

[22] The Kannisto method carries out a linear regression on the logit $m^G(x,t)$ over ages 80 to 90, and then extrapolates the fitted line to higher ages.

[23] This step is only relevant if death rates below age 40 are required in the liability valuations.

Use these rates to calculate cohort survival rates, and the portfolio present value at time $T$, $L(T)$.

23. Repeat steps 6 to 22 $N$ times ($N = 10,000$, say).

# B Required models

The process of generating $L(T)$ and $\tilde{L}(T)$ requires several models, each of which must be clearly articulated.

- The time zero liability valuation model to calculate the Solvency II technical reserve, $L(0)$.
- The two-population stochastic mortality model for the core ages.

  This model is used to simulate death rates in years 1 to $T$.
- The hedge instrument valuation model.

  This model and its method of recalibration at time $T$ is specified in the contract at time 0. It is used to project the general population death rates beyond time $T$.
- The liability valuation model.

  This model is used to project the specific population death rates beyond time $T$.

  The model used and the method of recalibration at $T$ should reflect current local practice and potential improvements.
- A model, or numerical method, for extrapolating death rates to higher and lower ages.

# C Capital relief in future years

The main paper has focused on the calculation of capital relief at the time the hedge is put in place. It is also necessary to assess how much capital relief there should be between time 0 and time $T$. The approach that we recommend is as follows:

- the time horizon for the stochastic analysis should remain at the maturity of the hedge;
- at each valuation date, $t$, the volatility of the stochastic mortality model should be recalibrated so that the 99.5% capital charge under the $T - t$ year stochastic simulation matches the 20% stress also carried out at time $t$ (and in both cases reflecting the data up to $t$).

Over time, the "moneyness" of the hedge will change, so that the capital relief with and without basis risk might change, and the haircut might rise significantly above 0. For example, if at time $t$, the exhaustion point was now close to the 99% quantile (so the hedge had moved further out of the money) then the haircut might be around 10% to 15% (see Figure 8). Concurrently, the capital relief is amortized over time (up to zero capital relief at maturity), the absolute impact of a potential haircut over time is therefore reduced due to the amortization of the capital relief itself.

# Acknowledgements

The authors would like to thank Philippe Coulomb for his input and support during the process of writing this paper, and Raymond Waucomont, Wilbert Ouburg and Alan Rae for their helpful comments in the early stages of the work. All opinions expressed in the paper are, however, those of the authors and not those of their employers or of the contributors above. AC acknowledges financial support from the Actuarial Research Centre of the Institute and Faculty of Actuaries.

# References

Antonio, K., Devolder, L., and Devriendt, S. (2015) The IABE 2015 mortality projection for the Belgian population.

Blake, D., Cairns, A., Coughlan, G., Dowd, D., MacMinn, R. (2013) The new life market. *Journal of Risk and Insurance*, 80: 501558.

Blake, D., Cairns, A.J.G., and Dowd, K. (2006) Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12: 153-197.

Cairns, A.J.G. (2013) Robust hedging of longevity risk. *Journal of Risk and Insurance*, 80: 621-648.

Cairns, A.J.G., Blake, D., and Dowd, K. (2008) Modelling and management of mortality risk: A review. *Scandinavian Actuarial Journal*, 2008: 79-113.

Cairns, A.J.G., Blake, D., Dowd, K., and Coughlan, G.D., (2014) Longevity Hedge Effectiveness: A Decomposition. Quantitative Finance, 14: 217-235.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., and Khalaf-Allah, M. (2011) Bayesian stochastic mortality modelling for two populations, *ASTIN Bulletin*, 41: 29-59.

Cairns, A.J.G., Kallestrup-Lamb, M., Rosenskjold, C.P.T., Blake, D., and Dowd, K., (2016) Modelling socio-economic differences in the mortality of Danish males using a new affluence index. Working paper, Heriot-Watt University. www.macs.hw.ac.uk/~andrewc/papers/ajgc72.pdf (accessed 9/8/2016).

Coughlan, G., Epstein, D., Sinha, A., and Honig, P. (2007) *q*-forwards: Derivatives for transferring longevity and mortality risk. Available at www.researchgate.net/profile/Guy_Coughlan (accessed 8/8/2016).

de Vries, A. (2015) Introduction to capital motivated longevity solutions. Presentation at 11th International Longevity Risk and Capital Markets Solutions Conference, Lyons, September 2015. www.cass.city.ac.uk/longevity-11/programme (accessed 8/8/2016).

Haberman, S., Kaishev, V., Millossovich, P., Villegas, A., Baxter, S., Gaches, A., Gunnlaugsson, and Sison, M. (2014). *Longevity Basis Risk: A Methodology for Assessing Basis Risk*. Sessional research paper, Institute and Faculty of Actuaries, 8 December 2014. Tech. rep., Cass Business School, City University London and Hymans Robertson LLP.

KAG (2014) Prognosetafel AG2014. Koninklijk Actuarieel Genootschap (Dutch Royal Actuarial Association). Available at http://www.ag-ai.nl/view.php?Pagina_Id=535 (accessed 8/8/2016).

Kessler, A., McCloskey, W.M., and Bensousson, A. (2015) The pension risk transfer market at $240 billion: innovation, globalization, and growth. In *Pension and Longevity Risk Transfer for Institutional Investors, Fall 2015.* Editor B.R. Bruce. Institutional Investor Journals. pp 18-27.

Lee, R.D., and Carter, L.R. (1992) Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87: 659-675.

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

Michaelson, A., and Mulholland, J. (2015) Strategy for increasing the global capacity for longevity risk transfer. In *Pension and Longevity Risk Transfer for Institutional Investors, Fall 2015.* Editor B.R. Bruce. Institutional Investor Journals. pp 28-37.