Delta-Hedging Longevity Risk under the M7-M5 Model: The Impact of Cohort Effect Uncertainty and Population Basis Risk

Kenneth Q. Zhou and Johnny Siu-Hang Li

Abstract: In a recent project commissioned by the Institute and Faculty of Actuaries and the Life and Longevity Markets Association, a two-population mortality model called the M7-M5 model is developed and recommended as an industry standard for the assessment of population basis risk. In this paper, we contribute a delta hedging strategy for use with the M7-M5 model, taking into account of not only period effect uncertainty but also cohort effect uncertainty and population basis risk. To enhance practicality, the hedging strategy is formulated in both static and dynamic settings, and its effectiveness can be evaluated in terms of either variance or 1-year ahead Value-at-Risk (the latter is highly relevant to solvency capital requirements). Three real data illustrations are constructed to demonstrate (1) the impact of population basis risk and cohort effect uncertainty on hedge effectiveness, (2) the benefit of dynamically adjusting a delta longevity hedge, and (3) the relationship between risk premium and hedge effectiveness.

Keywords: Cohort effect; Delta hedging; Population basis risk; The M7-M5 model; Value-at-Risk

1 Introduction

The Life and Longevity Markets Association (LLMA) is a non-profit organization funded by AVIVA, AXA, Deutsche Bank, J.P. Morgan, Morgan Stanley, Prudential PLC and Swiss Re to advance the development of a liquid longevity market. In 2012, the LLMA acquired ownership of the LifeMetrics Index (originally created by J.P. Morgan), on which standardized mortality-linked derivatives such as q-forwards can be written. The LLMA has also invested heavily on researching on population basis risk, the risk which arises from the fact that future mortality improvements of two different populations (the hedger’s portfolio of individuals and the reference population of the hedging instrument) are unlikely to be identical. It is believed by many that the lack of understanding of population basis risk is a major obstacle to market development.

In 2013, the LLMA and the Institute and Faculty of Actuaries (IFoA) jointly commissioned a project with an objective to develop a well-established methodology for assessing population basis risk. Phase I of the project was undertaken by a team of researchers from Cass Business School and practitioners from Hymans Robertson, who performed a systematic assessment of over 20 existing mortality models, aiming to identify the most suitable two-population mortality model for measuring population basis risk. Through a ‘best of breed’ selection process, the project team recommended the M7-M5 model, which can be regarded as a two-population extension of the Cairns-Blake-Dowd (CBD) family of models (Cairns et al., 2006, 2009). The selection process and estimation results can be found in the project report (Haberman et al., 2014).

Although the project team’s recommendation does not preclude the consideration of alternative models in future work, the M7-M5 model is likely to be regarded by market participants as an industry standard for assessing population basis risk. The M7-M5 model captures the most important drivers of the mortality dynamics of two related populations, including (1) the period (time-related) effect that applies to both populations, (2) the cohort (year-of-birth-related) effect that applies to both populations, (3) the period effect that applies to the mortality differential between the two populations, and (4) the interaction between age and period effects.

Users of index-based longevity hedges are challenged by the question of how to best use a collection of mortality-linked derivatives. In this paper, we attempt to seek an answer to this question, on the basis of the assumption that the true underlying mortality dynamics of the populations involved follow the M7-M5 model. Given the expected popularity of the M7-M5 model, the research problem we consider is important and practically relevant, but to our knowledge it has not been investigated seriously. Some related work has been performed by Villegas et al. (2017), who estimated hedge effectiveness under the M7-M5 model...
assumption using a hedge ratio of $h^* = \text{cov}(L, H)/\text{var}(H)$, where $L$ and $H$ represent the random present values of the unhedged liability and hedging instrument, respectively. It can be shown that the hedge ratio of $h = h^*$ minimizes the variance of the hedged position $L - hH$. However, this simple one-instrument hedging strategy is far from being adequate, and is not even valid when the hedger’s objective is not minimizing variance. It is also unclear as to how this strategy can be applied in a dynamic-setting with manageable computational efforts.

We propose to take a sensitivity-matching approach, in which the optimal hedging strategy is formulated by matching the longevity Greeks (derived specifically under the M7-M5 model assumption) of the liability being hedged and the portfolio of hedging instruments. The longevity Greeks, defined as partial derivatives with respect to the key model parameters, can be seen as a functional equivalent of the option Greeks that are used extensively in managing equity risk, so that practitioners should find them easy to interpret and understand. The approach we take is highly dissimilar from risk-minimization approaches (see, e.g., Coughlan et al., 2011; Cairns et al., 2014; Dahl and Møller, 2006; Dahl et al., 2008, 2011; Liu and Li, 2016a; Ngai and Sherris, 2011; Wong et al., 2014), in which one particular risk metric such as variance is minimized (but other risk metrics may have to be compromised). It is also different from duration-matching approaches (see, e.g., Tsai et al., 2010; Li and Hardy, 2011; Li and Luo, 2012; Lin and Tsai, 2013, 2014; Tsai and Jiang, 2011; Tsai and Chung, 2013), in which the sensitivities to the mortality rates themselves (rather than those to the parameters in the assumed model) are being matched.

Longevity Greeks have been studied previously by Luciano et al. (2012), Luciano and Regis (2014), Luciano et al. (2015) and De Rosa et al. (2017) in continuous-time settings and Cairns (2011), Cairns (2013), Liu and Li (2016b) and Zhou and Li (2017, 2018) in discrete-time settings. We choose to draw on the recent contributions of Cairns (2013) and Zhou and Li (2017), because they fit the fact that the M7-M5 model is defined in discrete-time and our ambition to develop not only static but also dynamic hedging strategies. As in the original work of Cairns (2013) and Zhou and Li (2017), we use q-forwards as hedging instruments and assume that the liability being hedged is a life annuity. However, to enhance flexibility, we permit the annuity liability to have a non-zero deferment period. Furthermore, unlike Cairns (2013) and Zhou and Li (2017) who focused on period effects only, we consider both period and cohort effects, incorporating the circumstances when the q-forwards and/or the annuity liability are subject to cohort effect uncertainty. Of course, we allow the annuity liability and q-forwards to be linked to different populations, leveraging the ability of the M7-M5 model to capture population basis risk. It should be emphasized, however, that we do not consider other sources of basis risk including sampling basis risk (which arises from the finite sample variation surrounding a fixed mortality trajectory) and structural basis risk (which arises from the differences in the payoff structures of the liability being hedged and the hedging instruments). These additional sources of basis risk have been studied by Li et al. (2018) and Li and Hardy (2011).

We first introduce a static hedging strategy by deriving the semi-analytical expressions for the longevity deltas (i.e., the first partial derivatives with respect to the most recently realized period and cohort effects) of the annuity liability and q-forwards. The expressions are semi-analytical in the sense that a large part of the calculations are accomplished using the structural and statistical properties of the assumed stochastic processes for the period and cohort effects.

We then extend the hedging strategy to a dynamic setting, in which the hedger is allowed to rebalance the portfolio of q-forwards periodically. To this end, we utilize the ‘approximation of survival functions’ method, considered previously by Cairns (2013) and Zhou and Li (2017), to avoid the nested simulations that would otherwise be required in the calculation of the following:

1. the values of the q-forwards at each time point when the hedge portfolio is adjusted, for each simulated sample path of future mortality;
2. the deltas of the q-forwards and annuity liability at each time point when the hedge portfolio is adjusted, for each simulated sample path of future mortality.

The application of the ‘approximation of survival functions’ method under the M7-M5 model assumption is significantly more complicated than that in the previous studies, due primarily to the fact that the M7-M5 model incorporates both period and cohort effects. To overcome this technical challenge, we systematically divide all possibly encountered survival functions into five cases, according to the duration, the starting age,
the starting time and the given information (filtration), and tailor a specific approximation (or calculation) method for each of the five cases.

In addition to aforementioned contributions, this paper adds value to the literature on longevity Greeks on the following aspects:

- Unlike Cairns (2011) and Zhou and Li (2017) who assume that the hedging instruments are costless to the hedger, we better mimic reality by allowing the counterparty of the q-forwards to charge a non-zero risk premium. Although the cost of hedging has no impact on hedge effectiveness measured in terms of variance reduction, its effect on asymmetric risk measures such as Value-at-Risk can be significant. Such an effect is examined in this paper.

- We consider not only cash flow hedges (of which the focus is the variability of cash flows) but also value hedges (of which the focus is the variability of the portfolio values at a certain future time point), in contrast to the existing work on discrete-time longevity Greek hedging which considers only the former. In line with Solvency II capital requirements, we measure the effectiveness of value hedges in terms of the reduction in the Value-at-Risk over a one-year horizon at a confidence level of 99.5%.

- We study the benefit of a dynamically adjusted hedge over a hedge that is left unadjusted over time. Cairns et al. (2008) has also investigated this issue, but their study takes no account of population basis risk and cohort effect uncertainty.

Our theoretical work is supplemented by three real data illustrations, which respectively demonstrate (1) the impact of cohort effect uncertainty and population basis risk on hedge effectiveness, (2) the benefit of dynamically adjusting a hedge portfolio in different market conditions, and (3) how the risk premium demanded by the counterparty may affect hedge effectiveness. The empirical work leads to several conclusions that may inform future studies of index-based longevity hedging. For example, it is found that if the liability being hedged is free of cohort effect uncertainty, then the effectiveness of a longevity hedge reduces as the extent of the cohort effect uncertainty surrounding the hedging instruments increases.

The rest of this paper is structured as follows. In Section 2, we specify the M7-M5 model and estimate it to real data. In Section 3, we define the ex post survival probability, from which most of the theoretical work in this paper is developed. In Section 4, we explain how survival probabilities can be approximated in different circumstances. The approximation methods are then applied in Section 5 where the valuation of the annuity liability and q-forwards is discussed, and in Section 6 where the longevity deltas of the annuity liability and q-forwards are derived. In Section 7, we define the metrics for evaluating hedge effectiveness, which are then used in Section 8 where the three real data illustrations are presented. Finally, Section 9 concludes the paper.

2 The M7-M5 Model

2.1 Model Specification

Proposed by Haberman et al. (2014), the M7-M5 model is a two-population stochastic mortality model that is formed by amalgamating the M5 model (the original Cairns-Blake-Dowd) model and the M7 model (the Cairns-Blake-Dowd model with quadratic age and cohort effects). In the M7-M5 model, one of the two populations being modelled is regarded as the dominant population (also referred to as the reference population), driving the mortality dynamics of both populations being modelled. The mortality dynamics of the reference population is assumed to follow an M7 model; that is,

\[
\ln \left( \frac{q_{x,t}^{(R)}}{1 - q_{x,t}^{(R)}} \right) = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + ((x - \bar{x})^2 - \sigma_x^2)\kappa_t^{(3)} + \gamma_{t-x},
\]

where

- \(x\) and \(t\) are integers representing age and time, respectively,
• \( q_{x,t}^{(R)} \) denotes the probability that an individual from the reference population \((R)\) dies in calendar year \(t\) (between time \(t - 1\) and time \(t\)), given that he/she has survived to age \(x\) at the beginning of year \(t\),

• \( \kappa_t^{(1)} \), \( \kappa_t^{(2)} \) and \( \kappa_t^{(3)} \) are the first, second and third period effects for calendar year \(t\), respectively,

• \( \gamma_{t-x} \) is the cohort effect for year-of-birth \(t - x\),

• \( \bar{x} \) is the mid-point of the age range to which the model is fitted, and

• \( \sigma_x^2 \) is the average value of \((x - \bar{x})^2\) over the age range to which the model is fitted.

The other population is called the book population. The mortality differential between the book and reference populations is assumed to follow an M5 model; that is,

\[
\ln \left( \frac{q_{x,t}^{(B)}}{1 - q_{x,t}^{(B)}} \right) - \ln \left( \frac{q_{x,t}^{(R)}}{1 - q_{x,t}^{(R)}} \right) = \kappa_t^{(1,B)} + (x - \bar{x})\kappa_t^{(2,B)}, \quad (2)
\]

where

• \( q_{x,t}^{(B)} \) denotes the probability that an individual from the book population \((B)\) dies in calendar year \(t\) (between time \(t - 1\) and time \(t\)), given that he/she has survived to age \(x\) at the beginning of year \(t\), and

• \( \kappa_t^{(1,B)} \) and \( \kappa_t^{(2,B)} \) are the period effects that determine the mortality differential in year \(t\).

Using equations (1) and (2), we have

\[
\ln \left( \frac{q_{x,t}^{(B)}}{1 - q_{x,t}^{(B)}} \right) = (\kappa_t^{(1)} + \kappa_t^{(1,B)}) + (x - \bar{x})(\kappa_t^{(2)} + \kappa_t^{(2,B)}) + ((x - \bar{x})^2 - \sigma_x^2)\kappa_t^{(3)} + \gamma_{t-x},
\]

which in turn means that the mortality dynamics of the book population also follow an M7 model, whose third period effect and cohort effect are identical to those in the M7 model for the reference population.

For ease of reading, we define \( y_{x,t}^{(i)} := \ln(q_{x,t}^{(i)}/(1 - q_{x,t}^{(i)}) \)), and express the M7-M5 model in a vector form as

\[
y_{x,t}^{(i)} = \beta_x^{(i)} \kappa_t^{(i)} + \gamma_{t-x}, \quad i = R, B,
\]

where

• \( \beta_x^{(R)} = (1, x - \bar{x}, (x - \bar{x})^2 - \sigma_x^2) \),

• \( \beta_x^{(B)} = (1, x - \bar{x}, (x - \bar{x})^2 - \sigma_x^2, 1, x - \bar{x}) \),

• \( \kappa_t^{(R)} = (\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)})' \) is the vector of period effects that are relevant to the reference population, and

• \( \kappa_t^{(B)} = (\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \kappa_t^{(1,B)}, \kappa_t^{(2,B)})' \) is the vector of period effects that are relevant to the book population.

We also define \( \kappa_t^{(R)} := (\kappa_t^{(1,B)}, \kappa_t^{(2,B)})' \) as the vector of period effects that determine the mortality differential.

### 2.2 Model Estimation

Let \([a, x_a] \) be the sample age range and \([t_a, t_b] \) be the sample period of the data set under consideration.

Fitting the M7-M5 model to such a data set yields estimates of \( \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \kappa_t^{(1,B)} \) and \( \kappa_t^{(2,B)} \) for \( t = t_a, \ldots, t_b \), and estimates of \( \gamma_{t-x} \) for \( t - x = t_a - x_a \) (the youngest cohort covered by the data set) to \( t - x = t_b - x_a \) (the oldest cohort covered by the data set). That being said, the period effects beyond calendar year \( t = t_b \) and the cohort effects beyond year-of-birth \( t_b - x_a \) have to be obtained by
extrapolations. Also, for ages \( x = x_b + 1, \ldots, \omega \), where \( \omega \) is the highest attainable age, the values of \( q_{x,t}^{(R)} \) and \( q_{x,t}^{(B)} \) are obtained by extrapolations with the built-in age functions in the M7-M5 model. We assume that \( \omega = 100 \) in our illustrations. The set-up is illustrated in Figure 1.

It is well known that the M7 model is subject to an identifiability problem. Following Haberman et al. (2014), we impose the following constraints to ensure parameter uniqueness in the M7 component:

\[
\sum_{c=x_a-x_b}^{t_b-x_a} \gamma_c = 0; \quad \sum_{c=x_a-x_b}^{t_b-x_a} c\gamma_c = 0; \quad \sum_{c=x_a-x_b}^{t_b-x_a} c^2\gamma_c = 0.
\]

With these constraints, the resulting values of \( \gamma_{t_a-x_b}, \ldots, \gamma_{t_b-x_a} \) would fluctuate around zero and exhibit no linear or quadratic trend. Note that the chosen constraints are also used in the paper by Cairns et al. (2009) where the M7 model is first introduced. We acknowledge that there exist other ways to formulate the identifiability constraints, but we choose to preserve the original setting specified by Haberman et al. (2014).

The illustrations in this paper are based on an M7-M5 model that is fitted to the historical mortality experience of the English and Welsh male population (the reference population) and U.K. male insured lives (the book population). The former population’s data are obtained from the Human Mortality Database, whereas that for the latter population’s data are obtained from the Institute and Faculty of Actuaries. The sample age range and sample period used are \([x_a, x_b] = [60, 89]\) and \([t_a, t_b] = [1961, 2005]\), respectively.

We estimate the M7 component using the method of binomial maximum likelihood, implemented with the R package StMoMo (Villegas et al., 2016), and estimate the M5 component using the method of least squares as in the original work of Cairns et al. (2006). Figure 2 shows the estimates of the period effects for calendar years \( t_a = 1961 \) to \( t_b = 2005 \), and the estimates of the cohort effects for years-of-birth \( t_a - x_b = 1961 - 89 = 1872 \) to \( t_b - x_a = 2005 - 60 = 1945 \).

Figure 1: A Lexis diagram showing the first and last period/cohort effects covered by the data sample, and an example of the period and cohort effects that have to be obtained by extrapolations.

---

1Haberman et al. (2014) estimated the M5 component with the method of binomial maximum likelihood instead. This alternative method yields very similar parameter estimates.
2.3 The Processes for the Period and Cohort Effects

Following Haberman et al. (2014), we assume that \( \kappa_t^{(R)} \) follows a tri-variate random walk,

\[
\kappa_t^{(R)} = \mu + \kappa_{t-1}^{(R)} + z_t^{(R)},
\]

where \( \mu \) is the drift vector, and \( z_t^{(R)} \) is the innovation vector which follows a tri-variate normal distribution with a zero mean vector and a constant covariance matrix of \( \Sigma^{(R)} \), and assume that \( \kappa_t^{(B)} \) follows a first-order vector-autoregressive process,

\[
\kappa_t^{(B)} = \theta_0 + \Theta_1 \kappa_{t-1}^{(B)} + z_t^{(B)},
\]

where \( \theta_0 \) is the offset vector, \( \Theta_1 \) is the 2-by-2 matrix of AR coefficients, and \( z_t^{(B)} \) is the innovation vector which follows a bivariate normal distribution with a zero mean vector and a constant covariance matrix of \( \Sigma^{(B)} \).

Haberman et al. (2014) do not assume any process for \( \gamma_{t-x} \), because their case study does not involve any unknown cohort effect. However, in our set-up, the liability being hedged and/or the hedging instruments may be associated with a year-of-birth that is beyond the last year-of-birth \( t_b - x_a \) covered by the data sample. Thus, a process for \( \gamma_{t-x} \) is needed to project the cohort effects for years-of-birth beyond \( t_b - x_a \). On the grounds that the estimated values of \( \gamma_{t-x} \) fluctuate around zero and exhibit no trend, we choose to assume a first-order autoregression for \( \gamma_{t-x} \):

\[
\gamma_{t-x} = \psi_0 + \psi_1 \gamma_{t-x-1} + z_{t-x},
\]

where \( \psi_0 \) and \( \psi_1 \) are the constant term and the autoregressive coefficient, respectively, and \( z_{t-x} \) is the innovation which follows a univariate normal distribution with a mean of zero and a constant variance of \( \sigma^2_\gamma \). It is further assumed that \( z_t^{(R)}, z_t^{(B)} \) and \( z_{t-x} \) are independently distributed and possess no serial
correlation.\textsuperscript{2} The estimates of the parameters (along with their standard errors) in the three processes are shown in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[\mu]_{1,1}$</td>
<td>$-1.7847 \times 10^{-2}$</td>
<td>$4.5310 \times 10^{-3}$</td>
</tr>
<tr>
<td>$[\mu]_{2,1}$</td>
<td>$3.9294 \times 10^{-4}$</td>
<td>$2.4359 \times 10^{-4}$</td>
</tr>
<tr>
<td>$[\mu]_{3,1}$</td>
<td>$3.8309 \times 10^{-5}$</td>
<td>$1.1701 \times 10^{-5}$</td>
</tr>
<tr>
<td>$[\theta_0]_{1,1}$</td>
<td>$-2.5406 \times 10^{-1}$</td>
<td>$2.8725 \times 10^{-2}$</td>
</tr>
<tr>
<td>$[\theta_0]_{2,1}$</td>
<td>$4.5888 \times 10^{-3}$</td>
<td>$1.2963 \times 10^{-3}$</td>
</tr>
<tr>
<td>$[\Theta_1]_{1,1}$</td>
<td>$4.6170 \times 10^{-1}$</td>
<td>$7.0964 \times 10^{-2}$</td>
</tr>
<tr>
<td>$[\Theta_1]_{1,2}$</td>
<td>$6.8288 \times 10^{-1}$</td>
<td>$3.1831$</td>
</tr>
<tr>
<td>$[\Theta_1]_{2,1}$</td>
<td>$-3.9184 \times 10^{-3}$</td>
<td>$3.2025 \times 10^{-3}$</td>
</tr>
<tr>
<td>$[\Theta_1]_{2,2}$</td>
<td>$1.8264 \times 10^{-2}$</td>
<td>$1.4365 \times 10^{-1}$</td>
</tr>
<tr>
<td>$\psi_0$</td>
<td>$-2.8093 \times 10^{-3}$</td>
<td>$3.9398 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\psi_1$</td>
<td>$9.0507 \times 10^{-1}$</td>
<td>$3.2408 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\sigma^2_\gamma$</td>
<td>$6.8077 \times 10^{-4}$</td>
<td>$5.6935 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

Table 1: Estimates (and their standard errors) of the parameters in the assumed processes of $\kappa^{(R)}_t$, $\kappa^{(B)}_t$ and $\gamma_{t-x}$.

Let $\mathcal{F}_t$ be the information concerning the evolution of mortality up to and including time $t$ (the end of year $t$) for $t = t_b, t_b + 1, \ldots$. For $t = t_b$, 

\[
\mathcal{F}_{t_b} = \{\kappa^{(R)}_{t_1}, \ldots, \kappa^{(R)}_{t_b}, \kappa^{(B)}_{t_1}, \ldots, \kappa^{(B)}_{t_b}, \gamma_{t_b-x_b}, \ldots, \gamma_{t_b-x_a}\}
\]

contains the values of the period and cohort effects that are estimated from the historical data, whereas for $t = t_b + 1, t_b + 2, \ldots$, 

\[
\mathcal{F}_t = \{\kappa^{(R)}_{t_1}, \ldots, \kappa^{(R)}_{t_b}, \kappa^{(R)}_{t_1}, \ldots, \kappa^{(R)}_{t_b}, \kappa^{(B)}_{t_1}, \ldots, \kappa^{(B)}_{t_b}, \gamma_{t_b-x_b}, \ldots, \gamma_{t_b-x_a}, \ldots, \gamma_{t-x_a}\}
\]

contains additionally the realized values of the period effects for calendar years $t_b + 1$ to $t$ and the realized values of the cohort effects for years-of-birth $t_b + 1 - x_a$ to $t - x_a$.

Given $\mathcal{F}_t$, $\kappa^{(R)}_t$, $\kappa^{(B)}_t$ and $\gamma_{t-x_a+s}$ for $s = 1, 2, \ldots$ can be expressed as

\[
\kappa^{(R)}_{t+s} = s\mu + \kappa^{(R)}_t + \sum_{v=0}^{s-1} z^{(R)}_{t+s-v},
\]

\quad \quad (4)

\[
\kappa^{(B)}_{t+s} = \sum_{v=0}^{s-1} \Theta^\gamma_1 \theta_0 + \Theta^\gamma_1 \kappa^{(B)}_t + \sum_{v=0}^{s-1} \Theta^\gamma_1 \gamma^{(B)}_{t-s-v},
\]

\quad \quad (5)

\textsuperscript{2}Haberman et al. (2014) made the assumption that $z^{(R)}_t$ and $z^{(B)}_t$ are independently distributed. We make the same assumption to preserve the original set-up. We have verified the validity of this assumption by computing the (empirical) pairwise correlations between the elements of $z^{(R)}_t$ and $z^{(B)}_t$ (from the fitted model) and by performing the Pearson’s correlation test for each correlation. The correlations (not shown for the sake of space) are found to be insignificant. We acknowledge that the correlations may be significant when another pair of populations is considered. If this is the case, then the user may choose to relax this assumption when simulating mortality scenarios for evaluating hedge effectiveness. Note that this assumption does not affect the calculation of longevity Greeks.
and
\[ \gamma_{t-x_u+s} = \sum_{i=0}^{s-1} \psi_1^i \psi_0 + \psi_1^i \gamma_{t-x_u} + \sum_{i=0}^{s-1} \psi_1^i \gamma_{t-x_u+s-v}, \] (6)
respectively. It follows that
\[ E[\kappa^{(R)}_{t+s}|\mathcal{F}_t] = s \mu + \kappa^{(R)}, \] (7)
\[ \text{Var}[\kappa^{(R)}_{t+s}|\mathcal{F}_t] = s \Sigma^{(R)}, \] (8)
\[ E[\kappa^{(B)}_{t+s}|\mathcal{F}_t] = (I - \Theta_1)^{-1}(I - \Theta_1^T) \theta_0 + \Theta_1^T \kappa^{(B)}, \] (9)
\[ \text{Var}[\kappa^{(B)}_{t+s}|\mathcal{F}_t] = (I - \Theta_1^2)^{-1}(I - \Theta_1^2) \Sigma^{(B)}, \] (10)
\[ E[\gamma_{t-x_u+s}|\mathcal{F}_t] = (1 - \psi_1)^{-1}(1 - \psi_1^s) \psi_0 + \psi_1^s \gamma_{t-x_u}, \] (11)
\[ \text{Var}[\gamma_{t-x_u+s}|\mathcal{F}_t] = (1 - \psi_1^2)^{-1}(1 - \psi_1^2) \sigma_\gamma^2, \] (12)
where \( I \) is a 2-by-2 identity matrix. Finally, as \( \kappa^{(R)}_{t+s}|\mathcal{F}_t \) is independent of \( \kappa^{(B)}_{t+s}|\mathcal{F}_t \), we have
\[ E[\kappa^{(B)}_{t+s}|\mathcal{F}_t] = \begin{pmatrix} E[\kappa^{(R)}_{t+s}|\mathcal{F}_t] \\ E[\kappa^{(B)}_{t+s}|\mathcal{F}_t] \end{pmatrix}, \]
and
\[ \text{Var}[\kappa^{(B)}_{t+s}|\mathcal{F}_t] = \begin{pmatrix} \text{Var}[\kappa^{(R)}_{t+s}|\mathcal{F}_t] & O_{3 \times 3} \\ O_{2 \times 2} & \text{Var}[\kappa^{(B)}_{t+s}|\mathcal{F}_t] \end{pmatrix}, \]
where \( O_{k \times k} \) is a \( k \)-by-\( k \) zero matrix. The results above are used extensively in later sections.

3 The Set-up

3.1 Survival Probabilities

We let
\[ S^{(i)}_{x,u}(T) = \prod_{s=1}^{T} (1 - q^{(i)}_{x+s-1,u+s}) \]
\[ = \prod_{s=1}^{T} \left( 1 + \exp(\beta^{(i)}_{x+s-1} \kappa^{(i)}_{u+s} + \gamma_{u-x+1}) \right)^{-1}, \]
for integer-valued \( x, u \) and \( T \), and \( i = R, B \), be the ex post probability that an individual (born in year \( u-x+1 \) from population \( i \), aged \( x \) at time \( u \) (the beginning of year \( u+1 \)) would have survived to the end of year \( u+T \). The second step in the above equation is achieved by applying the inverse-logit transform to \( q^{(i)}_{x+s-1,u+s} \). It is clear from the above expression that \( S^{(i)}_{x,u}(T) \) is a function of \( \kappa^{(i)}_{u+1}, \ldots, \kappa^{(i)}_{u+T} \) and \( \gamma_{u-x+1} \).

Using the definitions of \( \mathcal{F}_t \) and \( S^{(i)}_{x,u}(T) \), the following statements concerning the randomness surrounding \( S^{(i)}_{x,u}(T)|\mathcal{F}_t \) can be deduced.

- If \( t \leq u \), then none of the period effects in \( S^{(i)}_{x,u}(T) \) has been realized at time \( t \), and hence \( S^{(i)}_{x,u}(T)|\mathcal{F}_t \) depends on the unknown random values of \( \kappa^{(i)}_{u+1}|\mathcal{F}_t, \ldots, \kappa^{(i)}_{u+T}|\mathcal{F}_t \).
- If \( u < t < u+T \), then \( S^{(i)}_{x,u}(T)|\mathcal{F}_t \) depends on the realized values of \( \kappa^{(i)}_{u+1}|\mathcal{F}_t, \ldots, \kappa^{(i)}_{t}|\mathcal{F}_t \) and also the unknown random values of \( \kappa^{(i)}_{t+1}|\mathcal{F}_t, \ldots, \kappa^{(i)}_{u+T}|\mathcal{F}_t \).
- If \( t \geq u+T \), then all of the period effects in \( S^{(i)}_{x,u}(T) \) have been realized at time \( t \), and hence \( S^{(i)}_{x,u}(T)|\mathcal{F}_t \) is not subject to any period effect uncertainty.
• If \( t - x_a < u - x + 1 \), \( S^{(i)}_{x,u}(T) | \mathcal{F}_t \) depends on the unknown random value of \( \gamma_{u-x+1} | \mathcal{F}_t \), since \( \mathcal{F}_t \) contains cohort effects only up to and including \( \gamma_{t-x_a} \).

• If \( t - x_a \geq u - x + 1 \), then the cohort effect in \( S^{(i)}_{x,u}(T) \) has already been realized at time \( t \), and thus \( S^{(i)}_{x,u}(T) | \mathcal{F}_t \) is not subject to any cohort effect uncertainty.

It follows that \( S^{(i)}_{x,u}(T) | \mathcal{F}_t \) is a random variable if \( t < u + T \lor u - x + 1 + x_a \), and a known constant otherwise.\(^3\)

The expected value of \( S^{(i)}_{x,u}(T) \) given \( \mathcal{F}_t \),

\[
E[S^{(i)}_{x,u}(T) | \mathcal{F}_t],
\]

is crucially important in this study. In Sections 3.3 and 3.4, we explain how this conditional expectation is involved in the valuation of the liability being hedged and the hedging instruments. In Section 4, we discuss how this conditional expectation can be computed in different circumstances.

### 3.2 The \((u, \mathcal{F}_t)\)-Value of a Random Cash Flow Stream

The following definition is used throughout the rest of this paper.

**Definition 1.** The \((u, \mathcal{F}_t)\)-value of a cash flow stream is the conditional expectation of the discounted value of all of the cash flows beyond time \( u \), with \( u \) being the time point to which the cash flows are discounted and \( \mathcal{F}_t \) being the condition on which the expectation is taken.

### 3.3 The Liability Being Hedged

Suppose that the hedger wishes to establish a longevity hedge at an integer time point \( t_b \) (at the end of the last year of the sample period to which the M7-M5 model is fitted). The liability being hedged is a \( \tau \)-year deferred whole life annuity, payable to an individual from the book population \( B \), who has survived to age \( x_0 - \tau \) at time \( t_b \) (the beginning of year \( t_b + 1 \)). The annuity makes no payment in years \( t_b + 1, \ldots, t_b + \tau \), and pays $1 at the end of each year until death starting in year \( t_b + \tau + 1 \). Note that \( x_0 \) is the age of the annuitant at the beginning of the year in which the annuity begins to pay, and that the annuitant was born in year \( t_b - x_0 + \tau + 1 \).

Let \( r \) be the constant interest rate at which future payments are discounted. Discounting the liability cash flows to time \( t_b \) when the longevity hedge is established, we have

\[
\mathcal{L}^{(B)}(\tau) = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} S^{(B)}_{x_0-\tau,t_b}(\tau+s).
\]

It follows from Definition 1 that the \((t_b, \mathcal{F}_t)\)-value of the liability being hedged is

\[
L^{(B)}(\tau) := E[\mathcal{L}^{(B)}(\tau) | \mathcal{F}_t] = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} E[S^{(B)}_{x_0-\tau,t_b}(\tau+s) | \mathcal{F}_t],
\]

for \( t = t_b, t_b + 1, \ldots \).

### 3.4 The Hedging Instruments

Suppose that the hedging instruments used are q-forwards that are linked to the reference population \( R \). As the hedging instruments are not linked to the book population \( B \), the hedge is subject to population basis risk.

\(^3\)We use \( x \lor y \) to represent the minimum of \( x \) and \( y \).
Consider a q-forward whose date-of-issue is $t^f$, reference population is $R$, reference age is $x^f$ and time-to-maturity is $T^f$, where $t^f$, $x^f$ and $T^f$ are both integers. From the perspective of the fixed rate receiver, the payoff of this q-forward per $1$ notional at maturity is $q^f - q_x^{(R)}(t^f, x^f + T^f)$, where $q^f$ is the forward mortality rate which is a constant fixed at time $t^f$. We assume that all q-forwards used are issued on or after the day when the hedge is first established (i.e., $t^f \geq t_b$).

Discounting the q-forward’s payoff to time $t^f$ at a constant interest rate $r$, we have

$$Q^{(R)}(t^f, x^f, T^f) = (1 + r)^{-T^f}(q^f - q_x^{(R)}(t^f, x^f + T^f)) = (1 + r)^{-T^f}(S_{x^f, t^f + T^f, t^f - 1}^{(R)}(1) - (1 - q^f)).$$

It then follows from Definition 1 that the $(t^f, \mathcal{F}_t)$-value of the q-forward is given by

$$Q_t^{(R)}(t^f, x^f, T^f) := E[Q^{(R)}(t^f, x^f, T^f) | \mathcal{F}_t] = (1 + r)^{-T^f}(E[S_{x^f, t^f + T^f, t^f - 1}^{(R)}(1) | \mathcal{F}_t] - (1 - q^f)),$$

for $t = t^f, t^f + 1, \ldots$. Following Coughlan et al. (2007) and Li and Hardy (2011), we determine the forward mortality rate $q^f$ (fixed at time $t^f$ when the q-forward is launched) as a fraction of $E[q_x^{(R)}(t^f, x^f + T^f) | \mathcal{F}_t]$; that is

$$q^f = (1 - \lambda)(1 - E[S_{x^f, t^f + T^f, t^f - 1}^{(R)}(1) | \mathcal{F}_t]),$$

(13)

where $\lambda$ is a parameter that reflects the risk premium demanded by the counterparty.\(^4\)

4 Evaluation of $E[S_{x,u}^{(i)}(T) | \mathcal{F}_t]$

This section discusses the evaluation of $E[S_{x,u}^{(i)}(T) | \mathcal{F}_t]$. The evaluation method depends on the value of $t$ relative to the values of $u$ and $t_b$.

4.1 Computing $E[S_{x,u}^{(i)}(T) | \mathcal{F}_t]$ for $t = t_b$

This sub-section explains how $E[S_{x,u}^{(i)}(T) | \mathcal{F}_t]$ can be computed. We focus only on the situation when $u \geq t_b$, because the otherwise situation is never encountered in our set-up. For $u \geq t_b$, $S_{x,u}^{(i)}(T) | \mathcal{F}_{t_b}$ is a random variable with the following properties.

- All of the period effects in $S_{x,u}^{(i)}(T)$ are not contained in $\mathcal{F}_{t_b}$, so they are random as of $t = t_b$, depending on the value of $\kappa_{t_b}^{(i)}$ (by the Markov property of the period effect processes).
- If $t_b - x_a \geq u - x + 1$, then the cohort effect $\gamma_{u-x+1}$ in $S_{x,u}^{(i)}(T)$ is covered by $\mathcal{F}_{t_b}$ and is thus a known constant.
- If $t_b - x_a < u - x + 1$, then the cohort effect in $S_{x,u}^{(i)}(T)$ is not covered by $\mathcal{F}_{t_b}$, and is hence random as of $t = t_b$. Using the Markov property of the cohort effect process, it depends on the value of $\gamma_{t_b-x_a}$ (the latest cohort effect contained in $\mathcal{F}_{t_b}$).

\(^4\)As $q^f$ is merely a constant that determines the payoff at maturity, it carries no demographic meaning and is consequently not bounded by 0 and 1. In other words, equation (13) does not impose any restriction on $\lambda$. In principle, we may keep $q^f$ bounded by 0 and 1 (and impose no restriction on $\lambda$) with the aid of a distortion operator such as the Wang transform. We further remark that $\lambda$ can be formulated as an increasing function of $T^f$ to reflect the possibility that longer-dated q-forwards have larger risk premiums (see Li and Hardy (2011)).
These properties imply that

\[
E[S_{x,u}^{(i)}(T)|\mathcal{F}_{tb}] = \begin{cases} 
E[S_{x,u}^{(i)}(T)|\kappa_{tb}, \gamma_{tb-x_a}] & \text{if } t_b - x_a < u - x + 1 \\
E[S_{x,u}^{(i)}(T)|\kappa_{tb}, \gamma_{u-x+1}] & \text{if } t_b - x_a \geq u - x + 1,
\end{cases}
\]

which is equivalent to

\[
E[S_{x,u}^{(i)}(T)|\mathcal{F}_{tb}] = E[S_{x,u}^{(i)}(T)|\kappa_{tb}^{(i)}, \gamma_{tb-x_a \wedge u-x+1}].
\]  

This expectation is calculated by simulating sample paths of the period and/or cohort effects involved in \(S_{x,u}^{(i)}(T)\), given the values of \(\kappa_{tb}^{(i)}\) and/or \(\gamma_{tb-x_a}\).

### 4.2 Approximating \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_{t}]\) for \(t > t_b\)

At time \(t_b\) when the hedge is established (and when hedge effectiveness is evaluated), \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_{t}]\) for \(t > t_b\) is a random variable. The (empirical) distribution of \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_{t}]\) given \(\mathcal{F}_{tb}\) may be obtained using nested simulations.

The nested simulations involve a generation of \(N\) realizations of \(\mathcal{F}_t\) (i.e., \(N\) realizations of \(\kappa_{u}^{(i)}|\mathcal{F}_{tb}\) and \(\gamma_{u-x_a}|\mathcal{F}_{tb}\) for \(u = t_b + 1, \ldots, t\)). Also, for each of the \(N\) realizations of \(\mathcal{F}_t\), another \(M\) simulations are needed to obtain the value of \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\). In total, \(N \times M\) simulations are required to obtain an empirical distribution of \(N\) realizations of \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) given \(\mathcal{F}_{tb}\). As both \(N\) and \(M\) are typically large, the procedure is computationally demanding.

Drawing from the work of Cairns (2011), we develop methods to approximate \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) for any given realization of \(\mathcal{F}_t\), so that the second set of simulations is no longer required. The approximation method used depends on which one of the following three cases \(u\) and \(t\) fall into.

- **Case A:** \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) for \(t = u\)
- **Case B:** \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) for \(t < u\)
- **Case C:** \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) for \(t > u\)

The distinction between these three cases is illustrated in Figure 3. In the rest of this section, we explain how \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) is approximated in each case. The accuracy of the approximation methods is demonstrated in Appendix B.

#### 4.2.1 Case A: \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) for \(t = u\)

Using arguments similar to those made in Section 4.1, we obtain

\[
E[S_{x,u}^{(i)}(T)|\mathcal{F}_t] = E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x_a \wedge u-x+1}].
\]  

We approximate \(E[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) by applying a first-order Taylor expansion to the probit transform of \(E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x_a \wedge u-x+1}]\) around \(\hat{\kappa}_u^{(i)} := E[\kappa_{u}^{(i)}|\mathcal{F}_{tb}]\) and \(\hat{\gamma}_{u-x_a \wedge u-x+1} := E[\gamma_{u-x_a \wedge u-x+1}|\mathcal{F}_{tb}]\); that is,

\[
\Phi^{-1}(E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x_a \wedge u-x+1}]) \\
\approx d_{x,u,0}^{(i)}(T) + d_{x,u}(T)\prime(\kappa_{u}^{(i)} - \hat{\kappa}_u^{(i)}) + d_{x,u}^{(i)}(T)(\gamma_{u-x_a \wedge u-x+1} - \hat{\gamma}_{u-x_a \wedge u-x+1}),
\]  

where \(\Phi\) is the standard normal distribution function,

\[
d_{x,u,0}^{(i)}(T) = \Phi^{-1}(E[S_{x,u}^{(i)}(T)|\hat{\kappa}_u^{(i)}, \hat{\gamma}_{u-x_a \wedge u-x+1}]),
\]

\[
d_{x,u}^{(i)}(T) = (d_{x,u,1}^{(R)}(T), d_{x,u,2}^{(R)}(T), d_{x,u,3}^{(R)}(T))\prime,
\]

\[
d_{x,u}^{(i)}(T) = (d_{x,u,1}^{(R)}(T), d_{x,u,2}^{(R)}(T), d_{x,u,3}^{(R)}(T))\prime.
\]
Figure 3: A Lexis diagram illustrating the distinctions between Case A ($t = u$), Case B ($t < u$) and Case C ($t > u$).

and

$$d^{(B)}_{x,u}(T) = (d^{(B)}_{x,u,1}(T), d^{(B)}_{x,u,2}(T), d^{(B)}_{x,u,3}(T), d^{(B)}_{x,u,4}(T), d^{(B)}_{x,u,5}(T))',$$

with

$$d^{(i)}_{x,u,j}(T) = \frac{\partial}{\partial \kappa_{u,j}^{(i)}} \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1]) \bigg|_{\kappa_{u-x}^{(i)} = \kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1 = \gamma_{u-x} \wedge u-x+1},$$

$$d^{(B)}_{x,u,3+k}(T) = \frac{\partial}{\partial \kappa_{u,3+k}^{(B)}} \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(B)}, \gamma_{u-x} \wedge u-x+1]) \bigg|_{\kappa_{u-x}^{(B)} = \kappa_{u-x}^{(B)}, \gamma_{u-x} \wedge u-x+1 = \gamma_{u-x} \wedge u-x+1},$$

$$d^{(i)}_{x,u,\gamma}(T) = \frac{-1}{\partial \gamma_{u-x} \wedge u-x+1} \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1]) \bigg|_{\kappa_{u-x}^{(i)} = \kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1 = \gamma_{u-x} \wedge u-x+1}.$$ 

for $i = R, B$, $j = 1, 2, 3$ and $k = 1, 2$. Since the approximation is set around $\kappa_{u}^{(i)}$ and $\gamma_{u-x} \wedge u-x+1$, we evaluate $d^{(i)}_{x,u,j}(T)$, $d^{(B)}_{x,u,3+k}(T)$ and $d^{(i)}_{x,u,\gamma}(T)$ at $\kappa_{u}^{(i)} = \kappa_{u}$ and $\gamma_{u-x} \wedge u-x+1 = \gamma_{u-x} \wedge u-x+1$.

To derive $d^{(i)}_{x,u,j}(T)$, $d^{(B)}_{x,u,3+k}(T)$ and $d^{(i)}_{x,u,\gamma}(T)$, we use the fact that

$$\frac{\partial \Phi^{-1}(f(x))}{\partial x} = \frac{1}{\phi(\Phi^{-1}(f(x)))} \frac{\partial f(x)}{\partial x},$$

where $\phi$ is the standard normal probability density function. First, we have

$$d^{(i)}_{x,u,j}(T) = \frac{1}{\phi(\Phi^{-1}(E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1]))} \left( \frac{\partial}{\partial \kappa_{u,j}^{(i)}} E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1] \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1]) \right),$$

$$d^{(B)}_{x,u,3+k}(T) = \frac{1}{\phi(\Phi^{-1}(E[S_{x,u}^{(B)}(T)|\kappa_{u}^{(B)}, \gamma_{u-x} \wedge u-x+1]))} \left( \frac{\partial}{\partial \kappa_{u,x}^{(B)}} E[S_{x,u}^{(B)}(T)|\kappa_{u}^{(B)}, \gamma_{u-x} \wedge u-x+1] \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(B)}, \gamma_{u-x} \wedge u-x+1]) \right),$$

$$d^{(i)}_{x,u,\gamma}(T) = \frac{-1}{\partial \gamma_{u-x} \wedge u-x+1} \left( \frac{\partial}{\partial \gamma_{u-x} \wedge u-x+1} E[S_{x,u}^{(i)}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1] \Phi^{-1}(E[S_{x,u}(T)|\kappa_{u}^{(i)}, \gamma_{u-x} \wedge u-x+1]) \right).$$

(17)
In the above, \( \frac{\partial}{\partial \kappa_{i,u}^{(j)}} \prod_{s=1}^{T} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right) \) can be calculated recursively as

\[
\frac{\partial}{\partial \kappa_{i,u}^{(j)}} \prod_{s=1}^{T} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right) = \left\{ \begin{array}{ll}
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{i,u}^{(j)}} \exp(y_{x,u+1}^{(i)}) & \text{if } T = 1 \\
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{i,u}^{(j)}} \prod_{s=1}^{T-1} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right) + \left( 1 + \exp(y_{x+T-1,u+T}^{(i)}) \right) \frac{\partial}{\partial \kappa_{u}^{(j)}} \prod_{s=1}^{T-1} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right) & \text{if } T > 1
\end{array} \right.
\]

where

\[
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{i,u}^{(j)}} = \left\{ \begin{array}{ll}
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{u}^{(j)}} & \text{if } j = 1 \\
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{u}^{(j)}} = x + T - 1 - \bar{x} & \text{if } j = 2 \\
\frac{\partial y_{x+T-1,u+T}^{(i)}}{\partial \kappa_{u}^{(j)}} = (x + T - 1 - \bar{x})^2 - \sigma_x^2 & \text{if } j = 3
\end{array} \right.
\]

for \( T \geq 1 \) is obtained by using the chain rule and equations (3) and (4).

Second, \( y_{x,u,\gamma}^{(j)}(T) \) is obtained by replacing \( y_{x,u}^{(j)} \) in equations (17) and (18) with \( \kappa_{u}^{(j,B)} \). The result depends on the partial derivatives of \( y_{x,T-1,u+T}^{(j,B)} \) with respect to \( \kappa_{u}^{(1,B)} \) and \( \kappa_{u}^{(2,B)} \), which can be obtained by using the chain rule and equation (3) and (5):

\[
\frac{\partial y_{x,T-1,u+1}^{(j,B)}}{\partial \kappa_{u}^{(k,B)}} = \frac{\partial y_{x,T-1,u+1}^{(1,B)}}{\partial \kappa_{u}^{(1,B)}} + \frac{\partial y_{x,T-1,u+1}^{(2,B)}}{\partial \kappa_{u}^{(2,B)}} = \left[ \Theta_T^{1,k} \right]_i + (x + T - 1 - \bar{x})[\Theta_T^{1,k}]_j,
\]

where \( k = 1,2 \), where \( [\Theta_T^{1,k}]_i,j \) is the \( (i,j) \)th element in \( \Theta_T^{1,k} \).

Finally, \( y_{x,u,\gamma}^{(j)}(T) \) is obtained by replacing \( y_{x,u}^{(j)} \) in equations (17) and (18) with \( \gamma_{u-x_a,u-x+1} \). The result depends on the partial derivative of \( y_{x,T-1,u+T}^{(j)} \) with respect to \( \gamma_{u-x_a,u-x+1} \), which can be derived using the chain rule and equations (3) and (6):

\[
\frac{\partial y_{x,T-1,u+1}^{(j)}}{\partial \gamma_{u-x_a,u-x+1}} = \left\{ \begin{array}{ll}
\frac{\partial y_{x,T-1,u+1}^{(j)}}{\partial \gamma_{u-x_a,u-x+1}} & \text{if } u - x_a < u - x + 1 \\
\frac{\partial y_{x,T-1,u+1}^{(j)}}{\partial \gamma_{u-x_a,u-x+1}} = \psi_{u-x_a+1} - x + 1 & \text{if } u - x_a \geq u - x + 1
\end{array} \right.
\]

We can analytically calculate \( \hat{\kappa}_{u}^{(j)} \) and \( \hat{\gamma}_{u-x_a,u-x+1} \) using equations (7), (9) and (11). Given the values of \( \hat{\kappa}_{u}^{(j)} \) and \( \hat{\gamma}_{u-x_a,u-x+1} \), we can compute \( d_{x,u,0}^{(i)}(T), d_{x,u,j}^{(i)}(T), d_{x,u,3+k}^{(i)}(T) \) and \( d_{x,u,\gamma}^{(i)}(T) \) with a single set of (say \( M \)) simulations, which are notably independent of the (say \( N \)) simulations that are used to obtain realizations of \( F_u \). In effect, the total number of simulations needed to obtain an empirical distribution of \( N \) values of \( E[S_{x,u}^{(i)}(T)|F_u] \) is simply \( N + M \).

In contrast, if the approximation is not used, then the \( M \) simulations used to compute \( E[S_{x,u}^{(i)}(T)|F_u] \) would be dependent on the realization of \( F_u \). As a consequence, the total number of simulations needed to obtain an empirical distribution of \( N \) values of \( E[S_{x,u}^{(i)}(T)|F_u] \) would be \( N \times M \), which is substantially larger than \( N + M \) when both \( N \) and \( M \) are large.

### 4.2.2 Case B: \( E[S_{x,u}^{(i)}(T)|F_t] \) for \( t < u \)

As in Section 4.1 and 4.2.1, we use the Markov property of the assumed processes to obtain

\[
E[S_{x,u}^{(i)}(T)|F_t] = E[S_{x,u}^{(i)}(T)|\kappa_{t}^{(i)}, \gamma_{t-x_a,u-x+1}].
\]
In Case B, \( t \) and \( u \) take different values. Therefore, if the approximation method for Case A is used, then a specific approximation formula is needed for each \( u = t+1, t+2, \ldots \), thereby demanding significant computational effort. As such, for Case B, we employ a variant of Case A’s method, which yields the following approximation formula:

\[
E[S_{x,u}^{(i)}(T) | \kappa_t^{(i)}, \gamma_{t-x_a \wedge u-x+1}] \approx \Phi \left( \frac{-E[V_u^{(i)} | F_t]}{\sqrt{\text{Var}[V_u^{(i)} | F_t]}} \right), \quad i = R, B,
\]

where

\[
E[V_u^{(i)} | F_t] = -d_{x,u,0}^{(i)}(T) - d_{x,u}^{(i)}(T)'(E[\kappa_u^{(i)} | F_t] - \hat{\kappa}_u^{(i)}) - d_{x,u,\gamma}^{(i)}(T)(E[\gamma_{t-x_a \wedge u-x+1} | F_t] - \hat{\gamma}_{t-x_a \wedge u-x+1})
\]

and

\[
\text{Var}[V_u^{(i)} | F_t] = 1 + d_{x,u}^{(i)}(T)'\text{Var}[\kappa_u^{(i)} | F_t]d_{x,u}^{(i)}(T) + (d_{x,u,\gamma}^{(i)}(T))^2\text{Var}[\gamma_{t-x_a \wedge u-x+1} | F_t].
\]

The full derivation is presented in Appendix A. We remark that the derivation of equation (20) requires the use of the probit transformation. If a transformation other than probit is used, we may not be able to obtain an analytical approximation formula, as in equation (20).

Using equations (7) to (12), we can analytically calculate \( E[\kappa_u^{(i)} | F_t], \text{Var}[\kappa_u^{(i)} | F_t], E[\gamma_{t-x_a \wedge u-x+1} | F_t] \) and \( \text{Var}[\gamma_{t-x_a \wedge u-x+1} | F_t] \) for a given realization of \( F_t \). Therefore, as in Case A, with approximation formula (20), the number of simulations required to obtain an empirical distribution of \( N \) values of \( E[S_{x,u}^{(i)}(T) | F_t] \) for \( t > u \) is \( N + M \).

### 4.2.3 Case C: \( E[S_{x,u}^{(i)}(T) | F_t] \) for \( t > u \)

We further divide Case C into Cases C1 and C2 as follows:

- Case C1: \( E[S_{x,u}^{(i)}(T) | F_t] \) for \( t > u \) and \( t - x_a \geq u - x + 1 \);
- Case C2: \( E[S_{x,u}^{(i)}(T) | F_t] \) for \( t > u \) and \( t - x_a < u - x + 1 \).

In Case C1, the cohort effect in \( S_{x,u}^{(i)}(T) \) has already been realized at time \( t \), but in Case C2, \( S_{x,u}^{(i)}(T) | F_t \) still depends on the unknown random value of \( \gamma_{t-x+1} | F_t \). Figure 4 illustrates the distinction between Cases C1 and C2. The shaded area in the diagram represents the combinations of \( u \) and \( x \) that fall into Case C2.

**Case C1**

Let us first focus on Case C1. For \( u < t < u + T \), we have

\[
S_{x,u}^{(i)}(T) = S_{x,u}^{(i)}(t-u)S_{x+t-u,t}^{(i)}(T-t+u).
\]

As the relevant cohort effect has already been realized at time \( t \), \( S_{x,u}^{(i)}(t-u) \) given \( F_t \) is free of cohort effect uncertainty. Moreover, as \( F_t \) contains all of the period effects to which \( S_{x,u}^{(i)}(t-u) \) is linked, \( S_{x,u}^{(i)}(t-u) \) given \( F_t \) is also free of period effect uncertainty. Thus, \( S_{x,u}^{(i)}(t-u) \) given \( F_t \) is non-random, and we have

\[
E[S_{x,u}^{(i)}(T) | F_t] = E[S_{x,u}^{(i)}(t-u)S_{x+t-u,t}^{(i)}(T-t+u) | F_t]
\]

\[
= S_{x,u}^{(i)}(t-u)E[S_{x+t-u,t}^{(i)}(T-t+u) | F_t]
\]

\[
= S_{x,u}^{(i)}(t-u)E[S_{x+t-u,t}^{(i)}(T-t+u) | \kappa_t^{(i)}, \gamma_{u-x+1}]
\]

(21)

The last step in the above follows from equation (15) and Case C1’s condition that \( t - x_a \geq u - x + 1 \). The value of \( S_{x,u}^{(i)}(t-u) \) can be calculated using the period and cohort effects contained in \( F_t \), whereas the value of \( E[S_{x+t-u,t}^{(i)}(T-t+u) | \kappa_t^{(i)}, \gamma_{u-x+1}] \) can be approximated using the method introduced in Section 4.2.1. For \( t \geq u + T \), we have \( E[S_{x,u}^{(i)}(T) | F_t] = S_{x,u}^{(i)}(T) \), as all of the relevant period and cohort effects are contained in \( F_t \).
Figure 4: A Lexis diagram illustrating the distinction between Case C1 \((t > u and t - x_a \geq u - x + 1)\) and Case C2 \((t > u and t - x_a < u - x + 1)\). The shaded area represents the combinations of \(u\) and \(x\) that fall into Case C2.

**Case C2**

Next, we turn to Case C2. Given \(\mathcal{F}_t\), \(\gamma_{u-x+1}\) is still unknown and random; hence, given \(\mathcal{F}_t\), \(S_{x,u}^{(i)}(t-u)\) for \(u < t < u + T\) and \(S_{x,u}^{(i)}(T)\) for \(t \geq u + T\) are still random variables, even though all of the relevant period effects are contained in \(\mathcal{F}_t\). As a consequence, for \(u < t < u + T\), \(S_{x,u}^{(i)}(t-u)\) cannot be taken out from \(\mathbb{E}[S_{x,u}^{(i)}(t-u)S_{x+u}^{(i)}(T-t+u)|\mathcal{F}_t]\), and for \(t \geq u + T\), \(S_{x,u}^{(i)}(T)\) cannot be taken out from \(\mathbb{E}[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\). Thus, we have

\[
\mathbb{E}[S_{x,u}^{(i)}(T)|\mathcal{F}_t] = \begin{cases} 
\mathbb{E}[S_{x,u}^{(i)}(T)|\kappa_{u+1}^{(i)}, \ldots, \kappa_{t}^{(i)}, \gamma_{t-x_a}^{(i)}] & \text{if } u < t < u + T \\
\mathbb{E}[S_{x,u}^{(i)}(T)|\kappa_{u+1}^{(i)}, \ldots, \kappa_{u+T}^{(i)}, \gamma_{t-x_a}^{(i)}] & \text{if } t \geq u + T
\end{cases}
\]

An approximation formula for the above may be obtained by applying a first order Taylor’s expansion around \((\kappa_{u+1}^{(i)}, \ldots, \kappa_{t}^{(i)}, \gamma_{t-x_a}^{(i)})\) for \(u < t < u + T\) or \((\kappa_{u+1}^{(i)}, \ldots, \kappa_{u+T}^{(i)}, \gamma_{t-x_a}^{(i)})\) for \(t \geq u + T\). The approximation formula for Case C2 is inevitably much more tedious compared to those for the other cases. Fortunately, Case C2 is not encountered in our hedging set-up, given the assumptions made in the later sections.

**4.3 Summary**

Figure 5 summarizes the methods for calculating \(\mathbb{E}[S_{x,u}^{(i)}(T)|\mathcal{F}_t]\) in all possible circumstances. When \(t = t_b\), we use non-nested simulations. When \(t > t_b\), we use an approximation and the applicable approximation formula depends on the values of \(t\) and \(u\). In Case A \((t = u)\), we use approximation formula (16). In Case B \((t < u)\), we use approximation formula (20). In Case C1 \((t > u and t - x_a \geq u - x + 1)\), we decompose the expression using equation (21) and apply approximation formula (16). In Case C2 (which is not encountered in the rest of this paper), the approximation formula is significantly more involved.
5 Valuation of the Liability Being Hedged and the Hedging Instruments

5.1 The Liability Being Hedged

Let us revisit the annuity liability described in Section 3.3. Recall that the \((t_b, \mathcal{F}_t)\)-value of the annuity liability is

\[
L_t^{(B)}(\tau) = \sum_{s=1}^{\omega-x_0} (1+r)^{-(\tau+s)}E[S_{x_0-\tau,t_b}^{(B)}(\tau+s)|\mathcal{F}_t], \quad t = t_b, t_b+1, \ldots
\]

We calculate, for any given realization of \(\mathcal{F}_t\), the value of \(E[S_{x_0-\tau,t_b}^{(B)}(\tau+s)|\mathcal{F}_t]\) for each \(s = 1, \ldots, \omega - x_0\) using the methods developed in Section 4.

5.1.1 \(L_t^{(B)}(\tau)\) for \(t = t_b\)

For \(t = t_b\), we use equation (14) to get

\[
L_{t_b}^{(B)}(\tau) = \sum_{s=1}^{\omega-x_0} (1+r)^{-(\tau+s)}E[S_{x_0-\tau,t_b}^{(B)}(\tau+s)|\kappa_{t_b}(B), \gamma_{t_b-x_a\wedge t_b-x_0+\tau+1}],
\]

(22)

where, as described in Section 4.1, the value of \(E[S_{x_0-\tau,t_b}^{(B)}(\tau+s)|\kappa_{t_b}(B), \gamma_{t_b-x_a\wedge t_b-x_0+\tau+1}]\) for each \(s = 1, \ldots, \omega - x_0\) is obtained by non-nested simulations.

5.1.2 \(L_t^{(B)}(\tau)\) for \(t > t_b\)

When \(t > t_b\), \(E[S_{x_0-\tau,t_b}^{(B)}(\tau+s)|\mathcal{F}_t]\) falls into Case C, so its value is approximated by the methods presented in Section 4.2.3. In what follows, we assume that \(x_0 - \tau \geq x_a\), which equivalently means that the annuitant is aged at least \(x_a\) at time \(t_b\) (the beginning of year \(t_b+1\)). This working assumption is generally satisfied in practice, because the user may fit the M7-M5 model to an age range \([x_a, x_b]\) that covers the age of the youngest annuitant in his/her portfolio.

As \(t > t_b\) (the condition applicable to this sub-section) implies \(t - x_a \geq t_b - x_a + 1\) and \(x_0 - \tau \geq x_a\) (the assumption made) implies \(t_b - x_a + 1 \geq t_b - x_0 + \tau + 1\), we have \(t - x_a \geq t_b - x_0 + \tau + 1\). In other words, the
condition applicable to this sub-section and the assumption made in the previous paragraph imply that the cohort effect in \( S_{x_0-\tau,t_b}^{(B)} (\tau + s) \) is realized at or before time \( t \). Under this condition, \( E[S_{x_0-\tau,t_b}^{(B)} (\tau + s) | \mathcal{F}_t] \) always belongs to Case C1, and hence we can use the method for Case C1 to calculate the value of \( E[S_{x_0-\tau,t_b}^{(B)} (\tau + s) | \mathcal{F}_t] \) for each \( s = 1, \ldots, \omega - x_0 \).

Using equation (21) and approximation formula (16), \( L_t^B (\tau) \) for \( t_b < t \leq t_b + \tau \) can be calculated as follows:

\[
L_t^B (\tau) = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} S_{x_0-\tau,t_b}^{(B)} (t-t_b) E[S_{x_0-\tau+t-t_b,t}^{(B)} (\tau + s - t + t_b) | \mathcal{F}_t] \Delta (\kappa^{(B)}_t, \gamma_{t_b-x_0+\tau+1}),
\]

where

\[
E[S_{x_0-\tau+t-t_b,t}^{(B)} (\tau + s - t + t_b) | \mathcal{F}_t] \approx \Phi (d_{x_0-\tau+t-t_b,t,0} (\tau + s - t + t_b) + d_{x_0-\tau+t-t_b,t}^B (\gamma_{t_b-x_0+\tau+1} - \hat{\gamma}_{t_b-x_0+\tau+1})
\]

for \( s = 1, \ldots, \omega - x_0 \). Similarly, for \( t_b + \tau < t < t_b + \tau + \omega - x_0 \), we get

\[
L_t^B (\tau) = \sum_{s=1}^{t-t_b-\tau} (1 + r)^{-(\tau+s)} E[S_{x_0-\tau,t_b}^{(B)} (\tau + s) | \mathcal{F}_t] + \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} E[S_{x_0-\tau,t_b}^{(B)} (\tau + s) | \mathcal{F}_t]
\]

where

\[
E[S_{x_0-\tau+t-t_b,t}^{(B)} (s) | \mathcal{F}_t] \approx \Phi (d_{x_0-\tau+t-t_b,t,0} (s) + d_{x_0-\tau+t-t_b,t}^B (\kappa^{(B)}_t - \hat{\kappa}_t)
\]

for \( s = 1, \ldots, \omega - x_0 - t + t_b + \tau \). Lastly, for \( t \geq t_b + \tau + \omega - x_0 \), we have

\[
L^B_t (\tau) = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} S_{x_0-\tau,t_b}^{(B)} (\tau + s),
\]

since \( E[S_{x_0-\tau,t_b}^{(B)} (\tau + s) | \mathcal{F}_t] \) is realized at \( \tau + s = \omega - x_0 \).

### 5.2 The Hedging Instruments

As mentioned in Section 3.4, the hedging instruments used are \( q \)-forwards that are linked to the reference population \( R \). Recall that the \((t^f, \mathcal{F}_t)\)-value of a \( q \)-forward linked to the reference population \( R \) can be expressed as

\[
Q_t^R (t^f, x^f, T^f) = (1 + r)^{-T^f} E[S_{x^f,t^f+T^f-1(t^f)}^R (1) | \mathcal{F}_t] - (1 - q^f), \quad t = t^f, t^f + 1, \ldots,
\]

where \( t^f \) is the date of issue, \( T^f \) is the time to maturity, \( x^f \) is the reference age, and \( q^f \) is the forward mortality rate. Again, the values of \( E[S_{x^f,t^f+T^f-1(t^f)}^R (1) | \mathcal{F}_t] \) for \( t = t^f, t^f + 1, \ldots \) (including \( E[S_{x^f,t^f+T^f-1(t^f)}^R (1) | \mathcal{F}_{t^f}] \) in \( q^f \)) in the expression above are calculated with the methods developed in Section 4.
5.2.1 \( Q_t^{(R)}(t^f, x^f, T^f) \) for \( t = t_b \)

As \( t^f \geq t_b \) (an assumption made in Section 3.4), \( t = t_b \) (the condition applicable to this sub-section) and \( t \geq t^f \) (we only need to value a q-forward on or after its issue date), we have \( t^f = t_b \). Thus, we may use equation (14) to obtain

\[
Q_{t_b}^{(R)}(t^f, x^f, T^f) = (1 + r)^{-T^f}(E[S_{x^f, t^f + T^f - 1}(1)|\kappa_{t_b}^{(R)}, \gamma_{t_b - x_a + T^f - x^f}] - (1 - q^f)),
\]

and compute the value of \( E[S_{x^f, t^f + T^f - 1}(1)|\kappa_{t_b}^{(R)}, \gamma_{t_b - x_a + T^f - x^f}] \) using non-nested simulations (Case A, Section 4.1). When \( t^f = t_b \), the conditional expectation in \( q^f \) (specified by equation (13)) can also be calculated with non-nested simulations.

5.2.2 \( Q_t^{(R)}(t^f, x^f, T^f) \) for \( t > t_b \)

For \( t > t_b \), the approximation methods introduced in Section 4.2 are used. When \( t^f \leq t < t^f + T^f - 1 \), \( E[S_{x^f, t^f + T^f - 1}(1)|\mathcal{F}_t] \) falls into Case B, and hence we use approximation formula (20) to get

\[
Q_t^{(R)}(t^f, x^f, T^f) = (1 + r)^{-T^f}(E[S_{x^f, t^f + T^f - 1}(1)|\kappa_t^{(R)}, \gamma_{t - x_a + t^f + T^f - x^f}] - (1 - q^f))
\]

\[
\approx (1 + r)^{-T^f}\left(\Phi\left(-\frac{-E[V_{t+T^f-1}^{(R)}|\mathcal{F}_t]}{\sqrt{Var[V_{t+T^f-1}^{(R)}|\mathcal{F}_t]}}\right) - (1 - q^f)\right),
\]

where

\[
E[V_{t+T^f-1}^{(R)}|\mathcal{F}_t] = -d_{x^f, t^f + T^f - 1, 1.0}(1) - d_{x^f, t^f + T^f - 1, 1.1}(E[\kappa_{t+T^f-1}^{(R)}|\mathcal{F}_t] - \kappa_{t+T^f-1}^{(R)})
\]

and

\[
Var[V_{t+T^f-1}^{(R)}|\mathcal{F}_t] = 1 + d_{x^f, t^f + T^f - 1, 1.0}(1)^2Var[\kappa_{t+T^f-1}^{(R)}|\mathcal{F}_t]d_{x^f, t^f + T^f - 1, 1.0}(1) + (d_{x^f, t^f + T^f - 1, 1.1}(1))^2Var[\gamma_{t+T^f-1-x_a+t^f+T^f-x^f}|\mathcal{F}_t].
\]

As mentioned previously in Section 4.2.2, \( E[\kappa_{t+T^f-1}^{(R)}|\mathcal{F}_t] \), \( E[\gamma_{t+T^f-1-x_a+t^f+T^f-x^f}|\mathcal{F}_t] \), \( Var[\kappa_{t+T^f-1}^{(R)}|\mathcal{F}_t] \) and \( Var[\gamma_{t+T^f-1-x_a+t^f+T^f-x^f}|\mathcal{F}_t] \) in the above can be calculated analytically.

When \( t = t^f + T^f - 1 \), \( E[S_{x^f, t^f + T^f - 1}(1)|\mathcal{F}_t] \) falls into Case A. Hence, we can apply approximation formula (16) to obtain

\[
Q_t^{(R)}(t^f, x^f, T^f) \approx (1 + r)^{-T^f}\left(\Phi(d_{x^f, t^f + T^f - 1, 1.0}(1) + d_{x^f, t^f + T^f - 1, 1.1}(1)(\kappa_{t^f+T^f-1}^{(R)} - \kappa_{t^f+T^f-1}^{(R)})
\]

\[
+ d_{x^f, t^f + T^f - 1, 1.1}(1)(\gamma_{t^f+T^f-1-x_a+t^f+T^f-x^f} - \gamma_{t^f+T^f-1-x_a+t^f+T^f-x^f})) - (1 - q^f)\right).
\]

When \( t > t^f + T^f - 1 \), \( E[S_{x^f, t^f + T^f - 1}(1)|\mathcal{F}_t] \) falls into Case C. In what follows, we assume that \( x^f \geq x_a \), which equivalently means that the q-forward’s reference age is no smaller than \( x_a \). This working assumption is generally satisfied in practice, as the reference age of a q-forward is typically a pensionable age, which should be encompassed in the age range \([x_a, x_0]\) to which the M7-M5 model is fitted. As \( t > t^f + T^f - 1 \) implies \( t - x_0 \geq t^f + T^f - x_0 \) and \( x^f \geq x_a \) (the assumption made) implies \( t^f + T^f - x_0 \geq t^f + T^f - x^f \), we have \( t - x_0 \geq t^f + T^f - x^f \). In this condition, \( E[S_{x^f, t^f + T^f - 1}(1)|\mathcal{F}_t] \) always belongs to Case C1. Specifically, we have

\[
Q_t^{(R)}(t^f, x^f, T^f) = (1 + r)^{-T^f}(S_{x^f, t^f + T^f - 1}(1) - (1 - q^f))
\]

as \( E[S_{x^f, t^f + T^f - 1}(1)|\mathcal{F}_t] = S_{x^f, t^f + T^f - 1}(1) \) for \( t > t^f + T^f - 1 \).

Lastly, the value of \( q^f \) in equation (26) can be calculated as follows: if \( t^f = t_b \), use non-nested simulations; if \( t^f > t_b \) and \( T^f = 1 \), use approximation formula (16); if \( t^f > t_b \) and \( T^f > 1 \), use approximation formula (20).
6 Delta Hedging

In this section, we use the set-up and calculation methods developed previously to derive static and dynamic delta hedging strategies. The hedging strategies incorporate not only period effect uncertainty, but also cohort effect uncertainty and population basis risk.

6.1 Static Delta Hedging

A static delta hedge established at time $t_b$ is calibrated by matching the first-order partial derivatives of the $(t_b, \mathcal{F}_{t_b})$-values of the annuity liability and the q-forward portfolio with respect to the most recently realized period and cohort effects that are relevant to both the book and reference populations (i.e., $\kappa^{(1)}_{t_b}$, $\kappa^{(2)}_{t_b}$, $\kappa^{(3)}_{t_b}$ and $\gamma_{t_b-x_a}$). No adjustment is made to the q-forward portfolio after time $t_b$.

6.1.1 Sensitivities of the Hedging Instruments

In a static hedge, all of the q-forwards used are launched at time $t_b$. The partial derivatives of the $(t_b, \mathcal{F}_{t_b})$-value of a q-forward with an issue date of $t_b$, a reference age of $x^f$ and a time-to-maturity of $T^f$ with respect to $\kappa^{(1)}_{t_b}$, $\kappa^{(2)}_{t_b}$ and $\kappa^{(3)}_{t_b}$ are calculated as

$$
\frac{\partial Q_{t_b}(t_b, x^f, T^f)}{\partial \kappa^{(j)}_{t_b}} = (1 + r)^{-T^f} \frac{\partial}{\partial \kappa^{(j)}_{t_b}} E[S^{(R)}_{x^f, t^f+T^f-1}(1)|\mathcal{F}_{t_b}], \quad j = 1, 2, 3, \nonumber
$$

Also, if the q-forward is subject to cohort effect uncertainty (i.e., $t_b - x_a < t_b + T^f - x^f$), we calculate the partial derivative of $Q_{t_b}(t_b, x^f, T^f)$ with respect to $\gamma_{t_b-x_a}$ as

$$
\frac{\partial Q_{t_b}(t_b, x^f, T^f)}{\partial \gamma_{t_b-x_a}} = (1 + r)^{-T^f} \frac{\partial}{\partial \gamma_{t_b-x_a}} E[S^{(R)}_{x^f, t^f+T^f-1}(1)|\mathcal{F}_{t_b}]. \nonumber
$$

Otherwise, $\partial Q_{t_b}(t_b, x^f, T^f)/\partial \gamma_{t_b-x_a}$ is simply zero. The partial derivatives of the expectations in the expressions above are derived in Appendix C.

6.1.2 Sensitivities of Liability Being Hedged

The partial derivatives of the $(t_b, \mathcal{F}_{t_b})$-value of the annuity liability with respect to $\kappa^{(1)}_{t_b}$, $\kappa^{(2)}_{t_b}$ and $\kappa^{(3)}_{t_b}$ are calculated as

$$
\frac{\partial L_{t_b}^{(B)}(\tau)}{\partial \kappa^{(j)}_{t_b}} = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} \frac{\partial}{\partial \kappa^{(j)}_{t_b}} E[S^{(B)}_{x_0-\tau,t_b}(\tau+s)|\mathcal{F}_{t_b}], \quad j = 1, 2, 3, \nonumber
$$

Also, if the annuity liability is subject to cohort effect uncertainty (i.e., $t_b - x_a < t_b - x_0 + \tau + 1$), we calculate the partial derivative of $L_{t_b}^{(B)}(\tau)$ with respect to $\gamma_{t_b-x_a}$ as

$$
\frac{\partial L_{t_b}^{(B)}(\tau)}{\partial \gamma_{t_b-x_a}} = \sum_{s=1}^{\omega-x_0} (1 + r)^{-(\tau+s)} \frac{\partial}{\partial \gamma_{t_b-x_a}} E[S^{(B)}_{x_0-\tau,t_b}(\tau+s)|\mathcal{F}_{t_b}]. \nonumber
$$

Otherwise, $\partial L_{t_b}^{(B)}(\tau)/\partial \gamma_{t_b-x_a}$ is simply zero. The partial derivatives of the expectations in the expressions above are presented in Appendix C.

6.1.3 Calculating the Notional Amounts

We let $J$ be the number of q-forwards in the hedge portfolio, $h_j$ be the notional amount of the $j$-th q-forward, and $Q_{t_b}^{(R)}(t_b, x^f_j, T^f_j)$ be the $(t_b, \mathcal{F}_{t_b})$-value of the $j$-th q-forward with a reference age of $x^f_j$ and a
Appendix D (for derivatives of the expectations in the above expressions are presented in Appendix C (for otherwise, the partial derivative of \(Q_t\) is 

\[
(\text{purchased at time } t) \text{ are purchased. The process continues until the annuity liability completely runs off.}
\]

6.2 Dynamic Delta Hedging

6.2.1 Sensitivities of the Hedging Instruments

By assumption, all q-forwards purchased at time \(t\) have an issue date of \(t^f = t\). Let \(J_t\) be the number of q-forwards purchased at time \(t\). We use \(x_{t,j}^{f}\) and \(T_{t,j}^{f}\) to denote the reference age and time-to-maturity of the \(j\)-th q-forward purchased at time \(t\). Note that the \((t, \mathcal{F}_t)\)-value of \(j\)-th q-forward purchased at time \(t\) is \(Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})\).

To (re-)calibrate the hedge at time \(t = t_b, t_b + 1, \ldots\), we need the partial derivatives of \(Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})\) with respect to the most recently realized period and cohort effects as of time \(t\). We calculate the partial derivatives of \(Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})\) with respect to \(\kappa_t^{(1)}, \kappa_t^{(2)}\) and \(\kappa_t^{(3)}\) as

\[
\frac{\partial Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})}{\partial \kappa_t^{(j)}} = (1 + r)^{-T^f} \frac{\partial}{\partial \kappa_t^{(j)}} \mathbb{E}[S_{x_{t,j}^{f}, t+T_{t,j}^{f}-1}^{(R)}(1)|\mathcal{F}_t], \quad j = 1, 2, 3.
\]

Also, if the \(j\)-th q-forward purchased at time \(t\) is subject to cohort effect uncertainty (i.e., \(t - x_a < t + T_{t,j}^{f} - x_{t,j}^{f}\)), then we calculate the partial derivative of \(Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})\) with respect to \(\gamma_{t-x_a}\) as follows:

\[
\frac{\partial Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})}{\partial \gamma_{t-x_a}} = (1 + r)^{-T^f} \frac{\partial}{\partial \gamma_{t-x_a}} \mathbb{E}[S_{x_{t,j}^{f}, t+T_{t,j}^{f}-1}^{(R)}(1)|\mathcal{F}_t].
\]

Otherwise, the partial derivative of \(Q_t^{(R)}(t, x_{t,j}^{f}, T_{t,j}^{f})\) with respect to \(\gamma_{t-x_a}\) is simply zero. The partial derivatives of the expectations in the above expressions are presented in Appendix C (for \(t = t_b\) and Appendix D (for \(t > t_b\)).
6.2.2 Sensitivities of Liability Being Hedged

Using equations (22) to (24), we can rewrite the \((t_b, \mathcal{F}_t)\)-value of the annuity liability as

\[
L_t^{(B)}(\tau) = \\
\begin{cases}
FL_t^{(B)}(\tau), & \text{if } t = t_b \\
(1 + r)^{- (t-t_b)} S_{x_0-\tau,t_b}^{(B)}(t-t_b) FL_t^{(B)}(\tau), & \text{if } t_b < t \leq t_b + \tau \\
\sum_{s=1}^{\infty} (1 + r)^{- (\tau+s)} S_{x_0-\tau,t_b}^{(B)}(\tau+s) \\
+ (1 + r)^{- (t-t_b)} S_{x_0-\tau,t_b}^{(B)}(t-t_b) FL_t^{(B)}(\tau), & \text{if } t_b + \tau < t < t_b + \tau + \omega - x_0
\end{cases}
\]

where

\[
FL_t^{(B)}(\tau) = \\
\begin{cases}
\sum_{s=1}^{\infty} (1 + r)^{- (\tau+s)} E[S_{x_0-\tau,t_b}^{(B)}(\tau+s) | \mathcal{F}_t], & \text{if } t = t_b \\
\sum_{s=1}^{\infty} (1 + r)^{- (\tau+s-t+t_b)} E[S_{x_0-\tau+t-t_b,t}^{(B)}(\tau+s-t+t_b) | \mathcal{F}_t], & \text{if } t_b < t \leq t_b + \tau \\
\sum_{s=1}^{\infty} (1 + r)^{- (t+t_b+\tau)} (1 + r)^{- (t-x_0)} E[S_{x_0-\tau+t-t_b,t}^{(B)}(s) | \mathcal{F}_t], & \text{if } t_b + \tau < t < t_b + \tau + \omega - x_0
\end{cases}
\]

for \( t = t_b, t_b + 1, \ldots \) is the \((t, \mathcal{F}_t)\)-value of the annuity liability.

To (re-)calibrate the hedge at time \( t = t_b, t_b + 1, \ldots \), we need the partial derivatives of \( FL_t^{(B)}(\tau) \) with respect to the most recently realized period and cohort effects as of time \( t \). Using the information presented in Appendices C and D, we can easily compute the partial derivatives of \( FL_t^{(B)}(\tau) \) with respect to \( \kappa_t^{(1)} \), \( \kappa_t^{(2)} \) and \( \kappa_t^{(3)} \) for each \( t = t_b, t_b + 1, \ldots \), and the partial derivative of \( FL_t^{(B)}(\tau) \) with respect to \( \gamma_{t-x_a} \) for each \( t = t_b, t_b + 1, \ldots \) if \( FL_t^{(B)}(\tau) \) is still subject to cohort effect uncertainty (i.e., \( t - x_a < t_b - x_0 + \tau + 1 \)). If \( FL_t^{(B)}(\tau) \) is no longer subject to cohort effect uncertainty, its partial derivative with respect to \( \gamma_{t-x_a} \) is simply zero.

6.2.3 Calculating the Notional Amounts

Similar to a static hedge, the notional amounts of the q-forwards purchased at time \( t = t_b, t_b + 1, \ldots \) are determined by solving the following system of linear equations:

\[
\begin{pmatrix}
\frac{\partial Q_t^{(R)}(t,x_1,1,T_1)}{\partial \kappa_1^{(1)}} & \frac{\partial Q_t^{(R)}(t,x_2,1,T_1)}{\partial \kappa_1^{(1)}} & \frac{\partial Q_t^{(R)}(t,x_3,1,T_1)}{\partial \kappa_1^{(1)}} & \frac{\partial Q_t^{(R)}(t,x_4,1,T_1)}{\partial \kappa_1^{(1)}} \\
\frac{\partial Q_t^{(R)}(t,x_1,2,T_1)}{\partial \kappa_1^{(2)}} & \frac{\partial Q_t^{(R)}(t,x_2,2,T_1)}{\partial \kappa_1^{(2)}} & \frac{\partial Q_t^{(R)}(t,x_3,2,T_1)}{\partial \kappa_1^{(2)}} & \frac{\partial Q_t^{(R)}(t,x_4,2,T_1)}{\partial \kappa_1^{(2)}} \\
\frac{\partial Q_t^{(R)}(t,x_1,3,T_1)}{\partial \kappa_1^{(3)}} & \frac{\partial Q_t^{(R)}(t,x_2,3,T_1)}{\partial \kappa_1^{(3)}} & \frac{\partial Q_t^{(R)}(t,x_3,3,T_1)}{\partial \kappa_1^{(3)}} & \frac{\partial Q_t^{(R)}(t,x_4,3,T_1)}{\partial \kappa_1^{(3)}} \\
\frac{\partial Q_t^{(R)}(t,x_1,4,T_1)}{\partial \gamma_{t-x_a}} & \frac{\partial Q_t^{(R)}(t,x_2,4,T_1)}{\partial \gamma_{t-x_a}} & \frac{\partial Q_t^{(R)}(t,x_3,4,T_1)}{\partial \gamma_{t-x_a}} & \frac{\partial Q_t^{(R)}(t,x_4,4,T_1)}{\partial \gamma_{t-x_a}}
\end{pmatrix}
\begin{pmatrix}
h_{1,1} \\
h_{1,2} \\
h_{1,3} \\
h_{1,4}
\end{pmatrix}
= \\
\begin{pmatrix}
\frac{\partial FL_t^{(B)}(\tau)}{\partial \kappa_1^{(1)}} \\
\frac{\partial FL_t^{(B)}(\tau)}{\partial \kappa_1^{(2)}} \\
\frac{\partial FL_t^{(B)}(\tau)}{\partial \kappa_1^{(3)}} \\
\frac{\partial FL_t^{(B)}(\tau)}{\partial \gamma_{t-x_a}}
\end{pmatrix}. \\
(27)
\]

where \( h_{1,j}, j = 1, \ldots, J_t \), denotes the notional amount of the \( j \)-th q-forward purchased at time \( t \). All of the partial derivatives in the system of equations are evaluated at \( \kappa_t^{(1)} = \kappa_t^{(2)}, \kappa_t^{(3)} = \kappa_t^{(4)} \) and \( \gamma_{t-x_a} = \gamma_{t-x_a} \).

If both the annuity liability and the q-forwards are not subject to cohort effect uncertainty at time \( t \) (all partial derivatives with respect to \( \gamma_{t-x_a} \) are zero), then the above reduces to a system of three linear equations and only \( J_t = 3 \) q-forwards are needed.

7 Hedge Evaluation

In this section, we define two metrics for evaluating the effectiveness of the delta longevity hedges. The first metric, which is based on the reduction in variance, is suited for evaluating cash flow hedges of which the focus is the variability of (the discounted values of) the cash flows involved in the portfolio. The second metric, which is based on the reduction in Value-at-Risk, is developed for evaluating value hedges of which the focus is the variability of the values of the portfolio at a certain future time point.

21
7.1 Evaluation with Variance

7.1.1 Static Hedging

Let us first consider on a static hedge. Suppose that the hedge is evaluated at time \( t_b \), i.e., the time when the hedge is established. Discounted to time \( t_b \), the net cash outflows arising from the hedged and unhedged positions sum to

\[
L^{(B)}(\tau) - \sum_{j=1}^{J} h_{j} Q_{t_b+T_j^f}(t_b, x_{t_b}^f, T_j^f) \quad \text{and} \quad L^{(B)}(\tau),
\]

respectively. Therefore, the following metric quantifies the extent to which the static hedge reduces the variability of cash flows:

\[
HE = 1 - \frac{\text{Var}[L^{(B)}(\tau) - \sum_{j=1}^{J} h_{j} Q_{t_b+T_j^f}(t_b, x_{t_b}^f, T_j^f) | F_{t_b}]}{\text{Var}[L^{(B)}(\tau) | F_{t_b}]}.
\]

The closer to 1 the value of HE is, the better the hedge effectiveness is. The variances in the expression for HE can be calculated easily using non-nested simulations.

7.1.2 Dynamic Hedging

As with static hedging, \( L^{(B)}(\tau) \) still represents the annuity liability’s cash flows discounted to time \( t_b \) when the hedge is established and evaluated.

However, the cash flows from the hedge (the portfolio of \( q \)-forwards) are very different if it is dynamically adjusted instead of being kept unchanged. For a dynamic hedge, the following events occur for each \( t = t_b, \ldots, t_b + \tau + \omega - x_0 - 1 \):

- At time \( t \), write \( J_t \) freshly launched \( q \)-forwards. The notional amount for the \( j \)th \( q \)-forward (with a reference age of \( x_{t,j}^f \) and a time-to-maturity of \( T_{t,j}^f \)) is \( h_{t,j} \). Since all \( q \)-forwards written are freshly launched, this event does not incur any cash flow.

- At time \( t + 1 \), the \( J_t \) \( q \)-forwards written at time \( t \) are closed out. Per unit notional, the value of the \( j \)th \( q \)-forward at time \( t + 1 \) is \( Q_{t+1}^{(R)}(t, x_{t+1,j}^f, T_{t+1}^f) \). Therefore, this event generates a cash inflow of

\[
\sum_{j=1}^{J_t} h_{t,j} Q_{t+1}^{(R)}(t, x_{t,j}^f, T_{t,j}^f).
\]

Discounted to time \( t_b \), the cash inflows generated from the dynamically adjusted hedge sum to

\[
\sum_{t=t_b}^{t_b+\tau+\omega-x_0-1} J_t \sum_{j=1}^{J_t} (1+r)^{-(t-t_b)} h_{t,j} Q_{t+1}^{(R)}(t, x_{t+1,j}^f, T_{t+1,j}^f)
\]

and hence the net cash outflows arising from the hedged position sum to

\[
L^{(B)}(\tau) - \sum_{t=t_b}^{t_b+\tau+\omega-x_0-1} J_t \sum_{j=1}^{J_t} (1+r)^{-(t-t_b)} h_{t,j} Q_{t+1}^{(R)}(t, x_{t,j}^f, T_{t,j}^f) | F_{t_b}, \tag{28}
\]

It follows that the following metric quantifies the extent to which the dynamic hedge reduces the variability of cash flows:

\[
HE = 1 - \frac{\text{Var}[L^{(B)}(\tau) - \sum_{t=t_b}^{t_b+\tau+\omega-x_0-1} J_t \sum_{j=1}^{J_t} (1+r)^{-(t-t_b)} h_{t,j} Q_{t+1}^{(R)}(t, x_{t,j}^f, T_{t,j}^f) | F_{t_b}]}{\text{Var}[L^{(B)}(\tau) | F_{t_b}]}.
\]

A value of HE that is close to 1 indicates the hedge is effective.

To estimate HE, we simulate a large number of mortality scenarios (i.e., sample paths of \( \kappa_{t}^{(R)} | F_{t_b}, \kappa_{t}^{(B)} | F_{t_b} \) and \( \gamma_{t-t_b}^\omega | F_{t_b}, \) for \( t = t_b + 1, \ldots, t_b + \tau + \omega - x_0 \)). For each mortality scenario, we obtain the following:
1. A realization of $L^{(B)}(\tau)$

2. A realization of $h_{t,j}$ for $t = t_b, \ldots, t_b + \tau + \omega - x_0 - 1$ and $j = 1, \ldots, J_t$

In particular, for each $t = t_b, \ldots, t_b + \tau + \omega - x_0 - 1$, we calculate

$$
\frac{\partial Q^{(R)}_i(t, x^j_{t,j}, T^f_{t,j})}{\partial k^{(i)}_t}, \quad \frac{\partial Q^{(R)}_i(t, x^j_{t,j}, T^f_{t,j})}{\partial \gamma_{t,x}}, \quad \frac{\partial F^{(B)}_t(\tau)}{\partial k^{(i)}_t}, \quad \text{and} \quad \frac{\partial F^{(B)}_t(\tau)}{\partial \gamma_{t,x}}
$$

for $j = 1, \ldots, J_t$ and $i = 1, 2, 3$, using the procedures outlined in Appendices C and D. Note that the procedure in Appendix D utilizes the approximation methods presented in Section 4.2. With the calculated partial derivatives, we then compute $h_{t,j}$ for $j = 1, \ldots, J_t$ using equation (27).

3. A realization of $Q^{(R)}_{t+1}(t, x^j_{t,j}, T^f_{t,j})$ for $t = t_b, \ldots, t_b + \tau + \omega - x_0 - 1$ and $j = 1, \ldots, J_t$

Specifically, for $t = t_b, \ldots, t_b + \tau + \omega - x_0 - 1$, we calculate $Q^{(R)}_{t+1}(t, x^j_{t,j}, T^f_{t,j})$ with the approximation methods described in Section 5.2.2.

The algorithm above generates empirical distributions of $L^{(B)}(\tau)$ and (28), from which the variances in the expression for HE can be estimated.

The importance of the previously discussed approximation methods can now be clearly seen. Without the approximation methods, items 2 and 3 in the algorithm have to be calculated with simulations, thereby creating the situation of computationally demanding “simulations on simulations”.

### 7.2 Evaluation with Value-at-Risk

We now define a metric for measuring the effectiveness of a value hedge in terms of reduction in VaR. To avoid additional notation, we focus on a horizon of 1 year and a confidence level of 99.5% (consistent with Solvency II). With straightforward adaptations, other horizons and confidence levels can be considered.

First, note the following:

- The $(t_b, F_{t_b+1})$-value of the unhedged position is $L^{(B)}_{t_b+1}(\tau)$.
- The $(t_b, F_{t_b+1})$-value of the hedged position is $L^{(B)}_{t_b+1}(\tau) - \sum_{j=1}^J h_j Q^{(R)}_{t_b+1}(t_b, x^j_{t,b}, T^f_{j})$.

Given $F_{t_b}$, $L^{(B)}_{t_b+1}(\tau)$ and $Q^{(R)}_{t_b+1}(t_b, x^j_{t,b}, T^f_{j})$ are random variables, depending on the period and/or cohort effects realized at time $t_b + 1$. Note also that when viewed at time $t_b$, the expected value of the annuity liability is $E[L^{(B)}_{t_b+1}(\tau)|F_{t_b}] = L^{(B)}_{t_b}(\tau)$, which can be calculated easily with non-nested simulations.

It follows that the reduction in VaR (over the expected value of the annuity liability) produced by the delta hedge can be calculated as

$$
\text{HEVaR} = 1 - \frac{\text{VaR}_{99.5\%}[L^{(B)}_{t_b+1}(\tau) - \sum_{j=1}^J h_j Q^{(R)}_{t_b+1}(t_b, x^j_{t,b}, T^f_{j})|F_{t_b}] - L^{(B)}_{t_b}(\tau)}{\text{VaR}_{99.5\%}[L^{(B)}_{t_b+1}(\tau)|F_{t_b}] - L^{(B)}_{t_b}(\tau)}.
$$

A value of HEVaR that is close to 1 indicates that the hedge is effective in mitigating the variability (in terms of VaR) of the values of the portfolio one year from the time when the hedge is established.

The following procedure is used to calculate the two values of VaR$^{99.5\%}$ in HEVaR:

(i) Simulate realizations of $\kappa^{(R)}_t|F_{t_b}$, $\kappa^{(B)}_t|F_{t_b}$ and $\gamma_{t+x}|F_{t_b}$ for $t = t_b + 1$.

(ii) For each realization of the period and cohort effects, calculate the realized values of $L^{(R)}_{t_b+1}(\tau)$ and $Q^{(R)}_{t_b+1}(t_b, x^j_{t,b}, T^f_{j})$ for $j = 1, \ldots, J_t$, using the approximation methods described in Sections 5.1.2 and 5.2.2.

(iii) Using the result from the previous step, obtain empirical distributions of $L^{(B)}_{t_b+1}(\tau)$ and $L^{(B)}_{t_b+1}(\tau) - \sum_{j=1}^J h_j Q^{(R)}_{t_b+1}(t_b, x^j_{t,b}, T^f_{j})$, of which the 99.5th percentile yield the values of VaR$^{99.5\%}$ in the denominator and numerator in the expression for HEVaR, respectively.

Note that the approximations methods we developed also play a crucial role in the calculation of HEVaR (Step (ii) in the procedure above).
8 Illustrations

In this section, we provide three illustrations of the longevity hedging strategies developed in Section 6.

8.1 Illustration 1: Population Basis Risk and Cohort Effect Uncertainty

The first illustration provides the baseline results. It also analyzes the impact of cohort effect uncertainty and population basis risk on hedge effectiveness.

8.1.1 Assumptions

The following assumptions are used for Illustration 1:

- The hedge is a static hedge, which is established and evaluated at time \( t_b = 2005 \). Recall that the data sample used ends in year 2005.
- The liability being hedged is a deferred annuity-immediate sold at time \( t_b \). It has a deferment period of \( \tau \) years, and begins payment at the end of the year in which the annuitant attains age 65. The mortality experience of the annuitant is identical to that of the UK male insured lives.
- Two different values of \( \tau \) are considered:
  - \( \tau = 0 \): This value of \( \tau \) implies that the annuitant was born in 1941. As this year-of-birth is covered by the data sample, the annuity is not subject to any cohort effect uncertainty.
  - \( \tau = 9 \): This value of \( \tau \) implies that the annuitant was born in 1950. This year-of-birth is beyond the data sample, and thus the annuity is subject to cohort effect uncertainty.
- The q-forwards used are issued at the end of year 2005 (i.e., \( t_f = 2005 \)). We consider q-forwards with references ages 65, 69, 75 and 85, and times-to-maturity 5, 10 and 15 years. The table below shows the years-of-birth to which the q-forwards are linked. The number of asterisks (displayed next to the year-of-birth) indicates the relative level of cohort effect uncertainty.

<table>
<thead>
<tr>
<th>( T_f \times x_f )</th>
<th>65</th>
<th>69</th>
<th>75</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1945</td>
<td>1941</td>
<td>1935</td>
<td>1925</td>
</tr>
<tr>
<td>10</td>
<td>1950 (**)</td>
<td>1946 (*)</td>
<td>1940</td>
<td>1930</td>
</tr>
<tr>
<td>15</td>
<td>1955 (***)</td>
<td>1951 (**)</td>
<td>1945</td>
<td>1935</td>
</tr>
</tbody>
</table>

The q-forwards’ reference population is the English and Welsh population. As the q-forwards and the annuity liability are associated with different populations, population basis risk exists.

To focus on the goals of this illustration, we set \( \lambda \) to 0. This assumption is relaxed in Section 8.3 when the relationship between the cost of hedging and hedge effectiveness is studied.

- Three or four q-forwards are used in each hedge, depending on whether or not the sensitivity to \( \gamma_{t_b-x_a} \) is matched.
- The interest rate is at 5% per annum for all durations.

8.1.2 Baseline Results

The baseline results are displayed in Table 2. Hedges #1 to #3 are for the annuity liability with \( \tau = 0 \), whereas Hedges #4 to #8 are for the annuity liability with \( \tau = 9 \).
Table 2: The effectiveness of the static hedges for the annuity liabilities with $\tau = 0$ and $\tau = 9$. The specification of each q-forward used is presented by a triplet $(x^f, T^f, c^f)$, where $c^f = t_b + T^f - x^f$ is the cohort to which the q-forward is linked. The number of asterisks next to the triplet indicates the level of cohort effect uncertainty to which the q-forward is subject.

<table>
<thead>
<tr>
<th>Hedge</th>
<th>q-forward 1</th>
<th>q-forward 2</th>
<th>q-forward 3</th>
<th>q-forward 4</th>
<th>HE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#1</td>
<td>(65,5,1945)</td>
<td>(75,10,1940)</td>
<td>(85,15,1935)</td>
<td>-</td>
<td>0.83</td>
</tr>
<tr>
<td>#2</td>
<td>(75,5,1935)</td>
<td>(65,10,1950)(**)</td>
<td>(85,15,1935)</td>
<td>-</td>
<td>0.78</td>
</tr>
<tr>
<td>#3</td>
<td>(75,5,1935)</td>
<td>(85,10,1930)</td>
<td>(65,15,1955)(***)</td>
<td>-</td>
<td>0.68</td>
</tr>
<tr>
<td>$\tau = 9$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>(65,5,1945)</td>
<td>(75,10,1940)</td>
<td>(85,15,1935)</td>
<td>-</td>
<td>0.57</td>
</tr>
<tr>
<td>#5</td>
<td>(75,5,1935)</td>
<td>(65,10,1950)(**)</td>
<td>(85,15,1935)</td>
<td>-</td>
<td>0.60</td>
</tr>
<tr>
<td>#6</td>
<td>(75,5,1935)</td>
<td>(85,10,1930)</td>
<td>(65,15,1955)(***)</td>
<td>-</td>
<td>0.55</td>
</tr>
<tr>
<td>#7</td>
<td>(65,5,1945)</td>
<td>(75,10,1940)</td>
<td>(85,15,1935)</td>
<td>(65,10,1950)(*)</td>
<td>0.67</td>
</tr>
<tr>
<td>#8</td>
<td>(65,5,1945)</td>
<td>(75,10,1940)</td>
<td>(85,15,1935)</td>
<td>(69,10,1946)(*)</td>
<td>0.60</td>
</tr>
</tbody>
</table>

### Annuity with $\tau = 0$

We first focus on the annuity liability with $\tau = 0$. Since it is not subject to any cohort effect uncertainty, the sensitivity to $\gamma_{t_b-x_a}$ is not matched and three q-forwards are used.

Even though the sensitivity to $\gamma_{t_b-x_a}$ is not matched, the hedger may still use q-forwards that are subject to cohort effect uncertainty. In Hedge #1, all of the q-forwards are free of cohort effect uncertainty, but in Hedges #2 and #3, one of the q-forwards used is subject to cohort effect uncertainty.

The following findings are observed:

- Even when population basis risk exists, a static hedge with only three q-forwards can still achieve a reasonably high level of hedge effectiveness (up to 83% reduction in variance).
- Among the three hedges constructed, Hedge #1 (the only one that is free of cohort effect uncertainty), is the most effective. This outcome is not overly surprising, because using a q-forward that is exposed to cohort effect uncertainty (Hedges #2 and #3) would introduce cohort effect uncertainty to the hedger’s portfolio (which is originally free of such uncertainty). A similar reasoning can also explain why Hedge #2 yields a higher value of HE compared to Hedge #3.

### Annuity with $\tau = 9$

We now turn to the annuity liability with $\tau = 9$. Since it is subject to cohort effect uncertainty, we can match the sensitivity to $\gamma_{t_b-x_a}$, provided that at least one of the q-forwards used is also subject to cohort effect uncertainty. The following observations are made:

- By comparing Hedges #4, #5 and #6 with Hedges #1, #2 and #3, we can deduce that hedge effectiveness is significantly reduced if the annuity liability is subject to cohort effect uncertainty but the sensitivity to $\gamma_{t_b-x_a}$ is not matched.
- While both Hedges #5 and #6 contains a q-forward that is exposed to cohort effect uncertainty, Hedge #5 is more effective than Hedge #6. This outcome can be attributed to the fact that the second q-forward in Hedge #5 and the annuity liability are linked to the same cohort (year-of-birth 1950).
Hedges #4 and #7 are identical, except that Hedge #7 contains an additional q-forward that is subject to cohort effect uncertainty and is linked to the same cohort (year-of-birth 1950) as the annuity liability. It can be seen that by matching $\gamma_{t_b-x_a}$ using an additional q-forward that is also associated with year-of-birth 1950, the hedge effectiveness is significantly increased.

Hedge #8 also contains an additional q-forward that is subject to cohort effect uncertainty. However, the additional q-forward is not associated with year-of-birth 1950. As the cohorts to which the annuity liability and the additional q-forward are linked do not coincide, Hedge #8 is not as effective as Hedge #7. However, Hedge #8 is still more effective than Hedge #4, in which the sensitivity to $\gamma_{t_b-x_a}$ is not matched at all.

From these observations, we can conclude that when the liability being hedged is subject to cohort effect uncertainty, it is important to match the sensitivity to $\gamma_{t_b-x_a}$. Furthermore, whenever possible, the hedger should include in his/her portfolio a q-forward that is linked to the same cohort as the liability being hedged.

### 8.1.3 The Impact of Population Basis Risk

In the M7-M5 model, the extent of population basis risk is determined exclusively by the covariance matrix $\Sigma^{(B)}$ in the process for $\kappa^{(B)}_t$. To examine the impact of population basis risk, we now consider four hypothetical scenarios for which the specifications of $\Sigma^{(B)}$ are different. In Scenario 1, $\Sigma^{(B)} = 0$ so that the hedge is not subject to any population basis risk. In Scenarios 2 to 4, the diagonal elements in $\Sigma^{(B)}$ are increased to 2, 5 and 10 times their estimated values, respectively.

In Table 3, we compare, for each of eight hedges constructed, the values of HE in the four hypothetical scenarios with the baseline HE value. The following observations are made:

- Compared to the baseline HE value, the value of HE when population basis risk is completely absent (Scenario 1) is only slightly higher.
- Across Scenarios 1 to 4, the extent of population basis risk increases rapidly, but the hedge effectiveness decreases very slowly. Even if population basis risk is 10 times that in the baseline scenario, the hedges still perform reasonably well.
- The conclusions made in the previous sub-section are still valid under all of the scenarios.

<table>
<thead>
<tr>
<th>Hedge</th>
<th>Baseline</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.83</td>
<td>0.84</td>
<td>0.82</td>
<td>0.79</td>
<td>0.75</td>
</tr>
<tr>
<td>#2</td>
<td>0.78</td>
<td>0.79</td>
<td>0.77</td>
<td>0.75</td>
<td>0.71</td>
</tr>
<tr>
<td>#3</td>
<td>0.68</td>
<td>0.69</td>
<td>0.68</td>
<td>0.66</td>
<td>0.62</td>
</tr>
</tbody>
</table>


| #4    | 0.57     | 0.57       | 0.56       | 0.55       | 0.54       |
| #5    | 0.60     | 0.60       | 0.60       | 0.58       | 0.57       |
| #6    | 0.55     | 0.55       | 0.54       | 0.53       | 0.52       |
| #7    | 0.67     | 0.67       | 0.67       | 0.65       | 0.63       |
| #8    | 0.60     | 0.61       | 0.60       | 0.59       | 0.57       |

Table 3: The values of HE for the eight static hedges produced under the baseline specification of $\Sigma^{(B)}$ and the four alternatives specifications of $\Sigma^{(B)}$.

From these observations, we can conclude that if a longevity hedge is properly calibrated (using the strategies developed earlier in this paper), then the resulting hedge effectiveness is reasonably robust with respect to the level of population basis risk. In particular, with a proper calibration, the hedge effectiveness when there is a normal level of population basis risk is almost identical to that when population basis risk is completely absent.

---

*It should be noted that the hedge is still subject to structural basis risk.*
8.2 Illustration 2: Static vs. Dynamic

In this illustration, we focus on the difference in hedge effectiveness between a static hedge and a dynamic hedge.

8.2.1 Assumptions

The following assumptions are made for Illustration 2:

- Both hedges (static and dynamic) are established and evaluated at time $t_b = 2005$. The dynamic hedge is rebalanced annually until one year before the liability completely runs off (i.e., from time $t_b + 1 = 2006$ to $t_b + \tau + \omega - x_0 - 1 = 2044$).
- The liability being hedged is a deferred annuity-immediate sold at time $t_b = 2005$. It has a deferment period of 5 years, and begins payment at the end of the year in which the annuitant attains age 65. The mortality experience of the annuitant is identical to that of the UK male insured lives.
- These assumptions imply that the annuitant was born in 1946. This year-of-birth is not covered by the data sample, so at time $t_b$ the annuity is subject to cohort effect uncertainty. However, as the cohort effect for year-of-birth 1946 will be realized at time $t_b + 1$, the annuity liability will be free of cohort effect uncertainty one year after the hedges are established.
- The reference population of the q-forwards used is English and Welsh males. To focus on the goals of this illustration, $\lambda$ is again set to 0. All q-forwards used are freshly launched at the time when they are written.
- Two scenarios are considered. They differ in the specifications of the q-forwards used.
  - **Scenario I**
    For the static hedge, we use four q-forwards with $(x^f, t^f) = (65, 5), (69, 10), (75, 10), (85, 15)$. The q-forward with $(x^f, t^f) = (69, 10)$ is associated with the same year-of-birth (1946) as the annuity liability, and is subject to cohort effect uncertainty. The other three q-forwards are not subject to cohort effect uncertainty.
    For the dynamic hedge, we use four q-forwards with $(x^f, t^f) = (65, 5), (69, 10), (75, 10), (85, 15)$ when the hedge is first established at time $t_b = 2005$. As the annuity liability is free of cohort effect uncertainty after time $t_b + 1$, only three q-forwards, with $(x^f, t^f) = (65, 5), (75, 10), (85, 15)$, are used when the hedge is rebalanced. These three q-forwards are not subject to cohort effect uncertainty.
  - **Scenario II**
    For the static hedge, we use four q-forwards with $(x^f, t^f) = (69, 10), (70, 10), (75, 10), (80, 10)$. Note that they have the same time-to-maturity. The q-forward with $(x^f, t^f) = (69, 10)$ is associated with the same year-of-birth (1946) as the annuity liability, and is subject to cohort effect uncertainty. The other three q-forwards are not subject to cohort effect uncertainty.
    For the dynamic hedge, we use four q-forwards with $(x^f, t^f) = (69, 10), (70, 10), (75, 10), (80, 10)$ when the hedge is first established at time $t_b = 2005$. As the annuity liability is free of cohort effect uncertainty after time $t_b + 1$, only three q-forwards, with $(x^f, t^f) = (70, 10), (75, 10), (80, 10)$, are used when the hedge is rebalanced. These three q-forwards are not subject to cohort effect uncertainty.

8.2.2 Empirical Results

The resulting values of HE are tabulated in Table 4. The following observations and conclusions are made:

- In both scenarios, the dynamic hedge yields a significantly higher value of HE compared to the static hedge. This finding highlights the benefit from dynamically adjusting a longevity hedge over time.
• The static hedge in Scenario II (whereby all q-forwards have the same time-to-maturity) is substantially less effective than that in Scenario I. However, the dynamic hedges in both scenarios are similarly effective. This finding echoes that of Zhou and Li (2017), who identified and explained the empirical fact that the performance of a dynamically adjusted longevity hedge is robust with respect to the times-to-maturity of the q-forwards used.

This finding suggests that when the range of times-to-maturity in the q-forward market is limited, rebalancing a longevity hedge periodically is particularly beneficial. What is found in this illustration is practically relevant, because from the lessons learnt from the BNP/EIB longevity bond (see, e.g., Blake et al., 2006), we anticipate that it is unlikely that q-forwards with ultra-long times-to-maturity will become available in the market.

<table>
<thead>
<tr>
<th></th>
<th>Static Hedge</th>
<th>Dynamic Hedge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario I</td>
<td>0.74</td>
<td>0.97</td>
</tr>
<tr>
<td>Scenario II</td>
<td>0.57</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Table 4: The values of HE for the static and dynamic hedges in Scenarios I and II.

8.3 Illustration 3: Cost of Hedging

The third illustration investigates the impact of the cost of hedging. To this end, we now allow $\lambda$ (which reflects the risk premium demanded by the counterparty) to deviate from zero. Also, we now measure hedge effectiveness with HEVaR instead of HE, because the former takes the cost of hedging into account while the latter (which only measures the dispersion around the expected value) does not.

8.3.1 Assumptions

Other than the above assumptions regarding hedge evaluation and $\lambda$, the following assumptions are used in this illustration:

• The hedge is a static hedge, established and evaluated at time $t_b = 2005$.
• The liability being hedged is identical to that in Illustration 2.
• The q-forwards used are the same as those for the static hedge in Scenario 2 of Illustration 2.
• Four values of $\lambda$, namely 0, 0.005, 0.01 and 0.05, are considered.

8.3.2 Empirical Results

Table 5 shows the values of HEVaR for the four assumed values of $\lambda$. Figure 6 compares, for each of the four assumed values of $\lambda$, the empirical distributions of the $(t_{2005}, F_{2006})$-values of the hedged position (i.e., $L_{2006}^{(B)} - \sum_{j=1}^{4} h_j Q_{2006}^{(R)}$) and unhedged positions (i.e., $L_{2006}^{(B)}$) given $F_{2005}$. The following observations and conclusions can be made from Table 5 and Figure 6:

• When the q-forwards are costless (i.e., $\lambda = 0$), the value of HEVaR is 0.95, which equivalently means that the longevity hedge reduces the 1-year ahead 99.5% VaR (over the expected value of the annuity liability) by 95%.

• For all of the four assumed values of $\lambda$, the hedged positions are equally less dispersed than the unhedged position. This phenomenon confirms the fact that the cost of hedging does not affect the mitigation of the dispersion surrounding the expected value.

• However, as $\lambda$ increases, the distribution of $L_{2006}^{(B)} - \sum_{j=1}^{4} h_j Q_{2006}^{(R)}$ given $F_{2005}$ shifts rightwards, and hence has a higher 99.5th percentile. Consequently, the value of HEVaR decreases with $\lambda$. It can be deduced that beyond a certain threshold value of $\lambda$, the value of HEVaR would become negative, which in turn means that the longevity hedge would no longer be economically viable.
Table 5: The values of HEVaR for the four assumed values of $\lambda$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0</th>
<th>0.005</th>
<th>0.01</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEVaR</td>
<td>0.95</td>
<td>0.90</td>
<td>0.85</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Figure 6: The empirical distributions of the $(t^{2005}, F^{2006})$-values of the unhedged position (i.e., $L^{(B)}_{2006}$) and the hedged position (i.e., $L^{(B)}_{2006} - \sum_{j=1}^{4} h_j Q^{(R)}_{2006}$) given $F_{2005}$, obtained using the four assumed values of $\lambda$.

9 Conclusion

In this paper, we have contributed a discrete-time delta hedging strategy (constructed in both static and dynamic settings) for use with the M7-M5 model, a two-population mortality model that is recommended by the research team in charge of the population basis risk project (phase I) commissioned by the LLMA and IFoA. The hedging strategy takes into account of both cohort effect uncertainty and population basis risk, in contrast to those in previous work which ignore at least one of these two sources of randomness.

Under the M7-M5 model assumption, a survival function may involve both unrealized period effects and an unrealized cohort effect. As a consequence, the ‘approximation of survival function’ method, which is essential for keeping the computational effort entailed in a dynamic delta hedge manageable, is not straightforward to implement. We have overcome this technical challenge by systematically dividing all possibly encountered survival functions into five cases, and tailoring a specific approximation (or calculation) method for each.
We have defined several metrics for quantifying the effectiveness of cash flow hedges (of which the focus is the variability of cash flows) and value hedges (of which the focus is the variability of the portfolio values at a future time point). The metric based on the reduction in Value-at-Risk over a one-year horizon at a confidence level of 99.5% is particularly relevant to Solvency II capital requirements. Typically, the evaluation of such a metric requires simulations on simulations, but this need is waived by the approximation methods we consider.

We have provided three real data illustrations to supplement our theoretical work. The empirical work suggests the following points concerning index-based longevity hedging:

- If the liability being hedged is subject to cohort effect uncertainty, a delta hedge that matches the sensitivity to the most recently realized cohort effect significantly outperforms one that does not.
- Whether possible, the hedger should include in his/her hedge portfolio a q-forward that is linked to the same cohort as his/her annuity liability.
- If a longevity hedge is properly calibrated (with the strategies developed in this paper), then the resulting hedge effectiveness is reasonably robust with respect to the level of population basis risk.
- When the range of times-to-maturity of the q-forward market is limited, rebalancing a longevity hedge periodically is particularly beneficial.
- The hedge effectiveness measured in terms of Value-at-Risk reduces as the risk premium demanded by the counterparty of the q-forwards increases. When the risk premium becomes sufficiently high, a negative hedge effectiveness may be resulted.

Generally speaking, an index-based longevity hedge is subject to three sources of basis risk, namely (i) population basis risk (which arises from demographic or socioeconomic differences), (ii) sampling basis risk (which arises from the finite sample variation surrounding a fixed mortality trajectory), and (iii) structural basis risk (which arises from the differences in the payoff structures of the liability being hedged and the hedging instruments). In this paper, we have focused on (i) and addressed (iii) by deriving longevity deltas that are specific to the payoff structures of the annuity liability and q-forwards, but have paid no attention to (ii). Future research warrants a study that incorporates (ii), using the M7-M5 model in conjunction with a suitable death count process. Note that (ii) has no impact on the calculation of deltas, but may reduce hedge effectiveness depending on the size of the hedger’s portfolio.

We further remark that indexed-based longevity hedges are subject to some additional sources of risk that are not considered in this paper. The additional sources of risk include the following.

- **Recalibration risk**: We have used a model that is estimated to the data up to and including time \( t_b \) to evaluate hedge effectiveness and to calculate longevity Greeks at \( t = t_b, t_b + 1, \ldots \). In practice, one may re-estimate the model as new data becomes available. Any model recalibration beyond time \( t_b \) introduces additional uncertainty, as the revised parameter estimates may be different from the parameter estimates that are based on data up to and including time \( t_b \) only.

- **Parameter risk**: We have not taken into account the uncertainty surrounding the parameter estimates. This piece of uncertainty is particularly significant for the estimates of \( \gamma_c \) when \( c \) is close to \( t_a - x_b \) (the first year-of-birth covered by the data sample) and \( t_b - x_a \) (the last year-of-birth covered by the data sample), because such parameter estimates are based on very few data points only.

- **Model risk**: We have ignored model risk, as we have assumed that the M7-M5 model completely captures the evolution of mortality in the two populations. To assess the extent of model risk, one may assume that the hedge is constructed using the M7-M5 model but that the actual evolution of mortality is governed by another (two-population) mortality model.

We leave the study of these additional sources of risk in future research.
Acknowledgment

This work is supported by research grants from the Global Risk Institute, the Natural Sciences and Engineering Research Council of Canada (Discovery Grant RGPIN-356050-2013) and the Society of Actuaries Center of Actuarial Excellence Program.

References


A Deriving the Approximation Formula for Case B

Here we derive the approximation formula for $E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{u-x_u\wedge u-1}]$ when $t < u$. Let $Z$ be a standard normal random variable that is independent of the period and cohort effects. Using the approximation formula derived in Section 4.2.1; Case A) and when $t < u$ (Section 4.2.2; Case B). As discussed in the main text, the approximations are based on a first-order Taylor expansion that is applied to the probit transform of $E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{u-x_u\wedge u-1}]$ around $\kappa_u^{(i)} := E[\kappa_u^{(i)}|F_t]$ and $\hat{\gamma}_{u-x_u\wedge u-1} := E[\gamma_{u-x_u\wedge u-1}|F_t]$.

$$E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{u-x_u\wedge u-1}] \approx \Phi(d_{x,u,0}^{(i)}(T) + d_{x,u}^{(i)}(T)'(\kappa_u^{(i)} - \hat{\kappa}_u^{(i)}) + d_{x,u,\gamma}^{(i)}(T)(\gamma_{u-x_u\wedge u-1} - \hat{\gamma}_{u-x_u\wedge u-1}))$$

$$= \Pr(Z \leq d_{x,u,0}^{(i)}(T) + d_{x,u}^{(i)}(T)'(\kappa_u^{(i)} - \hat{\kappa}_u^{(i)}) + d_{x,u,\gamma}^{(i)}(T)(\gamma_{u-x_u\wedge u-1} - \hat{\gamma}_{u-x_u\wedge u-1}))$$

where $I_A$ is an indicator function which equals to one if event $A$ holds and 0 otherwise. The last step in the above follows from the Markov property of the assumed processes for the period and cohort effects. Using this result, we obtain

$$E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{t-x_u\wedge u-x+1}] = E[E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{u-x_u\wedge u-1}])|F_t]$$

$$\approx E\left[E\left[Z \leq d_{x,u,0}^{(i)}(T) + d_{x,u}^{(i)}(T)'(\kappa_u^{(i)} - \hat{\kappa}_u^{(i)}) + d_{x,u,\gamma}^{(i)}(T)(\gamma_{u-x_u\wedge u-1} - \hat{\gamma}_{u-x_u\wedge u-1}) | F_u \right] | F_t\right]$$

Again, the first step in the above is a consequence of the Markov property of the assumed processes. Let $V_u^{(i)} = Z - d_{x,u,0}^{(i)}(T) - d_{x,u}^{(i)}(T)'(\kappa_u^{(i)} - \hat{\kappa}_u^{(i)}) - d_{x,u,\gamma}^{(i)}(T)(\gamma_{u-x_u\wedge u-1} - \hat{\gamma}_{u-x_u\wedge u-1})$.

It is not hard to see that $V_u^{(i)}|F_t$ follows a normal distribution, and that

$$E[S_{x,u}(T)|\kappa_t^{(i)}, \gamma_{t-x_u\wedge u-x+1}] \approx \Phi \left( \frac{-E[V_u^{(i)}|F_t]}{\sqrt{\text{Var}[V_u^{(i)}|F_t]}} \right),$$

where

$$E[V_u^{(i)}|F_t] = -d_{x,u,0}^{(i)}(T) - d_{x,u}^{(i)}(T)'(E[\kappa_u^{(i)}|F_t] - \hat{\kappa}_u^{(i)}) - d_{x,u,\gamma}^{(i)}(T)(E[\gamma_{u-x_u\wedge u-1}|F_t] - \hat{\gamma}_{u-x_u\wedge u-1})$$

and

$$\text{Var}[V_u^{(i)}|F_t] = 1 + d_{x,u}^{(i)}(T)'\text{Var}[\kappa_u^{(i)}|F_t] d_{x,u}^{(i)}(T) + (d_{x,u,\gamma}^{(i)}(T))^2 \text{Var}[\gamma_{u-x_u\wedge u-1}|F_t].$$

B Evaluating the Accuracy of the Approximation Methods

In this appendix, we evaluate the accuracy of the methods for approximating $E[S_{x,u}(T)|F_t]$ when $t = u$ (Section 4.2.1; Case A) and when $t < u$ (Section 4.2.2; Case B). As discussed in the main text, the approximations are based on a first-order Taylor expansion that is applied to the probit transform of $E[S_{x,u}(T)|\kappa_u^{(i)}, \gamma_{u-x_u\wedge u-1}]$ around $\kappa_u^{(i)} := E[\kappa_u^{(i)}|F_t]$ and $\hat{\gamma}_{u-x_u\wedge u-1} := E[\gamma_{u-x_u\wedge u-1}|F_t]$.
evaluated. With a layout similar to Figure 7, Figure 8 shows the percentage errors of the approximation $\kappa_{2010}$, $\kappa_{2010}$, $\kappa_{2010}$, and $\gamma_{1946}$ given $F_{2005}$.

Case A: $E[S_{x,u}(T)|F_i]$ for $t = u$

The method presented in Case A approximates the value of $E[S_{x,u}(T)|F_u]$ for a given realization of $F_u$. For the sake of space, we only present the evaluation result for $E[S_{65,2010}(10)|F_{2010}]$, which is involved in the $(2005, F_{2010})$-value of the liability being hedged for Illustrations 2 and 3. The evaluation results for other combinations of $i, x, u$ and $T$ are similar.

It follows from equation (16) that the approximation formula for $E[S_{65,2010}(10)|F_{2010}]$ is a function of $\kappa_{2010}$, $\kappa_{2010}$, $\kappa_{2010}$, and $\gamma_{1946}$, all of which are random as of $t_b = 2005$ when the hedges in Illustrations 2 and 3 are evaluated. The contour lines in Figure 7 indicate the percentage errors (percentage deviations from the ‘actual’ values that are calculated using full nested simulations) of the approximation when it is applied to different combinations of the four variables. The dots in the figure represent 1,000 simulated values of the four variables given $F_{2005}$, so that the clouds of dots can be interpreted as the possible ranges of the four variables. It can be observed that within the boundary of each cloud, the maximum percentage error is only approximately 0.1%. We thereby conclude that the approximation method for Case A is reasonably accurate over the possible ranges of the four variables.

Case B: $E[S_{x,u}(T)|F_i]$ for $t < u$

The method presented in Case B approximates the value of $E[S_{x,u}(T)|F_t]$ for a realization of $F_t$ with $t < u$. Again, for the sake of space, we only show the evaluation results for $E[S_{70,2019}(1)|F_{2010}]$, which is involved in the $(2010, F_{2010})$-value of the q-forward that has a reference age of 70 and matures at the end of year 2020 (one of the q-forwards used in Illustration 2). For other combinations of $i, x, u, T$ and $t$, the evaluation results are similar.

It follows from equation (20) that the approximation formula for $E[S_{70,2019}(1)|F_{2010}]$ is a function of $\kappa_{2010}$, $\kappa_{2010}$, $\kappa_{2010}$, and $\gamma_{1950}$, all of which are random as of $t_b = 2005$ when the hedge in Illustration 2 is evaluated. With a layout similar to Figure 7, Figure 8 shows the percentage errors of the approximation
Figure 8: Contour plots showing the percentage errors in approximating $E[S_{70,2019}^{(i)}|F_{2010}]$. The dots in the figure represent simulated values of $\kappa_{2010}^{(1)}$, $\kappa_{2010}^{(2)}$, $\kappa_{2010}^{(3)}$ and $\gamma_{1950}$ given $F_{2005}$.

when it is applied to different combinations of the four variables. It can be observed that within the possible ranges of the four variables, the percentage errors are smaller than 0.01%, suggesting that the approximation method for Case B is highly accurate.

C Deriving the Sensitivities of $E[S_{x,u}^{(i)}(T)|F_t]$ for $t = t_b$

Here we derive the first-order partial derivatives of $E[S_{x,u}^{(i)}(T)|F_t]$ with respect to $\kappa_{t_b}^{(1)}$, $\kappa_{t_b}^{(2)}$ and $\kappa_{t_b}^{(3)}$ for $t = t_b$, $u \geq t_b$ and any possible values of $i$, $x$ and $T$, and that with respect to $\gamma_{t_b-x_a}$ for $t_b-x_a < u-x+1$.

The partial derivatives of $E[S_{x,u}^{(i)}(T)|F_t]$ with respect to $\kappa_{t_b}^{(j)}$ for $j = 1, 2, 3$ are given by

$$
\frac{\partial}{\partial \kappa_{t_b}^{(j)}} E[S_{x,u}^{(i)}(T)|F_t] = \frac{\partial}{\partial \kappa_{t_b}^{(j)}} \mathbb{E} \left[ \prod_{s=1}^{T} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right)^{-1} \bigg| F_{t_b} \right] \\
= -\mathbb{E} \left[ \left( S_{x,u}^{(i)}(T) \right)^2 \frac{\partial}{\partial \kappa_{t_b}^{(j)}} \prod_{s=1}^{T} \left( 1 + \exp(y_{x+s-1,u+s}^{(i)}) \right) \bigg| F_{t_b} \right].
$$
In the above, $\frac{\partial}{\partial \kappa^{(j)}_{t_b}} \prod_{s=1}^{T} \left(1 + \exp(y^{(i)}_{x+s-1,u+s})\right)$ can be calculated recursively as

$$
\frac{\partial}{\partial \kappa^{(j)}_{t_b}} \prod_{s=1}^{T} \left(1 + \exp(y^{(i)}_{x+s-1,u+s})\right) = \begin{cases} 
\frac{\partial y^{(i)}_{x+1,u+1}}{\partial \kappa^{(j)}_{t_b}} \exp(y^{(i)}_{x,u+1}) & \text{if } T = 1 \\
\frac{\partial y^{(i)}_{x+1,u+1}}{\partial \kappa^{(j)}_{t_b}} \prod_{s=1}^{T-1} \left(1 + \exp(y^{(i)}_{x+s-1,u+s})\right) + \left(1 + \exp(y^{(i)}_{x+T-1,u+T})\right) \frac{\partial y^{(i)}_{x,u+1}}{\partial \kappa^{(j)}_{t_b}} \prod_{s=1}^{T-1} \left(1 + \exp(y^{(i)}_{x+s-1,u+s})\right) & \text{if } T > 1
\end{cases}
$$

for $T \geq 1$ are obtained using the chain rule and equations (3) and (4).

The partial derivatives of $E[S^{(i)}_{x,u}(T) | \mathcal{F}_t]$ with respect to $\gamma_{t_b-x_a}$ for $t_b - x_a < u - x + 1$ can be obtained by replacing $\kappa^{(j)}_{t_b}$ with $\gamma_{t_b-x_a}$ and using the fact that

$$
\frac{\partial y^{(i)}_{x+T-1,u+T}}{\partial \gamma_{t_b-x_a}} = \frac{\partial y^{(i)}_{x+T-1,u+T}}{\partial \gamma_{u-x+1}} \frac{\partial \gamma_{u-x+1}}{\partial \gamma_{t_b-x_a}} = \psi_{1}^{u-x+1-t_b+x_a},
$$

which can be derived easily with the chain rule and equations (3) and (6).

**D Deriving the Sensitivities of $E[S^{(i)}_{x,u}(T) | \mathcal{F}_t]$ for $t > t_b$**

Here we derive the first-order partial derivatives of $E[S^{(i)}_{x,u}(T) | \mathcal{F}_t]$ with respect to $\kappa^{(1)}_{t_b}, \kappa^{(2)}_{t_b}, \kappa^{(3)}_{t_b}$ for $t > t_b$, $u \geq t$ and any possible values of $i$, $x$ and $T$, and that with respect to $\gamma_{-x_a}$ for $t-x_a < u-x+1$.

First, because of the Markov property of the assumed period and cohort effect processes, we have

$$
E[S^{(i)}_{x,u}(T) | \mathcal{F}_t] = E[S^{(i)}_{x,u}(T) | \kappa^{(i)}_t, \gamma_{-x_a} \wedge t-x_a].
$$

When $u = t$, we use approximation formula (16) to obtain

$$
\frac{\partial}{\partial \kappa^{(j)}_{t_b}} E[S^{(i)}_{x,u}(T) | \kappa^{(i)}_t, \gamma_{-x_a} \wedge u-x+1] \approx \frac{\partial}{\partial \kappa^{(j)}_{t_b}} \Phi \left(d^{(i)}_{x,t}(T) + d^{(i)}_{x,t,\gamma}(\kappa^{(i)}_t - \kappa^{(j)}_{t_b}) + d^{(i)}_{x,t,\gamma}(\gamma_{t-x_a} \wedge t-x-1 - \hat{\gamma}_{t-x_a} \wedge t-x+1)\right)
$$

for $j = 1, 2, 3$, and

$$
\frac{\partial}{\partial \gamma_{t-x_a}} E[S^{(i)}_{x,u}(T) | \kappa^{(i)}_t, \gamma_{-x_a} \wedge u-x+1] \approx \Phi \left(d^{(i)}_{x,t,0}(T) + d^{(i)}_{x,t}(T) + d^{(i)}_{x,t,\gamma}(\kappa^{(i)}_t - \kappa^{(i)}_t) + d^{(i)}_{x,t,\gamma}(\gamma_{t-x_a} \wedge t-x-1 - \hat{\gamma}_{t-x_a} \wedge t-x+1)\right)
$$

for $j = 1, 2, 3$, and

$$
\frac{\partial}{\partial \gamma_{t-x_a}} E[S^{(i)}_{x,u}(T) | \kappa^{(i)}_t, \gamma_{-x_a} \wedge u-x+1] \approx \Phi \left(d^{(i)}_{x,t,0}(T) + d^{(i)}_{x,t}(T) + d^{(i)}_{x,t,\gamma}(\kappa^{(i)}_t - \kappa^{(i)}_t) + d^{(i)}_{x,t,\gamma}(\gamma_{t-x_a} - \hat{\gamma}_{t-x_a})\right)
$$

for $j = 1, 2, 3$. 

36
when \( t - x_a < u - x + 1 \). When \( u > t \), we use approximation formula (20) to get

\[
\frac{\partial}{\partial \kappa_i^{(j)}} E[S_{x,u}^{(i)}(T) | \kappa_i^{(i)}, \gamma_{t-x_a \wedge u-x+1}] \approx \frac{\partial}{\partial \kappa_i^{(j)}} \Phi \left( \frac{-E[V_{u}^{(i)}|F_i]}{\sqrt{\text{Var}[V_{u}^{(i)}|F_i]}} \right)
\]

\[
= \phi \left( \frac{-E[V_{u}^{(i)}|F_i]}{\sqrt{\text{Var}[V_{u}^{(i)}|F_i]}} \right) \frac{d_{x,u,j}^{(i)}(T)}{\sqrt{\text{Var}[V_{u}^{(i)}|F_i]}}
\]

for \( j = 1, 2, 3 \), and

\[
\frac{\partial}{\partial \gamma_{t-x_a}} E[S_{x,u}^{(i)}(T) | \kappa_i^{(i)}, \gamma_{t-x_a \wedge u-x+1}] \approx \phi \left( \frac{-E[V_{u}^{(i)}|F_i]}{\sqrt{\text{Var}[V_{u}^{(i)}|F_i]}} \right) \frac{d_{x,u,j}^{(i)}(T)}{\sqrt{\text{Var}[V_{u}^{(i)}|F_i]}}
\]

when \( t - x_a < u - x + 1 \).