Astin Bulletin (2026), **56**, pp. 876–902
doi:10.1017/asb.2026.10095

International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

A Markov ageing-based multi-state group self-annuitisation

Rosario Maggistro[1], Mario Marino[1] and Annamaria Olivieri[2]

[1]Department of Economics, Business, Mathematics and Statistics "Bruno de Finetti", University of Trieste, Italy
[2]Department of Economics and Management, University of Parma, Italy
Corresponding author: Mario Marino; Email: mario.marino@deams.units.it

Received: 7 August 2025; **Revised:** 7 February 2026; **Accepted:** 26 March 2026; **First published online:** 24 April 2026

Keywords Pooled annuity funds; multi-state models; mortality heterogeneity; cross-subsidies; solidarity

Abstract

This paper extends the traditional group self-annuitisation framework by explicitly incorporating mortality heterogeneity among participants. Heterogeneity stems from multiple factors that lead individuals to age at different paces, despite being born in the same year. Ageing is modelled as a finite-state continuous-time Markov process where each state represents a distinct phase of physiological deterioration, and transitions capture the stochastic progression towards death. Benefits are differentiated by ageing state and, after issue, they are dynamically adjusted in response to the realised evolution of both ageing and mortality. Our design is novel in its use of the Markov ageing framework within a risk-sharing scheme and in how benefits are updated. Indeed, both benefits and their respective adjustment coefficients are state-specific. Through the explicit modelling of cross-subsidies across states, the design ensures that actuarial equivalence between benefits and available resources is preserved both at the pool level and within each ageing state. However, we find that benefit adjustments based on actuarial equivalence may display undesirable patterns in some ageing classes, when their size shrinks substantially; this happens, in particular, in the younger ageing states, which are likely to empty out. To contrast such effects, we introduce a design preserving a target level of differentiation across states that mitigates the unfavourable impact of a declining size for younger ages. In our analysis, we point out that such a design (which is desirable in many respects) implies solidarity effects across states. Such effects can be identified by comparing benefit amounts under the two assumptions (i.e., benefits adjusted according to actuarial equivalence or so to preserve a predefined level of differentiation). The proposed framework is tested using Australian mortality data.

1. Introduction

Many populations worldwide have experienced significant mortality improvements throughout the last decades, and human mortality has shifted to later ages (see Vaupel *et al.* 2021). The advancement of the longevity frontier implies compelling financial challenges for individuals aiming to manage their post-retirement income, as well as for annuity providers. Individuals living longer face the risk of outliving their wealth, and at least theoretically, they should look at the standard life annuity as an optimal vehicle for insuring against the individual longevity risk (Yaari 1965). But in reality, the voluntary annuity market remains thin because of rational and behavioural individual reasons driving the life annuities demand (see, e.g., Brown 2001; 2009). The downward trend in mortality rates, jointly with the uncertainty of financial markets and the evolving demographic structure of many populations, has led social security schemes and pension funds to revise pension benefits downwards. The shift from Defined Benefit to Defined Contribution plans has also taken hold. Consequently, individuals remain exposed to both individual investment and longevity risks; transferring such risks to a private annuity provider should be perceived as a beneficial strategy for building a suitable and stable retirement income. However, on their side, life insurers suffer the cost of financial and longevity guarantees traditionally embedded in life

© The Author(s), 2026. Published by Cambridge University Press on behalf of The International Actuarial Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

annuities, and they have become averse to offering non-adjustable longevity guarantees. The aggregate longevity risk cannot be diversified through risk pooling strategies and could require considerable equity capital backing due to risk-based solvency regulations. Then, academics and practitioners have examined alternative longevity risk management solutions. For instance, Blake *et al.* (2006) discuss how annuity providers can use mortality/longevity-linked securities to handle the risk of uncertain aggregate mortality. Even with great potential, the demand for the capital market securities proposed over the last decade has been very low, and annuity providers mainly relied on insurance-based solutions, such as buy-outs, buy-ins and insurance-based longevity swaps (Blake *et al.* 2019). Non-financial longevity reinsurance may also be considered, but it is usually costly and ultimately has to be funded by higher insurance premiums. As a workable alternative, longevity-linked arrangements have been proposed in the literature (see, e.g, Olivieri and Pitacco 2020). Beyond the specific design of these products, the underlying idea is that the insurer shares the risk with the policyholder by linking the life annuity benefits, directly or indirectly, to the mortality experience of a reference population. If unexpected mortality improvements occur, the insurer can mitigate its longevity risk exposure, but in the meantime, the benefits amount would decrease, likely hurting the demand for such insurance products.

Each longevity risk management solution has drawbacks, and insurers are motivated to modify their products by partially shifting risks to the policyholder. In the recent actuarial literature, notable attention has been given to pooling designs as innovative risk-sharing mechanisms to support the post-retirement income of individuals. Their common feature consists of grouping individuals to create a fund, from which each participant will receive periodical benefits adjusted according to the investment returns and mortality experienced by the pool, with no guarantee embedded. As a result, pool participants share among themselves the financial and longevity risk, and the insurer could play the role of administrator, not a guarantor (Dhaene and Milevsky 2024). Different pooling structures have been proposed in the literature, such as the pooled annuity funds (see, e.g., Donnelly et al., 2014; Donnelly, 2015), and tontine annuities (see, e.g., Milevsky and Salisbury, 2015; Chen et al., 2019; Weinert and Grüundl, 2021; Chen et al., 2022; Hieber and Lucas, 2022) among others.

In the present work, we refer to one of the earlier pooling structures, namely the group self-annuitisation (GSA). Introduced by Piggott *et al.* (2005), a GSA scheme is a mutual insurance plan wherein individuals are brought together to pool both their investment and longevity risks. Individuals joining the GSA scheme pay an initial contribution determined according to an annuity payout rate. Therefore, the benefits that GSA members receive periodically account for both the expected mortality dynamics and the expected return on investments. Considering that the pool is required to fully cover its liabilities with its own assets at all times, future benefits have to be adjusted if the actual mortality and/or the realised investments return diverge from their respective expectations. Qiao and Sherris (2013) further study the GSA plan by assessing its effectiveness when systematic longevity risk is also considered. Shedding light on the investment activity of GSA plans, Olivieri *et al.* (2022) propose a dynamic target volatility strategy to enhance the survival benefits of the pool participants. Kabuche *et al.* (2024) expand the GSA structure, allowing group participants to share both longevity and health risks. To this end, a multi-state GSA plan is proposed to pool heterogeneous individuals and provide annuity benefits updated according to both realised health transitions and mortality experienced. Aragona *et al.* (2025) introduce a redistributive GSA fund aiming to optimally share benefits among heterogeneous participants belonging to different socio-economic classes.

There is a growing attention, both in the market and literature, towards annuity and pooling designs providing differentiated benefits by health status, as a lever to foster annuitisation. The background assumption is that individuals who feel to have a shorter life expectancy, due to illnesses or lifestyle habits, may be reluctant to opt into a standard plan, while they might find it convenient to join arrangements that pay higher benefits should a shorter life expectancy be ascertained. Special rate (or underwritten or substandard) annuities are designed to respond to such a preference (see, for example, Pitacco and Tabakova 2022; Ramsay *et al.* 2025; de Andrés-Sánchez and González-Vila Puchades 2025); in the context of self-insured plans, the pooling arrangement proposed by Kabuche *et al.* (2024) is also suitable in this respect.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Against this background, we propose a multi-state framework that explicitly incorporates the mortality heterogeneity characterising GSA members, with the objective of defining a design admitting a benefit differentiation. Our focus, in particular, is on whether the GSA scheme should differentiate benefits so that frailer individuals (with shorter life expectancy and presumably greater current financial needs, due to higher healthcare or care-related expenses) receive higher payments than healthier individuals (who are expected to cash benefits for a longer time).

Individual mortality is influenced by various elements, such as current health status, lifestyles, socioeconomic conditions and climate characteristics. Each of these elements could significantly reshape the lifespan of individuals, and in particular, the timing of their ageing. In this paper, we refer to ageing as a continuous-time deterioration process of the human physiological functions necessary for survival. Individuals could bear a series of ageing phases before dying, and the speed at which ageing progresses varies from person to person, generating heterogeneity in the mortality of a population. While in a traditional GSA scheme benefits are adjusted based on aggregate realised mortality, we aim to design a GSA in which benefit adjustments account for both realised ageing transitions and mortality. Accordingly, a mortality model that explicitly accounts for the ageing process must be considered.

Individual's ageing does not necessarily move in lockstep with calendar time. The so-called biological age is an example of this, that is, the age indicating how old the human mechanism is at both the cellular and molecular levels (see, e.g., Jackson *et al.* 2003). Biological age is usually determined with a microscopic approach, that is, by collecting data concerning physiological and molecular variables for a large sample of people, and, through regression techniques, an increment or a reduction of the corresponding chronological age is inferred. We refer the interested reader to Li *et al.* (2023) for a discussion on the methods proposed in the literature for calculating biological age; we also refer the reader to Milevsky (2020) for an actuarial view on using biological age in mortality analysis. The biological age measurement requires high data granularity, and it is subject to both data collection and processing times. In contrast, annuity providers, as well as the administrators of a pooling fund, usually have access to mortality data only at an aggregate level, without having the possibility to observe biological factors. When there are no detectable biological factors other than age, the ageing process can still be represented, with the so-called Markov ageing model (see Lin and Liu 2007; Liu and Lin 2012; Su and Sherris 2012; Cheng *et al.* 2020). It is a finite-state continuous-time Markov process suitable for modelling the human ageing evolution. Each state of the Markov process identifies an ageing phase, and the overall ageing process is described through consecutive transitions from one ageing state to a higher one. An absorbing (death) state is also included, and the transition from any ageing state to the absorbing state represents the ageing process termination due to death. The Markov ageing model allows for linking directly the age-specific mortality pattern to the ageing process, and the intrinsic heterogeneity of a cohort can be expressed explicitly. In addition, the stochastic process has desirable analytical properties suitable for mortality analysis.

We define a pooling design, which we refer to as a Markov ageing GSA, in which individuals joining the pool are classified according to their initial ageing state, and their ageing trajectory is reassessed annually. Benefits are differentiated by state; upon transitioning to a higher ageing state, the individual's benefit is topped up to the level associated with the new state. Each individual contributes an initial capital assessed accounting for the initial state, the expected transition to higher ageing states and to the death state, as well as for the expected investment return. For the sake of simplicity, we assume a deterministic financial setting and a constant return on investments. This enables us to isolate and examine the effects on the benefit profile of the uncertain mortality dynamics in a context of a heterogeneous population. Similarly to a plain GSA, to preserve the actuarial equivalence between benefits and available resources, benefits are adjusted annually if experience deviates from expectation. In our setting, we account for both the realised ageing process pattern and actual mortality. Benefit adjustments are state-specific; any deficit or surplus arising from deviations between expected and realised experience is allocated across states in a way that reflects both the relative value of obligations in each state and the potential severity of deviations within each state. In a first phase of the analysis, benefit

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

differentiation and adjustments are defined so to ensure actuarial equivalence for each state. To this end, making cross-subsidy mechanisms between states explicit is crucial. However, as noted above, the ageing process is not directly observable at the individual level but must be inferred from proxy factors, such as the onset of disabling illnesses or the worsening of lifestyle behaviours. This is a critical issue, in respect of both suitability of the proxy risk factors and potential self-interested behaviours by individuals. Participants, in particular, may have specific targets regarding the degree of benefit differentiation, which are not necessarily ensured by an actuarially equivalent adjustment mechanism. In a second phase of the analysis, we explore the adoption of benefits that are differentiated by the states identified by proxy risk factors (rather than the ageing states) and that keep a predetermined scaling factor between them, chosen according to the desired level of differentiation. Both choices introduce solidarity mechanisms, which we can highlight by comparing the resulting benefits with the ones grounded on actuarial equivalence by ageing state. Our work differs from previous contributions on this topic, such as Kabuche *et al.* (2024), in two regards. First, the use of a Markov ageing model allows us to incorporate latent heterogeneity that cannot be directly measured. Second, we address explicitly possible solidarity mechanisms, identifying their magnitude. Understanding the degree of solidarity built into the system is fundamental to ensure transparency and to make any redistribution a deliberate and justifiable design choice. We develop our proposal referring to a GSA closed to the entry of new cohorts. This allows us to understand the extent to which accounting for mortality heterogeneity leads to efficient mortality risk pooling within a single cohort. We test our theoretical framework through a numerical application using a 5-state Markov ageing model and referring to Australian mortality data for both genders. Our findings support that individual longevity risk can be effectively pooled under population heterogeneity, allowing for differentiated benefits reflecting differences in participants' life expectancies. Our results also indicate that actuarial equivalence requires a sufficiently large population size, whereas implementing solidarity mechanisms may be justified. Indeed, the estimated benefits profile appears generally stable, apart from state-specific fluctuations when the cohort shrinks to a low size. The severity of ageing significantly drives the dynamics of benefits. The last ageing state's benefit profile is increasing on average; conversely, the intermediate ageing states' benefit profile, as well as the first, tends to decrease as the maximum duration of the scheme is approached. Such evidence bolsters the introduction of explicit solidarity mechanisms to avoid disincentivising the participation of individuals with better physiological statuses. Then, we numerically investigate the effect of solidarity rules on rearranging benefit levels and preserving the actuarial equivalence between benefits and available resources. Depending on the number of benefit levels the scheme would provide, the benefit profile may increase more or less marked over time and in a consistent way across states, while remaining reasonably differentiated among frailer and healthier individuals.

The paper is organised as follows. In Section 2, we describe in detail the Markov ageing model. In Section 3, we present the actuarial design of the multi-state GSA, in particular addressing cross-subsidies across states and the introduction of solidarity. In Section 4, a numerical application of the proposed theoretical framework is provided, as well as a discussion of the main findings. Finally, Section 5 concludes with some final remarks.

2. The Markov ageing model

Let $(A_t,\ t \geq 0)$ be a finite-state continuous-time homogeneous Markov process describing the ageing evolution of a population, and let $E = \{1, 2, \ldots, m\}$, with $m \in \mathbb{N}_{>1}$, be the set of transient states. Each of them identifies an ageing state or phase. The process takes values on the finite state-space $\mathcal{S} = E \cup \{m+1\}$, where $m+1$ represents the death (absorbing) state. The infinitesimal generator of $(A_t,\ t \geq 0)$ is

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Lambda} & \boldsymbol{q} \\ \mathbf{0} & 0 \end{pmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \tag{2.1}$$

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

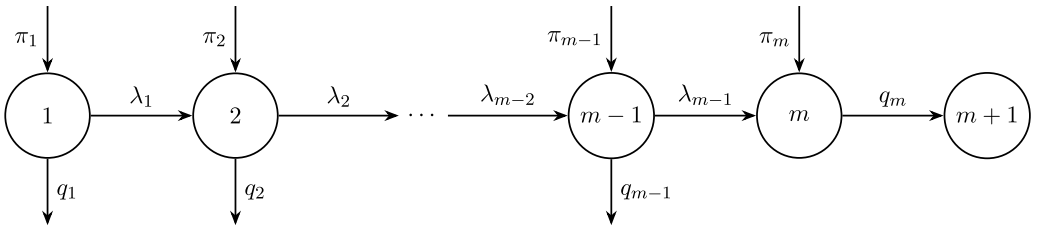


Figure 1. *Generalised Coxian distribution's phase diagram.*

where $\mathbf{\Lambda}$ is the $m \times m$ intensity matrix and $\boldsymbol{q} = -\mathbf{\Lambda}\boldsymbol{e}$ is a m-dimensional column vector of transition intensities to the death state, with $\boldsymbol{e}$ a column vector of ones.

Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_m)$ be a probability vector stating the initial distribution of the process, with $\pi_i := \mathbb{P}(A_0 = i) \geq 0,\ i \in E$, and $\mathbb{P}(A_0 = m+1) = 0$. Since the Markov process has only one absorbing state, the time until death is defined as

$$\tau := \inf\{t > 0 : A_t = m+1\}, \tag{2.2}$$

and follows a phase-type (PH) distribution with representation $(\boldsymbol{\pi}, \mathbf{\Lambda})$, that is, $\tau \sim \text{PH}(\boldsymbol{\pi}, \mathbf{\Lambda})$. In particular, we deal with a generalised Coxian PH distribution so that the intensity matrix is written in the following form:

$$\mathbf{\Lambda} = \begin{pmatrix} -(\lambda_1 + q_1) & \lambda_1 & 0 & \cdots & 0 \\ 0 & -(\lambda_2 + q_2) & \lambda_2 & \cdots & 0 \\ 0 & 0 & -(\lambda_3 + q_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -q_m \end{pmatrix}, \tag{2.3}$$

being $\lambda_i \geq 0$ the transition intensity from the ageing state i to the next ageing state, while $q_i > 0$ is the transition rate into the absorbing state. The generalised Coxian PH distribution is particularly attractive due to its mathematical tractability and explainability. As shown by the phase diagram in Figure 1, the ageing process is described through a sequence of ageing phases, following a lexicographical order, and the death event may occur with a certain probability in the meantime. The sojourn time in the ageing phase i is exponentially distributed with rate $\lambda_i + q_i$, and, for $i = 1, \ldots, m-1$, the process moves to the ageing phase $i+1$ with probability $\frac{\lambda_i}{\lambda_i + q_i}$; otherwise, the ageing process terminates due to the death occurrence with probability $\frac{q_i}{\lambda_i + q_i}$. As ageing progresses, the physiological status of the population deteriorates, implying a greater probability of approaching the absorbing state. In particular, when $i = m$, the probability of jumping to the death state is 1.

A key issue in adopting a PH distribution concerns the dimension of $\mathbf{\Lambda}$. The PH representation $(\boldsymbol{\pi}, \mathbf{\Lambda})$ is highly parameterised and non-uniqueness problems in parameter specifications could emerge (see, e.g., Asmussen *et al.* 1996). As m increases, numerical issues may also arise in estimating parameters λ_i and q_i. Focusing on modelling the human mortality schedule, the PH distribution could require a high dimension of the intensity matrix. For instance, Lin and Liu (2007) adopt the Coxian distribution to fit the mortality schedule from age 0 to age 100, suggesting that 200–250 ageing phases are needed. Cheng *et al.* (2020) impose specific structures on λ_i and q_i, finding that 100 phases are pretty reasonable in fitting mortality data. Conversely, Liu and Lin (2012) argue that if only part of the mortality schedule is relevant, for example, from age 65 onwards, a lower-dimensional PH distribution is enough to achieve an accurate fit. Such an approach could be suitable for our purposes, since we deal with ages falling within the post-retirement life of individuals. Further details concerning the Markov ageing model fitting will be provided in Section 4.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 1. *Notation, and related meaning, adopted for describing the Markov ageing model.*

Notation	Meaning
$(A_t, t \geq 0)$	Markov jump process describing ageing
$E = \{1, 2, \ldots, m\}$	Set of transient ageing states
$m+1$	Absorbing (death) state
$\mathcal{S} = E \cup \{m+1\}$	State space of the process
$\boldsymbol{\Sigma}$	Infinitesimal generator of of the process
$\boldsymbol{\Lambda}$	$m \times m$ transition intensity matrix
λ_i	Transition intensity from state i to $i+1$, $i \in E$
$\boldsymbol{q}$	$m \times 1$ column vector of transition intensities to the death state
q_i	Transition rate from ageing state i to the death state
$\boldsymbol{\pi}$	Initial distribution of the process
π_i	Probability that the ageing process starts in the state i, with $i \in E$
τ	Time until death (absorption time)
$\boldsymbol{P}(t) = (p_t^{i,j})_{i,j \in E}$	Transition probability matrix at time t
$p_t^{i,j}$	Transition probability from the state i to the state j

Denoting with $p_t^{i,j} = \mathbb{P}(A_t = j | A_0 = i)$ the time-homogeneous transition probability from the ageing state i to the ageing state j, the transition probability matrix, namely $\boldsymbol{P}(t) = \left(p_t^{i,j}\right)_{i,j \in E}$, satisfies the following Kolmogorov forward equation:

$$\frac{d}{dt}\boldsymbol{P}(t) = \boldsymbol{P}(t)\boldsymbol{\Lambda}, \tag{2.4}$$

with the initial condition $\boldsymbol{P}(0) = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. Equation (2.4) admits a unique solution given by

$$\boldsymbol{P}(t) = \exp(\boldsymbol{\Lambda} t), \tag{2.5}$$

where the exponential matrix is defined by series expansion, that is, $\exp(\boldsymbol{\Lambda} t) = \sum_{h=0}^{\infty} \frac{(t\boldsymbol{\Lambda})^h}{h!}$.

Under the PH distributions, many quantities of interest concerning the future lifetime are available in closed form. For instance, the survival function for a newborn can be written as:

$$S_0(t) = \boldsymbol{\pi}^\top \exp(\boldsymbol{\Lambda} t)\, \boldsymbol{e}, \quad t > 0. \tag{2.6}$$

Moreover, the conditional survival function given the survival at age $x > 0$ is

$$S_x(t) = \frac{S_0(x+t)}{S_0(x)} = \frac{\boldsymbol{\pi}^\top \exp(\boldsymbol{\Lambda}(x+t))\, \boldsymbol{e}}{\boldsymbol{\pi}^\top \exp(\boldsymbol{\Lambda} x)\, \boldsymbol{e}} = \boldsymbol{\pi}_x^\top \exp(\boldsymbol{\Lambda} t)\, \boldsymbol{e}, \quad t > 0, \tag{2.7}$$

where $\boldsymbol{\pi}_x^\top = \frac{\boldsymbol{\pi}^\top \exp(\boldsymbol{\Lambda} x)}{\boldsymbol{\pi}^\top \exp(\boldsymbol{\Lambda} x)\boldsymbol{e}}$ represents the distribution of the ageing phases, given the survival at age x.

Remark 2.1. In the Markov ageing model, once the process leaves an ageing state, it is impossible to return to it. In other words, the Markov ageing model keeps track of the ageing progression in the forward sense, and possible rejuvenations are not considered. Rejuvenation may be biologically possible only following various lifestyle interventions concerning nutrition, physical exercise and the social environment (see, e.g., Félix et al. 2024). Detailed information at the individual level is crucial to assessing the significance of possible rejuvenation transitions. As explained in Section 1, such data granularity is often unavailable to annuity providers, and rejuvenation transitions cannot be estimated.

To assist the reader, notation adopted in the current section is resumed in Table 1.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

3. The Markov ageing GSA arrangement

3.1 Pool structure

We consider a GSA scheme that provides post-retirement benefits to its members. Individuals pay an initial capital when joining the scheme; upon death, any residual amount is retained by the pool (i.e., no death benefit is included). In return for their initial contribution, members receive an annual lifetime income, whose amount is differentiated depending on the specific mortality profile of the individual. In fact, the GSA is structured with multiple levels of benefits, with higher benefit amounts provided to those with a shorter life expectancy. The individual's mortality profile is first assessed at entry and then reviewed annually. Upon a change of the mortality profile, for example, following the onset of a disease, the individual will be assigned the level of benefit consistent with his(her) new status. However, similarly to a traditional GSA, annual amounts for the various benefit levels are not guaranteed but are subject to annual adjustment based on the investment and longevity experience of the pool. In the following, we focus on longevity only, and we propose an approach to define such adjustments, which is motivated by actuarial arguments.

Mortality is modelled adopting the Markov ageing setting described in Section 2. We first assume that it is possible to classify individuals according to the ageing states and to differentiate benefits based on such states. However, as already pointed out in Section 1, biological age (which is at the root of ageing states) is not directly measurable. In the second step of the discussion (in particular, in Section 3.5), we assume that individuals are classified according to proxy risk factors. It is unlikely that risk classes identified this way correspond exactly to the ageing states of the Markov model. The implications are discussed in Section 3.5; for the moment, we assume that both at entry and at any time, a risk classification based on the ageing states of the Markov model can be performed. In particular, we assume that it is possible to describe the dynamics of each individual's mortality profile in terms of the transition to higher ageing states.

We consider a single cohort consisting of L_x individuals who join the scheme at time $t = 0$, at chronological age x, with L_x^i individuals placed in state i, $i \in E$ (clearly, $\sum_{i=1}^{m} L_x^i = L_x$). The maximum attainable age is ω for each individual; however, we assume that, to avoid the impact of major random fluctuations at later ages (when the cohort will shrink to low size), the scheme is committed to pay benefits up to time $t_{\max}$ (i.e., up to chronological age $x_{\max} = x + t_{\max} < \omega$). At time $t_{\max}$, the available funds will be transferred to an insurer willing to take charge of the surviving participants and their funds and guarantee them a lifetime annuity. The price and optimal time of such a transfer (which undoubtedly constitute interesting research questions) are not addressed in this paper, as our aim for now is to define acceptable benefit adjustment rules when multiple levels of benefits are involved; addressing also the transfer at time $t_{\max}$ would excessively broaden the scope of the present study. Thus, in this paper, we simply assume that the scheme will pay an income to survivors at times $t = 0, 1, \ldots, t_{\max}$, for a chosen (reasonable) $t_{\max}$, and the scheme's payout rate will not account for annuity costs concerning times $t > t_{\max}$.

Let B_0^i be the benefit set at time $t = 0$ for state $i \in E$. An individual entering the scheme in state i will initially get the benefit B_0^i; however, upon transition to a state $j > i$, $j \in E$, (s)he is entitled to cash the benefit B_0^j. Based on the usual actuarial arguments, the initial capital to be paid by an individual in state i at time $t = 0$ is then assessed as follows:

$$V_0^i = \sum_{j \geq i}^{m} B_0^j \, \ddot{a}_x^{i,j}, \tag{3.1}$$

where $\ddot{a}_x^{i,j} = \sum_{h=0}^{t_{\max}} (1+r)^{-h} {}_hp_x^{i,j}$ is the actuarial value of the annuity paid when in state j, computed using the transition probabilities ${}_hp_x^{i,j} = \exp(\mathbf{\Lambda} h)_{i,j}$ predicted by the mortality model and a given (deterministic) discount rate r. We point out that no loading is embedded in the choice of parameters of (3.1), as it is natural for a GSA scheme. We note also that the differentiation by state concerns not only the benefits but, consistently, the initial capital too.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

3.2 Pool funds

A GSA is grounded on the principle of funded liabilities. In other words, it is required that a balance be maintained at all times between available resources and the value of obligations. This is achieved by updating the benefit amount, upwards or downwards, depending on the experience with longevity and investments. We assume that the adjustment is applied annually, right before the benefit is paid at time t, $t \geq 1$. Let B_0^i be the benefit set at time $t = 0$ for state $i \in E$; reasonably, $B_0^1 < B_0^2 < \ldots < B_0^m$, since higher-indexed states correspond to lower life expectancy. Taking B_0^1 as the baseline benefit amount, it may be convenient to express the initial benefit amount for each state i, $i > 1$, as a proportion of B_0^1, as follows: $B_0^i = c^i B_0^1$, where the c^i's ($1 < c^2 < c^3 < \ldots < c^m$) are chosen according to some target. We let B_t^i denote the updated benefit amount at time t for state i, according to an adjustment rule which is described in Section 3.4. In the following, we understand $i \in E$ and $t = 0, 1, \ldots, t_{\max}$, unless otherwise stated (the adjustment of the benefit at time $t = 0$ is, of course, not under consideration; B_0^i is simply the starting chosen amount).

Following the adjustment of the benefit, the value of the obligation to an individual in state i at time t can be assessed as follows:

$$V_t^i = \sum_{j \geq i}^{m} B_t^j \, \ddot{a}_{x+t}^{i,j}, \tag{3.2}$$

with an obvious meaning of $\ddot{a}_{x+t}^{i,j}$ for $t > 0$. The total value of the obligations of the pool at time t is then

$$F_t^R = \sum_{i=1}^{m} L_{x+t}^i V_t^i, \tag{3.3}$$

where L_{x+t}^i is the (observed) number of individuals in state i at time t. We refer to F_t^R as the required pool fund at time t (in line with this, we will refer to V_t^i as the individual fund required in state i at time t).

The resources available to a GSA, which we call the available pool fund and denote as F_t^A at time t, are the money contributed by members, net of the payments made so far, plus interest. We focus on longevity risk only and assume a deterministic financial setting; we take r as the annual deterministic return on investments. Starting from $F_0^A = \sum_{i=1}^{m} L_x^i V_0^i$, the dynamics of the available pool fund can be described recursively for $t > 0$ as follows:

$$F_t^A = \left(F_{t-1}^A - \sum_{i=1}^{m} L_{x+t-1}^i B_{t-1}^i \right) (1 + r). \tag{3.4}$$

The basic requirement for running the GSA can now be written as:

$$F_t^R = F_t^A, \quad t \geq 0. \tag{3.5}$$

If benefits are undifferentiated (i.e., the same benefit amount is provided in all states), Equation (3.5) univocally defines the amount admitted for the benefit at time t (since, in such a case, there is only one unknown). When benefits are differentiated by state, Equation (3.5) admits an infinite number of solutions, as the number of unknowns is now m. The introduction of additional requirements is necessary to identify workable solutions. We point out that the benefit adjustment not necessarily will preserve the initial ratio between benefits, that is, not necessarily we find $B_t^i = c^i B_t^1$. If preserving such a ratio is a priority, then we need to impose $B_t^i = c^i B_t^1$ at all times t; in this case, Equation (3.5) has only one unknown, namely B_t^1. It would follow that, at any time $t > 0$, benefits would be differentiated by state, but they would all be adjusted in the same proportion.

At this point, it is convenient to examine the causes that may result in a mismatch between available and required pool funds. Disregarding financial risks, we may consider the following circumstances: mortality rates in the various states higher or lower than expected; ageing process (i.e., transition rates) higher or lower than expected. It should be taken into account that deviations between expected and observed rates may differ across states. Instead of adopting a common adjustment proportion, we may find it more appropriate to differentiate by state not only the benefits but also the adjustment coefficients.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

In the following, we investigate this solution. To this end, we split the available and required pool funds by state.

3.3 Splitting the pool funds by state

It is rather natural to identify

$$F_t^{i,R} = L_{x+t}^i \, V_t^i \tag{3.6}$$

as the pool fund required at time t for state i. Clearly, $\sum_{i=1}^m F_t^{i,R} = F_t^R$.

Let $F_t^{i,A}$ denote the pool fund made available to state i at time t, such that $\sum_{i=1}^m F_t^{i,A} = F_t^A$.

We now require that the principle of funded liabilities is applied by state, that is, we impose

$$F_t^{i,R} = F_t^{i,A}, \quad t \geq 0, \quad i \in E. \tag{3.7}$$

Such a requirement clearly also satisfies (3.5). To proceed further with (3.7), it is useful to note that the amount of resources available per state is affected by cross-subsidy effects between states. Part of these cross-subsidies are embedded in the individual fund required (and available) in state i at time $t-1$. To check this, we start by considering that

$$\ddot{a}_{x+t-1}^{i,i} = 1 + (1+r)^{-1} \, p_{x+t-1}^{i,i} \, \ddot{a}_{x+t}^{i,i},$$

and

$$\ddot{a}_{x+t-1}^{i,j} = (1+r)^{-1} \sum_{i \leq k \leq j} p_{x+t-1}^{i,k} \, \ddot{a}_{x+t}^{k,j}, \quad j > i.$$

Then, we can write the following recursion for the individual fund in state i:

$$V_{t-1}^i = B_{t-1}^i + \sum_{j \geq i} p_{x+t-1}^{i,j} \, V_{t^-}^j \, (1+r)^{-1}, \tag{3.8}$$

where

$$V_{t^-}^j = \sum_{k \geq j}^m B_{t-1}^k \, \ddot{a}_{x+t}^{j,k}$$

represents the amount of individual fund required in state j at time t, before the adjustment of the benefits at that time.

Let

$$F_{t^-}^{i,A} = \left(F_{t-1}^{i,A} - L_{x+t-1}^i B_{t-1}^i\right)(1+r) \tag{3.9}$$

denote the pool fund dragged to state i at time t, $t > 0$, by the pool fund available in such a state at time $t-1$, net of the benefits paid at time $t-1$ and credited with interest. Comparing Equation (3.9) with (3.4) it can be easily seen that $\sum_{i=1}^m F_{t^-}^{i,A} = F_t^A$.

Since because of (3.7) we have $F_{t-1}^{i,A} = F_{t-1}^{i,R}$, using (3.8) we can rearrange (3.9) as:

$$F_{t^-}^{i,A} = L_{x+t-1}^i \, p_{x+t-1}^{i,i} \, V_{t^-}^i + L_{x+t-1}^i \sum_{j>i} p_{x+t-1}^{i,j} \, V_{t^-}^j, \tag{3.10}$$

that we can interpret as follows. Given L_{x+t-1}^i, the quantity $L_{x+t-1}^i \, p_{x+t-1}^{i,i}$ represents the number of individuals expected to stay in state i during the year, while $L_{x+t-1}^i \, p_{x+t-1}^{i,j}$ represents the number of individuals expected to be in state j at the end of the year. Equation (3.10) then shows that $F_{t^-}^{i,A}$, that is, $F_{t-1}^{i,A}$ net of the benefits paid at the beginning of the year in state i and credited with interest, consists of the resources required for those expected to remain alive, either in state i or in a higher state $j, j > i$ (we point out that for $j = m+1$, that is, for the dead state, $V_{t^-}^j = 0$).

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Since $p^{i,i}_{x+t-1} = 1 - \sum_{j>i}^{m+1} p^{i,j}_{x+t-1}$, Equation (3.10) can be further rearranged as follows:

$$F^{i,A}_{t^-} = L^i_{x+t-1} \left(1 - p^{i,m+1}_{x+t-1}\right) V^i_{t^-} + L^i_{x+t-1} \sum_{j>i}^{m} p^{i,j}_{x+t-1} \left(V^j_{t^-} - V^i_{t^-}\right). \tag{3.11}$$

Equation (3.11) allows us to give a further useful interpretation to $F^{i,A}_{t^-}$. Since $L^i_{x+t-1}\, p^{i,m+1}_{x+t-1}$ is the expected number of deaths reported by those in state i at time $t-1$, the amount $L^i_{x+t-1}\, p^{i,m+1}_{x+t-1}\, V^i_{t^-}$ can be related to the expected mortality credits in such a state. Conversely, $L^i_{x+t-1} \sum_{j>i}^{m} p^{i,j}_{x+t-1}$ is the number of individuals expected to move from state i to a living state $j, j > i$, and $L^i_{x+t-1} \sum_{j>i} p^{i,j}_{x+t-1} \left(V^i_{t^-} - V^j_{t^-}\right)$ is a measure of the (total) net cross-subsidy 'cost' charged to state i for those expected to be alive, but in higher ageing states.

The annual dynamics of the population of state i can be described for $t > 0$ as follows:

$$L^i_{x+t} = L^i_{x+t-1} + \sum_{j<i} L^{j,i}_{x+t-1} - \sum_{j>i}^{m+1} L^{i,j}_{x+t-1}, \tag{3.12}$$

where we use the notation $L^{j,i}_{x+t-1}$ (or $L^{i,j}_{x+t-1}$) to denote the observed number of individuals in state i (j) at time t being in state j (i) at time $t-1$. Equation (3.12) makes it clear that the number of individuals in state i at time t is the result of those who left state i (moving to a living state j, $i < j \leq m$, or due to death, $j = m+1$) and those who have entered it from a state $j < i$. Clearly, $L^{j,i}_{x+t-1} = 0$ for $i = 1$. Note that situations such as $L^i_{x+t-1} = 0$ and $L^i_{x+t} > 0$ or $L^i_{x+t-1} > 0$ and $L^i_{x+t} = 0$ are possible, as a result of transitions into and out of state i. In the following, we use the notation $\widetilde{p}$ to denote the realised (observed) transition or permanence rates; for instance, $\widetilde{p}^{\,i,j}_{x+t-1} = \frac{L^{i,j}_{x+t-1}}{L^i_{x+t-1}}$. Such rates are obviously defined for states that are non-empty at time $t-1$, that is, for states i such that $L^i_{x+t-1} > 0$. Using this notation, Equation (3.12) can be rewritten as:

$$L^i_{x+t} = L^i_{x+t-1} \left(1 - \widetilde{p}^{\,i,m+1}_{x+t-1}\right) + \sum_{j<i} L^j_{x+t-1}\, \widetilde{p}^{\,j,i}_{x+t-1} - \sum_{j>i}^{m} L^i_{x+t-1}\, \widetilde{p}^{\,i,j}_{x+t-1}. \tag{3.13}$$

Let us now compare $F^{i,A}_{t^-}$ to the pool fund required at time t for state i. As stated by (3.6), such a fund is proportional to L^i_{x+t}. Its amount before the adjustment of the benefits is represented by:

$$F^{i,R}_{t^-} = L^i_{x+t}\, V^i_{t^-}. \tag{3.14}$$

If we plug (3.13) into (3.14), we obtain the expression:

$$F^{i,R}_{t^-} = L^i_{x+t-1} \left(1 - \widetilde{p}^{\,i,m+1}_{x+t-1}\right) V^i_{t^-} + \sum_{j<i} L^j_{x+t-1}\, \widetilde{p}^{\,j,i}_{x+t-1}\, V^i_{t^-} - \sum_{j>i}^{m} L^i_{x+t-1}\, \widetilde{p}^{\,i,j}_{x+t-1}\, V^i_{t^-}, \tag{3.15}$$

that we can compare with (3.11) for an interpretation. It is unlikely that $F^{i,A}_{t^-} = F^{i,R}_{t^-}$, for a number of reasons:

- The realised number of deaths, $L^i_{x+t-1}\, \widetilde{p}^{\,i,m+1}_{x+t-1}$, can be other than what expected, $L^i_{x+t-1}\, p^{i,m+1}_{x+t-1}$, and then the realised mortality credits in state i can be higher or lower than expected. We may consider it appropriate for any discrepancy to be absorbed by state i, given that this mismatch arises exclusively from mortality in state i.
- In face of individuals entering state i during the year, it is required to hold in this state an additional amount $\sum_{j<i} L^j_{x+t-1}\, \widetilde{p}^{\,j,i}_{x+t-1}\, V^i_{t^-}$, which is not already included in $F^{i,A}_{t^-}$. This is an amount that should be cross-subsidised by the previous states $j, j < i$.
- Individual funds $V^j_{t^-}$ for those moving to a living state $j, j > i$ are no longer required in state i and should be taken out from $F^{i,A}_{t^-}$, to cross-subsidise the next states. This is in line with what was discussed with regard to those entering state i (previous point).

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

- Finally, we note that when $L^i_{x+t} = 0$, then $F^{i,R}_{t^-} = 0$; it can happen, instead, that $F^{i,A}_{t^-} > 0$, given that this latter fund is assessed based on expected number of transitions, in contrast to $F^{i,R}_{t^-}$ which considers realised transitions. In such a circumstance, $F^{i,A}_{t^-}$ should be released to non-empty states.

Based on the above considerations, we define the pool fund available for state i at time t as follows:

$$F^{i,A}_t = \begin{cases} F^{i,A}_{t^-} + \sum_{j<i} L^j_{x+t-1}\, \widetilde{p}^{j,i}_{x+t-1}\, V^i_{t^-} - \sum_{j>i}^{m} L^i_{x+t-1}\, \widetilde{p}^{i,j}_{x+t-1}\, V^j_{t^-} + X^i_t & \text{if } L^i_{x+t} > 0; \\ 0 & \text{if } L^i_{x+t} = 0. \end{cases} \tag{3.16}$$

With X^i_t we denote the amount of resources left at time t by empty states, released to state i, so to ensure that $\sum_i F^{i,A}_t = F^A_t$. We assume a redistribution proportional to the required funds before the adjustment of the benefits; more precisely, we assume (for non-empty states i):

$$X^i_t = \left(F^A_t - \sum_{i:\, L^i_{x+t}>0} F^{i,A}_{t^-} \right) \frac{F^{i,R}_{t^-}}{F^R_{t^-}}.$$

As already noted, definition (3.16) ensures $\sum_i F^{i,A}_t = F^A_t$. The amounts $\sum_{j<i} L^j_{x+t-1}\, \widetilde{p}^{j,i}_{x+t-1}\, V^i_{t^-}$ and $\sum_{j>i}^{m} L^i_{x+t-1}\, \widetilde{p}^{i,j}_{x+t-1}\, V^j_{t^-}$ measure the cross-subsidy, respectively, in favour of and released by state i. These amounts are also present in F^A_t because, when we deal with multiple states, necessarily cross-subsidy effects between states are involved; however, such cross-subsidies are not explicit in (3.4) because they sum up to zero at the pool level.

3.4 Adjustment of the benefits by state

Once the pool fund available for state i at time t, that is $F^{i,A}_t$, is defined, we proceed to the adjustment of benefits imposing (3.7). We point out that this is a set of equations that must be solved backwards, from state m down to state 1. In particular, for $i = m$, we have

$$L^m_{x+t} B^m_t \ddot{a}^{m,m}_{x+t} = F^{m,A}_t,$$

and the unknown B^m_t straightforwardly follows:

$$B^m_t = \frac{F^{m,A}_t}{L^m_{x+t}} \frac{1}{\ddot{a}^{m,m}_{x+t}}.$$

For $i = m-1$, by imposing

$$L^m_{x+t} \left(B^{m-1}_t \ddot{a}^{m-1,m-1}_{x+t} + B^m_t \ddot{a}^{m-1,m}_{x+t} \right) = F^{m-1,A}_t$$

the unknown is B^{m-1}_t, being B^m_t previously determined, and it holds that:

$$B^{m-1}_t = \left(\frac{F^{m-1,A}_t}{L^{m-1}_{x+t}} - B^m_t \ddot{a}^{m-1,m}_{x+t} \right) \frac{1}{\ddot{a}^{m-1,m-1}_{x+t}}.$$

Proceeding iteratively, for the non-empty states, we find

$$B^i_t = \left(\frac{F^{i,A}_t}{L^i_{x+t}} - \sum_{j>i} B^j_t \ddot{a}^{i,j}_{x+t} \right) \frac{1}{\ddot{a}^{i,j}_{x+t}}, \tag{3.17}$$

while we assume that if $L^i_{x+t} = 0$, then $B^i_t = B^i_{t-1}$ (i.e., no adjustment is applied to benefits relating to empty states).

An interesting interpretation of the adjustment can be found in terms of individual fund V^i_t. For brevity, we develop it under the assumption $L^i_{x+t-1} > 0$ and $X^i_t = 0$ for all states i.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

If we equate (3.6)–(3.16), and rearrange, we get

$$V_t^i \left(L_{x+t}^i - \sum_{j<i} L_{x+t-1}^j \, \widetilde{p}_{x+t-1}^{j,i} \right) = L_{x+t-1}^i \, p_{x+t-1}^{i,i} \, V_{t^-}^i + \sum_{j<i} L_{x+t-1}^j \, \widetilde{p}_{x+t-1}^{j,i} \, (V_{t^-}^i - V_t^i) - L_{x+t-1}^i \sum_{j>i}^{m} \left(\widetilde{p}_{x+t-1}^{i,j} - p_{x+t-1}^{i,j} \right) V_{t^-}^j. \tag{3.18}$$

Since (3.13) can also be written as

$$L_{x+t}^i = L_{x+t-1}^i \, \widetilde{p}_{x+t-1}^{i,i} + \sum_{j<i} L_{x+t-1}^j \, \widetilde{p}_{x+t-1}^{j,i},$$

from (3.18) we obtain

$$V_t^i = \frac{p_{x+t-1}^{i,i}}{\widetilde{p}_{x+t-1}^{i,i}} \, V_{t^-}^i - \frac{\sum_{j<i} L_{x+t-1}^j \, \widetilde{p}_{x+t-1}^{j,i}}{L_{x+t-1}^i \, \widetilde{p}_{x+t-1}^{i,i}} \, (V_t^i - V_{t^-}^i) - \frac{\sum_{j>i}^{m} \left(\widetilde{p}_{x+t-1}^{i,j} - p_{x+t-1}^{i,j} \right)}{\widetilde{p}_{x+t-1}^{i,i}} \, V_{t^-}^j. \tag{3.19}$$

When adjusting the benefits, clearly, we also need to adjust the individual required funds. Equation (3.19) suggests the following interpretation. The individual fund required at time t for state i must be adjusted because:

- The proportion of individuals remaining in state i differs from what was expected (see the first term of Equation (3.19));
- The cross-subsidy received for individuals entering state i from lower states $j, j < i$ is based on $V_{t^-}^i$, but it must be updated to V_t^i after the adjustment (see the second term of Equation (3.19));
- The cross-subsidy transferred to higher states $j, j > i$ is based on the realised transition rates, whereas the amount set aside in $F_{t^-}^{i,A}$ is based on the expected transition rates (see the third term of Equation (3.19)).

We would like to emphasise that the first term recalls the mortality experience adjustment (MEA) factor introduced in Piggott *et al.* (2005). It expresses the benefit update implied by the discrepancy between expected and realised mortality. In our multi-state framework, such an adjustment applies state-specific, accounting for ageing and mortality deviations. The second and third terms in Equation (3.19) correct the state-specific adjustment by accounting for, respectively, cross-subsidy inflows and outflows. Naturally, cross-subsidy cash flows can be defined in alternative ways; for instance, their amounts may be shaped based on the expected transition rates rather than realised values. Beyond the possible choices and consequent algebraic expressions, individual funds and benefits must be updated when experience differs from expectations, that is, realised cross-subsidies differ from those expected.

3.5 Smoothing benefits across states

The reasons that may lead an individual to prefer participation in a GSA scheme over the purchase of a life annuity are extensively discussed in the literature (see also Section 1); above all, cost savings, albeit at the expense of guarantees. When dealing with a multi-state GSA, one should additionally wonder why an individual should choose it, rather than one offering identical benefits across all states.

Similarly to a standard annuity, an individual who is aware of being in substandard health or maintaining unhealthy lifestyle patterns might feel disadvantaged by the use of undifferentiated annuity rates (see, e.g., Olivieri and Tabakova 2024 for the case of annuities and Kabuche *et al.* 2024 for the case of a GSA). On the other hand, an individual currently in good health may feel better protected by a scheme that guarantees an upgrade in the benefit in the event of a deterioration in their health status. In this

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

preference, one may recognise an interest akin to that observed for long-term care insurance products or disability income annuities.

Although GSA members are expected to accept benefit fluctuations across states depending on experience, they may reasonably expect a consistent and coherent ordering of those benefits, that is, $B_t^1 < B_t^2 < \ldots < B_t^m$ at all times $t \geq 0$. At time $t = 0$, this is ensured by the initial choice of the benefit amounts; however, Equation (3.17) does not necessarily maintain such an ordering at later times ($t > 0$), as it does not explicitly consider the relative differences among benefit levels. We also note that, even when the ordering of benefits is preserved, the distance between them may fall below desirable thresholds. These considerations suggest the introduction of corrective mechanisms to the adjustment rule described in Section 3.4.

There is a further circumstance that may require modifications to the adjustment rule. As we have discussed earlier (see Sections 1, 2 and 3.1), ageing states are usually not observable. The current mortality profile of an individual can reasonably be detected based on proxy risk factors, which may lead to a classification of individuals that does not precisely replicate the intended ageing states. Moreover, either by design or in connection with the available proxy risk factors, it could be considered more appropriate to structure benefits across m^* classes, not necessarily matching the number of ageing states, that is, $m^* \leq m$. In settings where m is large, adopting fewer benefit classes, $m^* < m$, may offer a more practicable implementation. Illustrative proxy risk factors include medical events like the development of diabetes or a heart attack, as well as milder but indicative conditions such as hypertension, elevated glucose or cholesterol levels, or macular disease. They may also encompass functional or cognitive impairments, such as difficulties in performing activities of daily living or the onset of dementia. It should be noted that, given the absence of guarantees, the scheme may allow for a classification based on medical conditions or diseases that are typically excluded from standard insurance coverage. However, since more severe conditions result in access to a higher benefit class, safeguards must be anyhow in place to avoid potential misuse or fraud.

In the following, we still refer to E as the set of ageing states and to $E^* = \{1, 2, \ldots, m^*\}$ as the set of benefit classes. Benefits B_t^i assessed as in Equation (3.17) ensure the actuarial equivalence between benefits and available resources by ageing state and for the whole pool, as discussed in Section 3.3. In the following, we refer to them as the notional benefits (or benefits by ageing state). We note that

$$B_t^{TOT} = \sum_{i \in E} L_{x+t}^i \, B_t^i \tag{3.20}$$

measures the total payout admitted by the actuarial equivalence condition (3.7) at time t, $t \geq 0$. We denote with $\overline{B}_t^k$ the benefit for class $k \in E^*$, and to keep the actuarial equivalence at the pool level, we require

$$\sum_{k \in E^*} l_{x+t}^k \, \overline{B}_t^k = B_t^{TOT}, \tag{3.21}$$

where l_{x+t}^k is the number of survivors placed in benefit class k at time t (note that $\sum_{i \in E} L_{x+t}^i = \sum_{k \in E^*} l_{x+t}^k$). We refer to $\overline{B}_t^k$ also as the actual benefits. Based on the motivations mentioned above, we introduce the further requirement:

$$\overline{B}_t^k = u^k \, \overline{B}_t^1, \tag{3.22}$$

where $\overline{B}_t^1$ is the benefit amount for the lowest benefit class, and u^k, $k = 2, 3, \ldots, m^*$, is a coefficient chosen to achieve the desired differential between the benefit in class k and class 1. Reasonably, $1 < u^2 < u^3 < \ldots < u^{m^*}$. To avoid potential misunderstandings when comparing the coefficients c^i's introduced in Section 3.2 and the coefficients u^k's, we stress that, by design, u^k measures the ratio at any time t between the actual benefit in class k in respect of class 1. The coefficient c^i, instead, describes the initial level of differentiation across ageing-state benefits at time 0; at the following times t, $t = 1, 2, \ldots$, it may turn out $B_t^i \gtreqless c^i B_t^1$, as a result of the different benefit dynamics by ageing state. Based on (3.22), Equation (3.21) admits only one solution. After noting that requirement (3.22) introduces a kind of smoothing

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

of benefits, we point out that benefits subjected to (3.22) bring us back to the solution discussed in Section 3.2 after Equation (3.5). Indeed, they may differ from that solution because of the number m^* of benefit classes (and the specific choice of the proportion among benefits), but basically, they follow the same idea. Yet, the analysis conducted in Sections 3.3 and 3.4 is not redundant, as it enables us to assess the extent of solidarity entailed by this choice.

As already noted, while actual benefits ensure actuarial equivalence at the pool level, the actuarial equivalence at the ageing state level will clearly be lost. This means that by adopting benefits $\overline{B}_t^k$'s instead of B_t^i's, solidarity mechanisms are introduced in the benefit design. With reference to an individual in ageing state i, placed in benefit class k, the size and direction of solidarity can be understood by comparing the actual benefit $\overline{B}_t^k$ with the notional benefit B_t^i. Clearly, for some individuals $\overline{B}_t^k < B_t^i$, meaning that they are contributing solidarity, whereas for others $\overline{B}_t^k > B_t^i$, that is, they are receiving solidarity. We underline that this constitutes a design choice, albeit one that, in specific contexts, is imposed by the limited information available on individual mortality patterns. As pointed out in Section 1, understanding the solidarity effects introduced enables a rational justification and communication of design decisions, which is particularly important in arrangements where participants are exposed to the full range of shared risks. A numerical assessment is discussed in Section 4.2, which also clarifies why the introduction of solidarity effects may prove to be recommended.

4. Numerical application

In the following, we delve into the proposed theoretical framework through a numerical application.

Mortality rates provided by the Human Mortality Database are used to fit the Markov ageing model, and different mortality experiences are selected to test the robustness of the fitting procedure. In particular, we refer to mortality rates of the 1920, 1930 and 1940 Australian cohorts, for both genders, considering ages 65 and above. We denote by $m_{x,c}$ the observed mortality rate at age x for the cohort c, with $c = 1920, 1930, 1940$. The age sets we deal with are $\mathcal{X}_{1920} = \{65, 66, \ldots, 100\}$, $\mathcal{X}_{1930} = \{65, 66, \ldots, 90\}$ and $\mathcal{X}_{1940} = \{65, 66, \ldots, 80\}$ for the 1920, 1930 and 1940 cohorts, respectively.

In Section 4.1, the procedure adopted for fitting the Markov ageing model will be discussed, as well as the strategy employed to simulate transitions between the process's states. The numerical implementation of the Markov ageing GSA will be exposed in Section 4.2. For the sake of conciseness, some numerical results are collected in the Supplementary Material accompanying the present manuscript.

4.1 Model fitting and simulation

To fit the Markov ageing model, both the initial distribution $\boldsymbol{\pi}$ and the intensity matrix $\boldsymbol{\Lambda}$ of the ageing process must be estimated, and the choice of fitting method may be affected by the data availability. At first glance, the maximum-likelihood estimation (MLE) approach represents a natural candidate for fitting the Markov ageing model. To perform the MLE procedure, independent trajectories of the ageing process must be completely observed up to the time of death. Each path should consist of the sequence $\{(i_k, s_k)\,,\ k = 1, 2, \ldots, n\}$, where $i_1, i_2, \ldots, i_n$ are the ageing states visited before death, and $s_1, s_2, \ldots, s_n$ are the corresponding sojourn times. However, due to the latent nature of the ageing states, trajectories of the ageing process are not detectable, and the likelihood function maximisation cannot be directly executed. Nonetheless, if realisations of i.i.d. times until death are accessible, the MLE can be carried out via the expectation-maximisation (EM) algorithm. Assuming that an underlying ageing process generates such absorption times, the conditional expected value of the log-likelihood function can be computed and then maximised w.r.t. model parameters. In case no information on the times until death is available, they could be randomly drawn from a valid PH distribution. The simulation exercise requires that several combinations of the initial probability vector and the intensity matrix must be tried, as well as different numbers of EM iterations should be considered. It could be computationally expensive, and

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

there is no guarantee that the maximum of the likelihood function is achieved. Indeed, while the EM algorithm always converges, it does not necessarily converge to the ML estimator as it can be trapped in a local maximum. The EM algorithm can be exploited to its full effectiveness, for instance, when survival data are observable; however, this is not the case for the present analysis because only cohort mortality rates are available to us.

Given the above considerations, we opt for an alternative fitting method to MLE. For the sake of completeness, we refer the reader to the Section 1 of the Supplementary Material, where technical details on the EM-based fitting of the Markov ageing model are provided.

We proceed to fit the Markov ageing model by adopting a target approach, that is, a general procedure allowing for dealing directly with the observed mortality or survival curve. Let $\boldsymbol{\theta}_c = (\boldsymbol{\pi}_c, \boldsymbol{\Lambda}_c)$ be the parameter characterising the Markov ageing model for the cohort c. For each cohort, the deviations between the observed mortality data (the target) and the corresponding value provided by the ageing model are measured through a proper loss function $\mathcal{L}$. We choose the mean squared error as the loss function, while our target is the t-year survival probability at age 65. Then, we have

$$\mathcal{L}(\boldsymbol{\theta}_c) = \frac{1}{|\mathcal{X}_c|} \sum_{t=1}^{|\mathcal{X}_c|} \left({}_tp_{65,c} - S_{65}\,(t; \boldsymbol{\theta}_c) \right)^2, \tag{4.1}$$

where ${}_tp_{65,c}$ is the t-year survival probability at age 65 for the cohort c and computed as ${}_tp_{65,c} = \prod_{h=0}^{t-1} \left(1 - \frac{m_{65+h,c}}{1+0.5\, m_{65+h,c}}\right)$, while $S_{65}\,(t; \boldsymbol{\theta}_c)$ is the survival function for the same cohort and its expression is given in Equation (2.7). We find the estimate $\hat{\boldsymbol{\theta}}_c$ as the solution of the following constrained mean squared error minimisation problem:

$$\min_{\boldsymbol{\theta}_c} \mathcal{L}(\boldsymbol{\theta}_c), \quad \text{s.t.} \quad \pi_{i,c} \in [0,1]\ \forall\, i \in E,\ \sum_{i \in E} \pi_{i,c} = 1, \tag{4.2}$$

which we solve numerically for each cohort and gender.

As mentioned in Section 2, two other aspects deserve consideration in fitting the Markov ageing model: the number of ageing phases, m, and possible non-uniqueness problems in parameter specifications. Concerning the latter, along the lines of Su and Sherris (2012), we impose the following specification for intensity matrix elements in (2.3):

$$\lambda_i = \lambda \quad \text{and} \quad q_i = \gamma + \alpha\, \mathrm{e}^{\beta\, i}, \quad \forall\, i \in E,\ \gamma, \alpha, \beta > 0. \tag{4.3}$$

A constant transition intensity between states means that, on average, a linear decline of various physiological functions is acknowledged. The transition towards the death state accounts for a background rate, namely γ, and an ageing-dependent component, $\alpha\, \mathrm{e}^{\beta\, i}$. The former allows for the representation of ageing termination due to accidental death, while the latter describes an exponential increase in mortality as ageing proceeds. Regarding the number of ageing phases, an accurate fit of the entire mortality schedule could be attained for high values of m (see, e.g., Lin and Liu 2007; Cheng *et al.* 2020); however, if only part of the mortality schedule is of interest, a low-dimensional PH distribution furnishes a satisfactory fit (see, e.g., Liu and Lin 2012). Therefore, we select the number of phases empirically, starting with low m values and increasing them until the loss function reduces. In Table 2, mean squared error values are reported for different numbers of ageing phases, distinguishing by cohorts and genders. For the 1920 and 1930 Australian cohorts, we find that $m = 15$ ageing phases grant the best model fitting, for both genders; the same occurs for the 1940 cohorts with $m = 7$ ageing phases. Figure 2 furnishes a graphical visualisation for the comparison between the observed t-years survival probabilities and the corresponding fitted values of the PH's survival function. The latter are determined according to the m value providing the lowest mean squared error for each cohort and gender. We observe a satisfactory fit across all cohorts and both genders. The greater fitting accuracy is achieved for the younger cohorts, since their survival curve consists of a few observations outlining a non-complex trend over time.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 2. *Mean squared error values by increasing the number of ageing phases, m, and distinguishing by cohorts and genders. The lowest mean squared error values are in bold.*

Cohort	1920		1930		1940	
m	Female	Male	Female	Male	Female	Male
3	1.882084×10^{-2}	1.32322×10^{-3}	7.5310×10^{-4}	5.5678×10^{-4}	4.68×10^{-6}	1.505×10^{-5}
5	1.47505×10^{-3}	4.9801×10^{-4}	2.3156×10^{-4}	1.5141×10^{-4}	2.8×10^{-7}	7.1×10^{-7}
7	7.9133×10^{-4}	2.5435×10^{-4}	1.0336×10^{-4}	5.016×10^{-5}	$\mathbf{1.1 \times 10^{-7}}$	$\mathbf{1.5 \times 10^{-7}}$
10	4.0559×10^{-4}	1.0251×10^{-4}	2.149×10^{-5}	2.285×10^{-5}	7.9×10^{-7}	1.8×10^{-7}
15	$\mathbf{2.8797 \times 10^{-4}}$	$\mathbf{3.364 \times 10^{-5}}$	$\mathbf{1.421 \times 10^{-5}}$	$\mathbf{3.13 \times 10^{-6}}$	4.0×10^{-7}	5.0×10^{-7}
20	4.8811×10^{-4}	4.1846×10^{-4}	7.665×10^{-5}	2.1845×10^{-4}	1.9×10^{-7}	2.0×10^{-7}

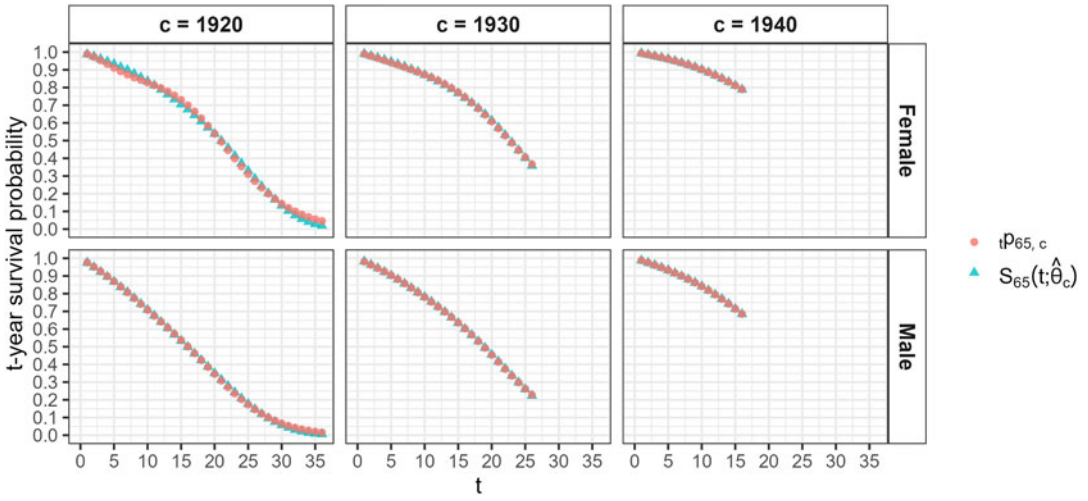


Figure 2. *Comparison between the observed t-years survival probabilities (red dots) and the corresponding fitted values of the PH's survival function (blue triangles). The latter are determined according to the m value providing the lowest mean squared error for each cohort and gender (see Table 2).*

For the sake of simplicity, from now on, we focus on the 5-state Markov ageing model for the 1920 Australian cohort, considering both genders. Using a low-dimensional Markov ageing model facilitates the readability of both the estimated parameters and the results of the Markov ageing GSA presented in Section 4.2. Comments we draw in the following discussions can be easily extended to high-dimensional Markov ageing models, as well as to the other cohorts. We refer the reader to Section 2 of the Supplementary Material for the parameters estimate concerning the befitting Markov ageing models.

Table 3 collects the 5-state Markov ageing mode's parameters estimate concerning the 1920 Australian cohort. We note that, for both genders, the initial distribution of the process is mainly concentrated in the first ageing state, and it is residually located in the other states, in particular the last one. Then, 65-year-old individuals cannot be attributed to a unique ageing status; individuals differ in physiological state, even though they have the same chronological age. For the male gender, the probability of being in the first ageing state is lower than for the female gender; the opposite occurs if we consider the fifth state. Therefore, the physiological status of females chronologically aged 65 should be better than that of males. Such evidence is accompanied by the gender-based estimates of λ. Females present a λ value lower than that for males, indicating a slower decline in their physiological functions. Considering the parameters embedded in the transition intensity to the death state, the

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 3. *Estimate of the 5-state Markov ageing model's parameters for the 1920 Australian cohort, both genders. We recall that: α is an ageing-related scale parameter, β denotes the rate of change of q_i as ageing progresses, γ is the rate measuring the ageing termination due to accidental deaths, λ is the transition intensity among states, and π_i is the probability that the process starts in the state i, for $i = 1, \ldots, 5$.*

Parameter	Females	Males
α	7.0115×10^{-6}	3.9205×10^{-4}
β	2.076213	1.298222
γ	1.0581×10^{-11}	2.3084×10^{-21}
λ	0.2117737	0.2260998
π_1	0.9438316	0.8677341
π_2	3.3452×10^{-8}	1.7088×10^{-7}
π_3	5.2938×10^{-8}	1.762×10^{-7}
π_4	1.9091×10^{-7}	1.7961×10^{-6}
π_5	0.05616817	0.1322637

background rate estimate appears negligible for both genders. Indeed, the background rate γ measures the effects of accidental deaths, which mainly occur at adult ages, while we are referring to older ages. The β parameter represents the rate at which the transition intensity to the death state varies as the ageing phases progress. The larger the value of β, the more likely the transition to the death state will be. As shown in Table 3, the female gender has a higher β than that for the male gender, but at the same time, there is a compensation with the respective α parameter values. On the whole, we find that the female gender tends to experience a slower ageing process and a longer lifespan, compared to the male gender.

Following the model fitting, the evolution over time and by state of the initial population can be simulated. Similarly to Kabuche *et al.* (2024), the realised number of individuals transitioning from state i to state j in the interval $(t-1, t]$ is randomly drawn from a multinomial distribution, that is,

$$L^{i,j}_{x+t-1} \sim \text{Multinom}\left(L^{i}_{x+t-1}, p^{i,j}_{x+t-1}\right), \quad i \in E,\ j \in \mathcal{S}. \tag{4.4}$$

In (4.4), L^{i}_{x+t-1} refers to the actual number of individuals belonging to the ageing state i at time $t-1$, while $p^{i,j}_{x+t-1}$ is the expected 1-year transition probability, from state i to state j, stemming from the calibrated Markov ageing model. At each time step, we run 10,000 simulations to generate population transitions among states. The total number of individuals attaining state j at time t is used for generating population transitions at time $t+1$. In addition, for each random draw, the simulated number $L^{i,j}_{x+t-1}$ is used to compute the realised 1-year transition probabilities as $\widetilde{p}^{i,j}_{x+t-1} = \frac{L^{i,j}_{x+t-1}}{L^{i}_{x+t-1}}$.

4.2 Implementation of the Markov ageing GSA

4.2.1 Pool composition and population dynamics

We assume that at time $t = 0$ the GSA scheme is joined by a cohort of $L_x = 10,000$ individuals (either females or males), chronological age $x = 65$; we further assume that the scheme closes upon reaching time $t_{\max} = 30$. Consistent with the Markov ageing model calibrated as described in Section 4.1, individuals are distributed in five ageing states, with initial composition as in Table 4. We point out that such a composition (for both genders) naturally follows from the estimated initial distribution of the Markov process; see Table 3.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 4. *Initial pool composition, by gender.*

	L^1_{65}	L^2_{65}	L^3_{65}	L^4_{65}	L^5_{65}
Females	9438	0	0	0	562
Males	8677	0	0	0	1323

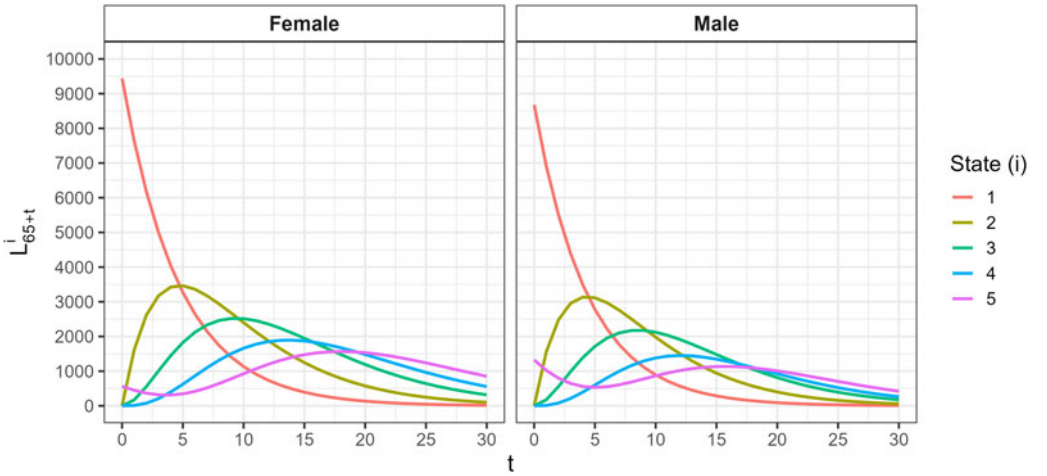


Figure 3. *Population dynamics (average value) by time (annual basis) and state for both genders. We refer the reader to the Section 3 of the Supplementary Material for the graphical representation of the population transitions among states, for both genders.*

Figure 3 plots the average number of survivors in the five ageing states across time, separately for females and males; this way, we have a representation of the population dynamics. The average paths of the five ageing states are similar for both genders: the decreasing pattern of state 1 is explained by the fact that individuals can only exit this state (either by jumping to a higher living state or because of death). The non-monotonic trajectories of the other states are driven by the net balance of transitions in and out of each state. States representing higher ageing levels reach their maximum size later than the others, and the peak population tends to decrease with increasing state severity. Nevertheless, at older chronological ages, and especially at the final time $t_{\max}$, the higher ageing states are larger than the lower ones. In particular, state 1 exhibits a population size that almost disappears at later times, although for both genders, it does not die out by time $t_{\max}$. When comparing these patterns between females and males, the greater average longevity of the former results in higher average sizes of the lower-index ageing states for females, at any given time.

4.2.2 Available pool fund by state

The GSA scheme pays annual benefits to survivors at the beginning of the year, from time $t = 0$ until time $t_{\max} = 30$. We set $B^1_0 = 1.00$ account unit and $B^i_0 = 1.2^{i-1}\, B^1_0$; Table 5 lists the initial benefit amounts by ageing state and the corresponding initial capital, assessed according to Equation (3.1). The discount rate is set to $r = 0.02$. For each gender, the higher initial capital required from members in ageing state 1 reflects their longer life expectancy, notwithstanding their lower initial benefit. Similarly, for a given state, higher initial capitals are charged to females as a consequence of their higher expected lifetime.

Figure 4 shows the time profile of the average values of the pool fund available by state for females. The time profiles for males are similar and, for brevity, are not shown here, but they can be found in the Section 3 of the Supplementary Material.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 5. *Initial notional benefit amount and capital, by gender.*

		Notional initial capital V_0^i	
Ageing state i	Notional benefit amount B_0^i	Females	Males
1	1.00	23.5532	20.3720
2	1.20	21.9534	18.3469
3	1.44	18.8619	15.2994
4	1.728	14.3547	11.6479
5	2.0736	9.5081	8.5352

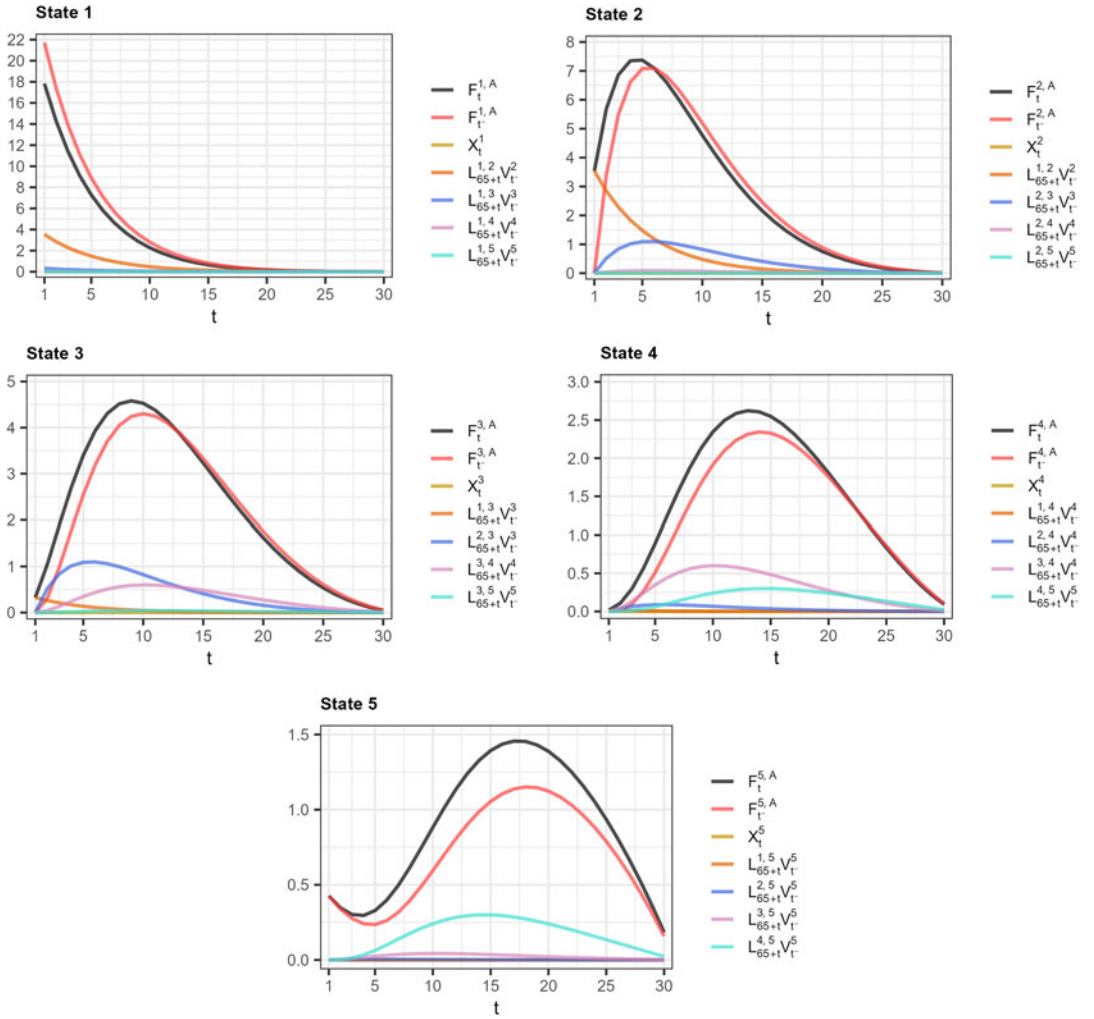


Figure 4. *Available fund (average value), and its components, by time (annual basis) and state for the female gender. y-axis values are on* 10^{-4} *scale. Since the fund value ranges differ significantly across the various states, we adopted separate y-axis scales to enhance the readability of the fund patterns for states* $i > 1$. *For the male gender, refer to the Section 3 of the Supplementary Material.*

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

The time profiles can be interpreted, in particular, considering the several components of the available pool fund by state, as defined in Equation (3.16). For each state, these components are included in the picture. In this regard, we point out that all quantities are plotted in absolute value; however, cross-subsidies may be either incoming or outgoing, depending on the states j considered in relation to state i.

In general, the time profile of each $F_t^{i,A}$ resembles that of survivors in the ageing state, that is, L^i_{65+t}, apart from scaling factors due to the benefit amounts. As the population ages, resources are shifted towards the more severe ageing states. In particular, cross-subsidy is prevalent between consecutive states. For instance, ageing state 2 receives people from state 1 increasing its heap of individual funds, that is, $L^{1,2}_{65+t}V^2_{t^-}$, and transfers part of its individuals (and the relative funds) primarily to state 3, that is, $L^{2,3}_{65+t}V^3_{t^-}$. In turn, ageing state 3 acquires people and funds from state 2, and part of its population is mainly relocated to ageing state 4, and so on. Both cross-subsidy inflows and outflows by state are reported in Table 6. It is worth noting that the cross-subsidy obtained by state 1 is null for every $t > 0$, as is the cross-subsidy offered by state 5. Since only forward ageing transitions are allowed in the Markov ageing model, the lowest ageing state subsidises all the next ones; conversely, the highest ageing state solely receives from the others. Given the initial population composition (see Table 4), the ageing state 1 brings individuals and related funds to subsequent states during the first year. Its cross-subsidy outflow at $t = 1$ is fully allocated among states $j > 1$, corresponding to their respective inflows. As individuals age (or die), state 1 tends to empty, and fewer funds will be transferred. Meanwhile, ageing states $j > 1$ experience an initial increase in incoming cross-subsidies, which then gives way to greater outflows as ageing progresses.

Finally, we notice that X_t^i is null over time and for all states. Such a result arises from state populations never becoming extinct; therefore, no state leaves resources to other states due to population depletion.

4.2.3 Notional benefits

Figure 5 plots the time profile of the average value of the notional benefits by gender. For all states, benefits on average are quite stable; however, at higher chronological ages, they show either an increasing or decreasing path. Such a behaviour can be interpreted in view of cross-subsidy flows, arising from deaths, entries into the state and transitions to higher living states. We note, in particular, an increasing trend in state 5 (which can be exited only upon death) and a decreasing trend for the other states (which are likely to become thinnier, as a result of deaths and ageing). The effect of cross-subsidies for the various states can also be understood in light of the dynamics of the available pool funds by state (see Figure 4). We note that the male population in state 1 is the one that shrinks the most (due to deaths and ageing), see Figure 3; then, the average time profile of benefits in state 1 for males may suggest that cross-subsidies rooted in actuarial equivalence lead to undesirable benefit patterns when the number of participants becomes very low.

Indeed, the benefit dynamics in ageing state 1 deserves careful consideration. While it is true that individuals in this state are exposed to the risk of jumping to a higher state (where they would receive a higher benefit amount), it is also the case that those who remain in state 1 often do so as a result of healthier lifestyles and well-being-oriented behaviours (that may also prove to be more expensive than more passive, sedentary or unhealthy lifestyles). A reduction in benefits for state 1 (especially when, at the same time, benefits in the other states are increasing) may be perceived as a form of penalisation and a disincentive for behaviours that should instead be encouraged due to their social value. This is one of the reasons why it may be appropriate to introduce an adjustment to actuarially equivalent cross-subsidies, by admitting some kind of solidarity among states (see Section 3.5). Numerical results, in particular, suggest that the introduction of solidarity should be considered in favour of those in ageing state 1, that is, individuals with longer life expectancy, who spend their lives in better ageing conditions. This conclusion may appear counterintuitive, as we are naturally inclined to believe that financial support should be directed from healthier individuals to those in poorer health. Thanks to our modelling framework,

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 6. *Ingoing and outgoing cross-subsidies (average value) by state. Values refer to the female gender and, for conciseness, are reported for selected years. Note that, by means of Equation (3.16), the ingoing cross-subsidy for the state i at time $t > 0$ is given by $\sum_{j<i} L^{j}_{x+t-1}\, \widetilde{p}^{j,i}_{x+t-1}\, V^{i}_{t^-} + X^{i}_{t}$, while the outgoing cross-subsidy is quantified as $\sum_{j>i}^{m} L^{i}_{x+t-1}\, \widetilde{p}^{i,j}_{x+t-1}\, V^{j}_{t^-}$. For the male gender, we refer the reader to the Section 3 of the Supplementary Material.*

	Cross-subsidy									
	State 1		State 2		State 3		State 4		State 5	
t	Ingoing	Outgoing	Ingoing	Outgoing	Ingoing	Outgoing	Ingoing	Outgoing	Ingoing	Outgoing
1	0	38,756.487	35,339.848	0	3238.628	0	171.594	0	6.417	0
5	0	16,201.842	14,765.114	11,819.459	12,261.149	3728.129	4434.500	635.132	923.799	0
10	0	5266.055	4783.225	8930.154	8682.250	6429.195	6699.831	2406.938	2867.037	0
15	0	1593.822	1440.639	4451.658	4228.595	5181.581	5188.092	2999.401	3369.136	0
20	0	419.850	375.970	1725.111	1611.437	2954.640	2889.787	2401.432	2623.837	0
25	0	81.187	72.163	472.588	432.710	1164.596	1112.655	1344.759	1445.602	0
30	0	4.387	3.876	33.406	29.944	117.582	107.272	226.012	240.295	0

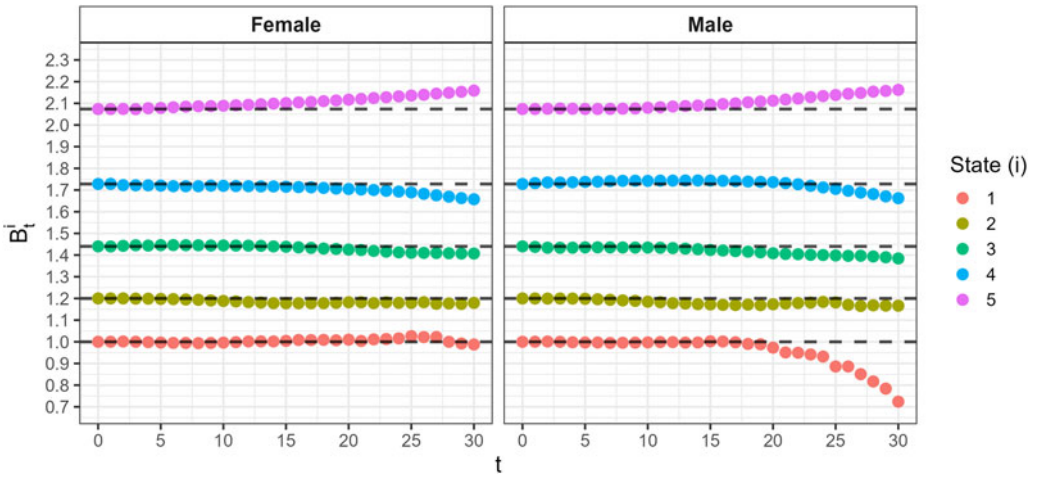


Figure 5. *Individual notional benefits (average value) by time (annual basis) and ageing state, for both genders. For each ageing state, the black dashed line represents the initial benefit level, that is, B_0^i for $i = 1, \ldots, 5$, outlined in Table 5.*

we are able to gain deeper insight into this issue, so as to enable more informed decision-making with respect to the GSA design.

4.2.4 Actual benefits

Notional benefits can be repackaged into actual benefits in a number of ways, each of which implies some degree of solidarity that individuals may perceive as either acceptable or not. For instance, if we consider a plain GSA, that is, a scheme with undifferentiated benefits, it is not difficult to conclude from Figure 5 that the common benefit amount falls between the lowest and the highest notional benefit amounts. Individuals in lower ageing states would implicitly receive solidarity from those in higher ageing states. The difference between notional and actual benefits in this case can be significant enough to lead certain categories of people to opt out of the scheme (provided that participation is not mandatory).

We examine numerically some benefit smoothing solutions that may be seen as suitable by various population groups. We consider four cases:

- $m^* = 1$, meaning that a unique benefit class is acknowledged. In other words, we look at a plain GSA wherein all the ageing states are grouped since the scheme provides undifferentiated benefits;
- $m^* = 2$, that is, two classes of benefit, assuming that $u^2 = 1.5$, $l^1_{x+t} = L^1_{x+t} + L^2_{x+t}$ and $l^2_{x+t} = L^3_{x+t} + L^4_{x+t} + L^5_{x+t}$. This way, ageing states 1 and 2, on the one side, and ageing states 3, 4, and 5, on the other, are grouped. These groupings are based on the assumption that individuals in adjacent ageing states are unlikely to clearly distinguish between their respective mortality characteristics;
- $m^* = 3$, that is, three classes of benefit, assuming that $u^2 = 1.5$, $u^3 = 2$, $l^1_{x+t} = L^1_{x+t} + L^2_{x+t}$, $l^2_{x+t} = L^3_{x+t} + L^4_{x+t}$ and $l^3_{x+t} = L^5_{x+t}$. In this case, ageing class 5 is kept by itself;
- $m^* = 5$, that is, five classes of benefit, assuming that $u^2 = 1.25$, $u^3 = 1.5$, $u^4 = 1.75$, $u^5 = 2$ and $l^k_{x+t} = L^i_{x+t}$ for every $k, i = 1, \ldots, 5$, with $k = i$. We point out that this case corresponds to the solution described in Section 3.2, after Equation (3.5).

The numerical results are plotted in Figure 6, for females (for males, see Section 3 of the Supplementary Material). The smoothing effect is apparent in all cases (as it can be easily seen when

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

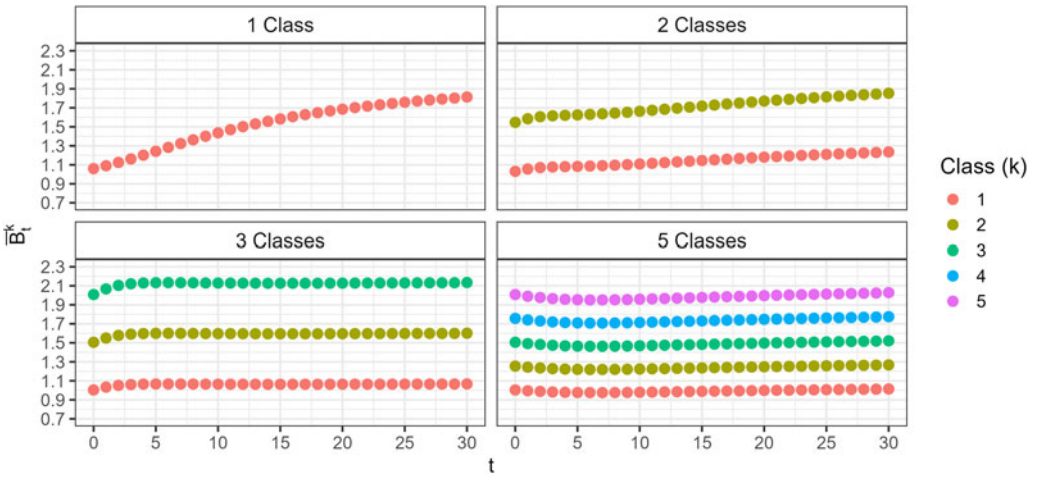


Figure 6. *Actual benefits (average value) by time (annual basis) for the female gender, by considering different numbers $m^\star$ of benefit classes. For the male gender, refer to Section 3 of the Supplementary Material.*

comparing Figure 5 with Figure 6); in particular, as a consequence of the imposed linking among the various benefit classes, the benefit trajectory is identical across all classes, apart from the chosen scaling factor. It should be noted that, in any case, the total payout of the pool remains the same (i.e., B_t^{TOT}) at all times and is determined by requiring funded liabilities by ageing state (i.e., imposing (3.7)). Solidarity contributions can be understood by comparing the notional with the actual benefit for each ageing state. In Table 7, both notional and actual benefits are presented by considering different numbers of benefit classes. In general, higher ageing states offer solidarity to lower ones. The size of such contributions depends on the specific choice of the benefit classes and the chosen proportions among benefit amounts. For instance, if a unique benefit class is acknowledged (i.e., $m^\star = 1$), then lower ageing states receive solidarity from higher ageing states along the entire scheme duration. In particular, it holds that $B_t^i < \overline{B}_t^1$, for $i = 1, 2$ and for every $t > 0$, while ageing states 3 and 4 offer solidarity, respectively, until the end of the 10th and 25th year. State 5, instead, only contributes to solidarity each year. Now, let us consider $m^\star = 2$ benefit classes. We recall that individuals belonging to the first and second ageing states, that is, L_{x+t}^1 and L_{x+t}^2, are relocated into the first benefit class since $l_{x+t}^1 = L_{x+t}^1 + L_{x+t}^2$. As emerges in Table 7, individuals in the first ageing state will earn solidarity, being $B_t^i < \overline{B}_t^1$ for every $t > 0$; conversely, individuals in the ageing state 2 obtain an actual benefit lower than the corresponding notional ones until the twentieth year. The second benefit class pays the actual benefits $\overline{B}_t^2$ to individuals residing in the states 3, 4 and 5. Comparing the actual benefits with the notional benefits B_t^3, B_t^4 and B_t^5, it is easy to see that states 4 and 5 give solidarity to state 3. By also inspecting $m^\star = 3$ and $m^\star = 5$ cases, we observe that as the number of benefit classes increases, the magnitude of the solidarity from higher ageing states to lower ones shrinks. Then, if the scheme aims to provide many benefit classes, a constraint to the solidarity mechanism between ageing states is implicitly introduced.

Once again, we note that the analysis of notional benefits allows us to assess both the magnitude and direction of solidarity effects, which would otherwise remain unspecified. Comparing actual with notional benefits can help identify the most workable smoothing rules, especially in light of the implied solidarity mechanisms.

Finally, we emphasise that in the case of a traditional GSA (i.e., $m^\star = 1$), the actual benefits are increasing over time, but they are undifferentiated. Then, individuals belonging to greater ageing phases perceive the same benefit as individuals in better health statuses, and in the meantime, they offer solidarity. However, if differentiated benefits are provided, more aged individuals continue to contribute solidarity, but they can access greater benefit levels.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Table 7. *Notional benefits (average value), B_t^i, $i = 1, \ldots, 5$, and their related actual benefits, $\overline{B}_t^k$, $k = 1, \ldots, m^\star$, under different numbers $m^\star$ of benefit classes. For conciseness, values are reported for every 5 years up to scheme maturity, and they refer to the female gender. For males, refer to Section 3 of the Supplementary Material.*

						Actual benefits										
	Notional benefits					$m^\star = 1$	$m^\star = 2$		$m^\star = 3$			$m^\star = 5$				
t	B_t^1	B_t^2	B_t^3	B_t^4	B_t^5	$\overline{B}_t^1$	$\overline{B}_t^1$	$\overline{B}_t^2$	$\overline{B}_t^1$	$\overline{B}_t^2$	$\overline{B}_t^3$	$\overline{B}_t^1$	$\overline{B}_t^2$	$\overline{B}_t^3$	$\overline{B}_t^4$	$\overline{B}_t^5$
0	1.000	1.200	1.440	1.728	2.074	1.060	1.031	1.547	1.004	1.506	2.008	1.004	1.255	1.506	1.757	2.008
5	0.997	1.198	1.446	1.721	2.079	1.242	1.083	1.625	1.067	1.600	2.133	0.977	1.221	1.465	1.709	1.953
10	0.996	1.188	1.444	1.719	2.089	1.436	1.109	1.663	1.065	1.597	2.130	0.979	1.224	1.469	1.714	1.959
15	1.004	1.178	1.438	1.715	2.101	1.583	1.145	1.718	1.063	1.595	2.127	0.989	1.236	1.483	1.730	1.978
20	1.010	1.182	1.425	1.705	2.117	1.685	1.180	1.771	1.064	1.596	2.128	0.998	1.248	1.497	1.747	1.997
25	1.016	1.180	1.413	1.693	2.132	1.745	1.204	1.807	1.065	1.598	2.130	1.005	1.256	1.507	1.759	2.010
30	0.988	1.179	1.407	1.658	2.159	1.814	1.236	1.853	1.068	1.601	2.135	1.014	1.268	1.521	1.775	2.028

5. Conclusions

Individuals face increasing exposure to financial and longevity risks in their post-retirement life, and compelling transfer solutions must be developed to support their post-retirement income. In this regard, longevity risk-sharing arrangements have been well accepted in some markets and have raised growing attention in the literature. Such mutual insurance plans can be preferable to traditional insurance designs, mainly because of cost savings and enhanced flexibility, although at the expense of guarantees (as all risks are borne by the group of the plan's participants).

In this paper, we consider one of the earlier pooling structures described in the literature, that is, the GSA arrangement, and we focus on longevity risk. We are particularly concerned with the heterogeneity in mortality among individuals and the diverse financial needs this may generate regarding a suitable post-retirement income. This issue has been only partially explored in the literature.

Mortality heterogeneity stems from multiple risk factors leading individuals to age at different paces, although they are equally aged chronologically. If such risk factors are not measurable, human ageing must be managed as a latent process with a material impact on individuals' observable time to death. To this end, we represent the ageing process through the Markov ageing model. Each process state represents a distinct ageing phase; transitions capture the random progression towards death. The multi-state structure of the ageing model is well-suited for allowing the design of differentiated benefits, and consequently for meeting the greater financial needs of frailer individuals; to our knowledge, the Markov ageing implementation in a risk-sharing context with multiple levels of benefit is new in the literature.

The Markov ageing GSA scheme, which we propose, provides differentiated benefits to participants based on their ageing profiles. Individuals in more vulnerable conditions (presumably with a lower life expectancy) are entitled to higher benefits than those with better physiological statuses (and likely a longer lifespan). In a multi-state model, each ageing state is associated with a specific benefit level. For each individual, benefits are updated dynamically in two respects. First, as an individual's ageing trajectory evolves, (s)he moves to a new state; a different benefit level is provided accordingly. This is anticipated at entry, since the initial capital contributed by the individual is assessed to match the value of the sequence of benefits (s)he is expected to cash back. Second, state-specific benefits are adjusted depending on the possible discrepancy between the state experience and the related initial expectation. Such an adjustment is not anticipated when calculating the initial capital. It becomes necessary when the realised ageing and mortality trajectories deviate from those assumed in the calculation of the expected payout rate.

Similar to a plain GSA scheme, the management of a multi-state GSA relies on cross-subsidies. While in a plain scheme the cross-subsidy occurs between deceased and surviving members, in a multi-state framework it also applies to transitions between ageing states. Identifying these cross-subsidies is crucial for the formulation of appropriate benefit adjustment rules, an aspect that has so far received limited attention in the literature about GSAs. In this paper, we focus in-depth on the definition of cross-subsidies in a multi-state setting and discuss how this allows us to design benefit adjustment rules that are free from solidarity effects. We also address the introduction of solidarity mechanisms, which may be required in various contexts. For instance, risk classification must be based on proxy risk factors, and the resulting risk classes do not fully coincide with the real ones. Likewise, due to regulation or market expectations, it could be necessary to stick to a target level of benefit differentiation. We believe that, even when their inclusion is inevitable, the availability of a framework to quantify their magnitude can improve the transparency of the pooling arrangement and support a constructive communication of design choices to participants.

We test our theoretical framework on Australian data and find some interesting results. In particular, we identify circumstances that may require the introduction of solidarity mechanisms in the design of multi-state benefits. The analysis we develop allows us to characterise both the extent and the direction of these effects.

We stress that in the Markov ageing model, transitions are only admitted to higher states, that is, backwards transitions to lower states are excluded. Accordingly, the actuarial design of the GSA we

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

propose only considers forward transitions; however, it can be extended to also account for backwards transitions. The algebraic expressions of the updated benefits and cross-subsidies would be different, but the same logic can be preserved. Similarly, the actuarial structure can be adapted to a multi-state mortality model accounting for observable risk factors only.

We believe that our paper offers stimulating insights for future research. A first step forward is generalising the actuarial design of the proposed Markov ageing GSA. For example, one could consider a pool open to multiple cohorts. Consequently, cross-subsidy and solidarity mechanisms must be studied even among different generations of participants. Furthermore, it would be interesting to test the impact of different risk-sharing rules on the benefits profile and possible alternative criteria for applying solidarity. From the perspective of mortality uncertainty modelling, the Markov ageing model can be boosted by considering a non-homogeneous Markov process (see, e.g., Albrecher *et al.* 2022) or a subordinated Markov process (see, e.g., Liu and Lin 2012). In the latter case, the Markov ageing GSA could be assessed in the presence of systematic longevity risk.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/asb.2026.10095

Acknowledgements. The authors sincerely thank Hansjöerg Albrecher for his valuable suggestions on using the Markov ageing-based mortality modelling.

Financial support. The authors acknowledge the support of MIUR-PRIN 2022 project 'Building resilience to emerging risk in financial and insurance market, grant 2022FWZ2CR. R. Maggistro and M. Marino are grateful to the Society of Actuaries, Research Institute, for its support through the funded project: 'A biological inspired perspective in longevity risk management: which implications for actuaries?', grant N. KP199.

Dedication. This research is dedicated to Anna Rita Bacinello, who passed away prematurely during the early stages of this project. Her intellectual contribution, heartfelt friendship and unwavering support are deeply missed.

Competing interests. The authors have no potential competing or conflicting interests to declare.

References

Albrecher, H., Bladt, M., Bladt, M. and Yslas, J. (2022) Mortality modeling and regression with matrix distributions. *Insurance: Mathematics and Economics*, **107**, 68–87.

Aragona, M., Regis, L. and Vigna, E. (2025) A redistributive GSA scheme to cope with socio-economic mortality differentials. *ASTIN Bulletin*, **55**(3), 757–774.

Asmussen, M., Nerman, O. and Olsson, M. (1996) Fitting phase-type distributions via em algorithm. *Scandinavian Journal of Statistics*, **23**(4), 419–441.

Blake, D., Cairns, A. and Dowd, K. (2006) Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, **12**(1), 153–197.

Blake, D., Cairns, A., Dowd, K. and Kessler, A. (2019) Still living with mortality: The longevity risk transfer market after one decade. *British Actuarial Journal*, **24**(1), e1.

Brown, J. (2001) Private pensions, mortality risk, and the decision to annuitize. *Journal of Public Economics*, **82**(1), 29–62.

Brown, J. (2009) Understanding the role of annuities in retirement planning. In *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs*. Chicago: University of Chicago Press.

Chen, A., Chen, Y. and Xu, X. (2022) Care-dependent tontines. *Insurance: Mathematics and Economics*, **106**, 69–89.

Chen, A., Hieber, P. and Klein, J.K. (2019) Tonuity: A novel individual-oriented retirement plan. *ASTIN Bulletin*, **49**(1), 5–30.

Cheng, B., Jones, B., Liu, X. and Ren, J. (2020) The mathematical mechanism of biological aging. *North American Actuarial Journal*, **25**(1), 73–93.

de Andrés-Sánchez, J. and González-Vila Puchades, L. (2025) Longevity risk and annuitisation decisions in the absence of special-rate life annuities. *Risks*, **13**(2), 37.

Dhaene, J. and Milevsky, M.A. (2024) Egalitarian pooling and sharing of longevity risk a.k.a. can an administrator help skin the tontine cat? *Insurance: Mathematics and Economics*, **119**, 238–250.

Donnelly, C. (2015) Actuarial fairness and solidarity in pooled annuity funds. *ASTIN Bulletin*, **45**(1), 49–74.

Donnelly, C., Guillén, M. and Nielsen, J.P. (2013) Exchanging uncertain mortality for a cost. *Insurance: Mathematics and Economics*, **52**(1), 65–76.

Donnelly, C., Montserrat, G. and Nielsen, J.P. (2014) Bringing cost transparency to the life annuity market. *Insurance: Mathematics and Economics*, **56**, 14–27.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press

Félix, J. Martnez de Toda, I., Daz-Del Cerro, E., González-Sánchez, M. and De la Fuente, M. (2024) Frailty and biological age. Which best describes our aging and longevity? *Molecular Aspects of Medicine*, **98**, 101291.
Hieber, P. and Lucas, N. (2022) Modern life-care tontines. *ASTIN Bulletin*, **52**(2), 563–589.
Jackson, S.H.D., Weale, M.R. and Weale, R.A. (2003) Biological age-what is it and can it be measured? *Archives of Gerontology and Geriatrics*, **36**(2), 103–115.
Kabuche, D., Sherris, M., Villegas, A.M. and Ziveyi, J. (2024) Pooling functional disability and mortality in long-term care insurance and care annuities: A matrix approach for multi-state pools. *Insurance: Mathematics and Economics*, **116**, 165–188.
Li, Z., Zhang, W., Duan, Y., Niu, Y., Chen, Y., Liu, X., Dong, Z., Zheng, Y., Chen, X., Feng, Z., Wang, Y., Zhao, D., Sun, X., Cai, G., Jiang, H. and Chen, X. (2023) Progress in biological age research. *Frontiers in Public Health*, **11**, 1074274.
Lin, S.X. and Liu, X. (2007) Markov aging process and phase-type law of mortality. *North American Actuarial Journal*, **11**(4), 92–109.
Liu, X. and Lin, S.X. (2012) A subordinated Markov model for stochastic mortality. *European Actuarial Journal*, **2**, 105–127.
Milevsky, M.A. (2020) Calibrating Gompertz in reverse: What is your longevity risk-adjusted global age? *Insurance: Mathematics and Economics*, **92**, 147–161.
Milevsky, M.A. and Salisbury, T.S. (2015) Optimal retirement income tontines. *Insurance: Mathematics and Economics*, **64**, 91–105.
Olivieri, A. and Pitacco, E. (2020) Linking annuity benefits to the longevity experience: Alternative solutions. *Annals of Actuarial Science*, **14**(2), 316–337.
Olivieri, A. and Tabakova, D. (2024) Stochastic assessment of special-rate life annuities. *Decisions in Economics and Finance*, **48**, 913–932.
Olivieri, A., Thirurajah, S. and Ziveyi, J. (2022) Target volatility strategies for group self-annuity portfolios. *ASTIN Bulletin*, **52**(2), 591–617.
Piggott, J., Valdez, E. and Detzel, B. (2005) The simple analytics of a pooled annuity fund. *The Journal of Risk and Insurance*, **72**(3), 497–520.
Pitacco, E. and Tabakova, D.Y. (2022) Special-rate life annuities: Analysis of portfolio risk profiles. *Risks*, **10**(3), 65.
Qiao, C. and Sherris, M. (2013) Managing systematic mortality risk with group self-pooling and annuitization schemes. *The Journal of Risk and Insurance*, **80**(4), 949–974.
Ramsay, C.M., Oguledo, V.I. and Krutto, A. (2025) The refinancing enhanced annuity rider (rear): A novel annuity product design and its contribution to solving the annuity puzzle. *North American Actuarial Journal*, **29**(3), 758–790.
Stamos, M.Z. (2008) Optimal consumption and portfolio choice for pooled annuity funds. *Insurance: Mathematics and Economics*, **43**(1), 56–68.
Su, S. and Sherris, M. (2012) Heterogeneity of Australian population mortality and implications for a viable life annuity market. *Insurance: Mathematics and Economics*, **51**, 322–332.
Vaupel, J.W., Villavicencio, F. and Bergeron-Boucher, M.-P. (2021) Demographic perspectives on the rise of longevity. *Proceedings of the National Academy of Sciences*, **18**(9), e2019536118.
Weinert, J. and Gründl, H. (2021) The modern tontine. *European Actuarial Journal*, **11**, 49–86.
Yaari, M. (1965) Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, **32**(2), 137–150.

https://doi.org/10.1017/asb.2026.10095 Published online by Cambridge University Press