Practical analysis of a pooled annuity fund with integrated bequest

Practical analysis of a pooled annuity fund with integrated bequest

Catherine Donnelly*

https://risk-insight-lab.com/

Risk Insight Lab,
Department of Actuarial Mathematics and Statistics,
Heriot-Watt University, Edinburgh, Scotland EH14 4AS

June 17, 2022

Abstract

A pooled annuity fund with an integrated bequest, introduced by Bernhardt and Donnelly (2019), is investigated for a finite pool of participants. In this fund, participants each receive an income while they are alive and a bequest payment upon their death, in exchange for paying their pension savings into the fund when they join.

The fund is placed in a discrete-time setting in this paper and the expected present value of a regular payment for life from the fund is derived. Using this annuity value, a numerical simulation is done of the fund under a stochastic mortality model. Each year, a new group of people join the fund, each bringing the same amount of money and with lives assumed to be independent random variables.

It is found that the distribution of income paid to each cohort varies, with the last cohort having the widest income distribution when they are old. Around 100 participants are needed to join every year to gain most of the benefits of pooling longevity risk, based on the model used for the simulations. These results hold whether there is a bequest or not in the fund.

The results show that a bequest could be integrated into pooled annuity fund, and still pay a stable income, albeit at a cost of a lower income compared to the fund without a bequest.

Keywords: tontine; longevity risk; mortality; retirement; decumulation.

*C.Donnelly@hw.ac.uk

1. Introduction

With millions and millions of people worldwide saving billions and billions of dollars into defined contribution pension schemes, the looming question is: how will they convert their pension pot into a retirement income for life? The two standard options currently on offer sit at opposite ends of the risk spectrum. Conventional life annuities remove virtually all risk from the retiree, but at a cost that makes it unattractive to most retirees. Income drawdown – keeping your pension pot invested and withdrawing money to live on, which is also called phased withdrawal or self-insurance – leaves all risk and decisions with the pensioner on an ongoing basis. Withdraw too much and you risk being ruined. Withdraw too little and you risk being miserable due to too little money to live the life you want. However, a third option, pooled annuity funds, has gained increasing interest from academics, industry and regulators. Can this third option offer a middle ground for those for whom life annuities are not seen as offering value-for-money, but for whom income drawdown is too much of a confusing gamble?

The purpose of the pooled annuity fund is to pay a lifetime income to its participants, by using longevity risk-sharing. In this context, longevity risk-sharing simply means that everyone who participates in the fund becomes the beneficiary of each other. More precisely, if a participant dies, their money in the fund which is allocated to longevity risk-sharing is shared out among the other participants, with the share calculated by a defined formula. The allocated amount of money, called a longevity credit, is earned on top of investment returns. It enables a higher expected income to be paid to the survivors in the fund compared to those who follow the same investment strategy but do not engage in any longevity risk-sharing.

Here a pooled annuity fund with a bequest feature is examined in a discrete-time setting. It was proposed and studied in continuous time, assuming an infinite number of participants, by Bernhardt and Donnelly (2019). This innovative structure pays an income while the participant is alive and a bequest to their estate when they die. The innovation comes from the linkage between the income and the bequest, which is explained below. It is shown here that, at least in the theoretical setting, such a fund can provide a lifelong income and a death bequest when implemented among a finite pool of participants. This is the laboratory proof-of-concept of the structure. To provide a lifelong income requires using an appropriate withdrawal rate, which is derived here as a further contribution.

A numerical study of the income paid from such a fund, and of one without a bequest feature, in the presence of systematic longevity risk, is then done. Qiao and Sherris (2013) also studies a pooled annuity fund, although one without a bequest feature, when systematic longevity risk exists. One of their most important results is that it is better to pool longevity risk across generations than only within a generation. Doing so narrows the distribution of the income paid significantly. This is related to the finding that pooling longevity risk is highly important at higher ages, as discussed in Qiao and Sherris (2013).

However, they did not study a fund which eventually closes to new members. This is the fund studied in this paper, as eventually all funds must close. It brings up new questions such as, how does each generation to join the fund fare compared to other generations? It is found that the last generation has the poorest outcomes in terms of a wide income distribution when they are old.

In this paper, a numerical simulation of the income distribution faced by each generation to join the fund is done and discussed, for a fund with and without a bequest. Having many people with whom to pool longevity risk when you are old reduces the volatility of your income. However, as the last generations to join the scheme age, they have few people left alive with whom to pool their longevity risk. This means that they experience a much higher income volatility when they are old, compared to earlier generations. Moreover, the last survivors in the fund bear the cumulative consequences of paying out too much or too little over time to themselves and earlier participants. This is the manifestation of systematic longevity risk. Overall, the result imply that risk management for the last generations in the fund needs to be considered in advance and managed appropriately.

A further contribution of the paper is on how many people are required to adequately pool longevity risk. This is of critical importance to the implementation of pooled annuity funds. The problem has been studied in Bernhardt and Donnelly (2020), for a fund in which there is only idiosyncratic longevity risk, i.e. the distribution of future mortality rates was known, and no bequest feature. In their setting of a single-cohort fund (i.e. only one generation is in the fund), of the order of 1 000 participants are needed to eliminate most of the idiosyncratic longevity risk, each participant an independent and identical copy of the others at the start (Donnelly and Bernhardt, 2021, Figure 2). However, allowing for generations to enter each year, the results in this paper show that of the order of 100 participants are needed to join each year (with the caveat that the longevity risk pooling is not sufficient for the oldest survivors of the last generations). This holds under the assumption that participants bring the same amount of money to the fund and that their mortality follows the Cairns-Blake-Dowd mortality model, first detailed in Cairns et al. (2006), applied in this paper. Furthermore, it is observed that the number of participants required to pool sufficiently longevity risk is not affected by the bequest feature.

2. Pooled annuity fund literature review

Structures in which longevity risk is pooled directly among participants have increasingly gained industry and academic interest in recent years. Various ways of doing longevity risk-sharing have been proposed. Some of these work for single-cohort pools in which every member is an independent and identical copy of each other (for example, Milevsky and Salisbury 2015; Stamos 2008) and some are intended for multi-cohort pools (for example, Piggott et al. 2005; Stamos 2008; Sabin 2010; Qiao and Sherris 2013; Donnelly et al. 2014; Milevsky and Salisbury 2016). Some calculate what proportion of the funds of those who have died should be received by each participant − whether explicitly (Stamos, 2008; Sabin, 2010; Donnelly et al., 2014) or implicitly (Piggott et al., 2005; Qiao and Sherris, 2013). Both of these explicit and implicit schemes then use an annuity value to calculate the income paid out. In the latter schemes, which assume that retirees prefer a fairly constant income to a power utility-optimized one, if there are enough participants in the fund to eliminate idiosyncratic longevity risk and systematic longevity risk does not exist, then the income payments from the fund should be constant over the lifetime of each participant.

Milevsky and Salisbury (2015) take a different tack and calculate the payout from a pooled annuity fund which maximizes the expected discounted value of lifetime consumption. Thus the approach followed by Milevsky and Salisbury (2015) makes different assumptions about what is attractive to individuals. In their setting, the optimal payout then varies according to the problem considered, as can be seen in various papers employing this approach (for example, Chen et al. 2020, 2021). The approach of Milevsky and Salisbury (2015) to calculating the income from a pooled annuity fund is potentially more consistent with any subsequent analysis of the fund which employs expected utility theory to assess the attractiveness of different options.

In a highly relevant paper, though not targeted specifically to pooled annuity funds, Denuit and Robert (2021) present various fair linear risk-sharing rules, and a conditional mean risk-sharing rule, in a more general, unifying setting, and study their convergence. Interestingly, Weinert and

Gründl (2021) derive a distribution to model the longevity credits paid from the pooled annuity fund, rather than modelling directly the mortality experience of the pool of participants.

In this paper, the longevity risk-sharing method proposed in Qiao and Sherris (2013) is applied, chosen for its relative simplicity and ease-of-calculation. As inflation is not considered here, the assumed ideal situation is to make the same payments at each age. Thus the fund value of each survivor is divided by the expected present value of a stream of constant payments of 1 unit per annum, valued within the relevant pooled annuity fund, to obtain the income paid out. The calculation is re-done at annual intervals. The numerical study observes the distribution of the payments. The motivation for aiming for constant payments at each age is that they are the *de facto* income profile from pension plans or annuity products, whether constant in nominal or real terms.

The demand for pooled annuity fund compared to life annuities has been studied by various authors (for example, Piggott et al. 2005; Valdez et al. 2006; Donnelly et al. 2013; Hanewald et al. 2013; Milevsky and Salisbury 2016; Chen et al. 2021). The results show that the attractiveness of pooled annuity funds increases as the risk aversion of the retiree reduces. This is because the less risk averse retirees are happier to bear the volatility of pooled annuity funds in exchange for their higher expected return. Pooled annuity funds become increasingly preferred to life annuities as the loadings on the latter increase.

Turning to the bequest aspect, some individuals value being able to leave money to their dependents. This may mean leaving money ring-fenced for that purpose, buying a life insurance contract or perhaps there is no plan to leave a specific amount, but simply what is left at the time of the individual's death.

Hanewald et al. (2013) go back to financial economic basics to show the demand for pooled annuity funds relative to typical post-retirement investments, when systematic longevity risk is present. One of their findings is that people with a bequest motive will buy less of the life annuity and replace it with a risk-free bond, in order to provide the bequest. As the loadings on the life annuity increase, they will also replace it with investment in a pooled annuity fund.

Chen and Rach (2022) allow for a bequest through bundling a life insurance contract with a pooled annuity fund. They take the perspective of the insurer rather than the individual. They find that their bundled product may lead to lower risk margins than the products sold separately. This is a different product to that in the present paper, since the bequest is provide through a separate contract in Chen and Rach (2022) and is not affected by the pooled annuity fund as it is in this paper.

To provide both an income and a bequest, Zhou et al. (2021) propose a portfolio consisting of a life annuity contract and a product that pays out the return earned on a principal amount while leaving the principal untouched (called the "natural income"). The bequest benefit is then the principal amount. This set-up allows the individual to choose their own attractive combination of the two products. Due to natural hedging, they find that the interest rate risk can be effectively hedged by coupon-bearing government bonds.

Bernhardt and Donnelly (2019) propose and study optimal strategies for the pooled annuity fund with integrated bequest which is analysed in this paper. They determine optimal investment and consumption strategies when there is no idiosyncratic longevity risk. Their model does not allow for systematic longevity risk. They find that, the stronger the bequest motive, the more money that is allocated to the bequest account. Dagpunar (2021) extends the analysis by allowing the proportion of longevity credits funnelled to the bequest account to vary. He suggests a way of choosing the parameter that indicates the strength of the bequest motive.

For the purposes of this paper, the conclusion is that there a theoretical demand for pooled annuity funds with a bequest. However, other theoretical aspects of their implementation, in markets where they do not exist, must be studied. For example, how many people are needed to join the funds? How should the income be calculated? By how much can the income fall? These are all questions asked by practitioners and which this paper attempts to address, for the given fund structure and membership profile.

3. Income payments in a pooled annuity fund with bequest

In the pooled annuity fund with integrated bequest, the retiree's savings in the pooled annuity fund are allocated to two accounts: the tontine account and the bequest account. The tontine account is ear-marked as part of a pooled annuity fund. The tontine account is shared out among the pooled annuity fund participants upon the retiree's death. The bequest account is not part of the pooling arrangement. Thus it is not shared out among the other participants upon death. Instead, it is paid to the estate of the account-holder when they die.

As only the tontine account is shared out among the other participants upon the retiree's death, only the tontine account earns longevity credits. However, a fixed fraction of the longevity credits, earned by the tontine account, is re-directed to the bequest account. This is the motivation for called it an 'integrated bequest'. It means that less money is available to the pool as a whole compared to a fund in which all money is allocated to the tontine account. When anyone dies in the fund with integrated bequest, their estate gets their bequest account and this money is lost forever to the other participants.

Otherwise, both accounts earn investment returns at the same rate and have money withdrawn at the same rate to meet the retiree's spending needs. (Neither of these assumptions are necessary but they are simplifying.) Thus both the bequest account and the tontine account are used to fund the retiree's spending requirements. Yet the fact that some longevity credits are channelled from the tontine account to the bequest account has two implications:

- The rate of sustainable, lifetime income is lower compared to that from a pooled annuity without the bequest account; and
- The annuity value, used to calculate how much can be sustainably withdrawn from a retiree's total pension savings, is higher.

In this section, the expected present value of a life annuity paid from a pooled annuity fund with bequest is derived, for payments made in advance each year. It is analogous to the standard expected present value, represented by the international actuarial notation $\ddot{a}_x$.

3.1. Derivation of the discretely paid life annuity value

Suppose that, in the pooled annuity fund with bequest, a proportion $\alpha \in [0, 1]$ of the longevity credits earned on each participant's tontine account at the end of each calendar year are diverted to that participant's bequest account. As both accounts are assumed to earn investment returns at the same rate and have income drawn down from them at the same rate, this is equivalent to re-balancing the tontine account and bequest account at the end of the year so that a proportion $\alpha \in [0, 1]$ of the participant's savings are in the tontine account, with the remainder in the bequest account.

All participants in the fund are assumed to be the same age with the same amount of money in the fund when they join. Assuming that there are $\ell \geq 1$ survivors in the fund at the end of the calendar year of death of the participant, they receive the fraction $1/\ell$ of the recently deceased's tontine account as their longevity credit. The exception is for the last person to die: their estate gets both the bequest account and tontine account (in any case, they had no-one with whom to pool their longevity risk).

The expected present value of a payment of one unit paid annually in advance from a pooled annuity fund with bequest, denoted by $\ddot{a}_x^{(\alpha)}$, is shown next.

Lemma 3.1. *Suppose an individual has brought savings of amount $\ddot{a}_x^{(\alpha)}$ to a pooled annuity fund with integrated bequest when they are age x years. By assumption, this amount is enough to pay the individual 1 unit per annum in advance for life. Equivalently, the expected present value of paying 1 unit per annum in advance for life to the individual, is denoted by $\ddot{a}_x^{(\alpha)}$. Assume that there is no systematic longevity risk.*

Investment returns are earned at the constant effective rate i per annum. Let q_{x+m} denote the probability of an individual, who is alive at age $x+m$, dying by age $x+m+1$. Assume there exists a limiting integer age $x_\infty \in [0,\infty)$, at which the individual must die at or before, i.e. $q_{x_\infty - 1} = 1$.

Then

$$\ddot{a}_x^{(\alpha)} = 1 + \sum_{k=1}^{\infty} (1+i))^{-k} \prod_{m=0}^{k-1} \frac{1 - q_{x+m}}{1-(1-\alpha)q_{x+m}}. \tag{1}$$

Proof. Assume that the individual is in a fund with infinitely many identical and independent copies of themselves, so that idiosyncratic longevity risk is fully diversified away. First the longevity credit is calculated, using the simple rule of sharing out the total value of the tontine accounts of those who have died equally amonf the survivors.

The value in the tontine account at age x of each participant is $\alpha\ddot{a}_x^{(\alpha)}$, which falls to $\alpha(\ddot{a}_x^{(\alpha)} - 1)$ just after the first annuity payment is made. Suppose $d_{x+1} \geq 0$ participants die between age x and $x+1$, leaving $\ell_{x+1} > 0$ survivors at age $x+1$. This means that a total amount of $\alpha(\ddot{a}_x^{(\alpha)} - 1)d_{x+1}$ is released by the deaths over the first year and is shared out equally among the ℓ_{x+1} survivors when they are age $x+1$. Mathematically, the longevity credit paid to a surviving individual when they are age $x+1$ is

$$\alpha\left(\ddot{a}_x^{(\alpha)} - 1\right)(1+i)\frac{d_{x+1}}{\ell_{x+1}}.$$

The expected present value of the annuity at age $x+1$, denoted $\ddot{a}_{x+1}^{(\alpha)}$, is the accumulation of $\alpha(\ddot{a}_x^{(\alpha)} - 1)$ plus the longevity credit gained at age $x+1$, i.e.

$$\ddot{a}_{x+1}^{(\alpha)} = \left(\ddot{a}_x^{(\alpha)} - 1\right)(1+i) + \alpha\left(\ddot{a}_x^{(\alpha)} - 1\right)(1+i)\frac{d_{x+1}}{\ell_{x+1}}.$$

Rearranging and using that $d_{x+1}/\ell_x = q_x$ and $\ell_{x+1}/\ell_x = 1 - q_x$, gives the recursion

$$\ddot{a}_x^{(\alpha)} = 1 + \frac{1}{1+i}\frac{1-q_x}{1-(1-\alpha)\,q_x}\,\ddot{a}_{x+1}^{(\alpha)}.$$

As participants must all die at some finite age, it follows that $\lim_{N\to\infty} \ddot{a}(x+N) = 0$. Using this limit and applying the last recursion repeatedly, results in

$$\ddot{a}_x^{(\alpha)} = 1 + \sum_{k=1}^{\infty} (1+i))^{-k} \prod_{m=0}^{k-1} \frac{1 - q_{x+m}}{(1-(1-\alpha)q_{x+m})}.$$

□

Thus the discretely-paid expected present value of a lifetime payment of 1 unit from the modern tontine with bequest has been derived. Setting $\alpha = 0$ results in the present value of a perpetuity and setting $\alpha = 1$ results in the expected present value of a conventional life annuity.

A term-by-term comparison in the summand shows that $\ddot{a}_x^{(\alpha)}$ is a decreasing function of α. The higher the proportion of the fund value that is assigned to the tontine account, the cheaper it becomes to provide a pension of 1 unit per annum. This is because more longevity credits are obtained, when a higher fraction of the fund value is allocated to the tontine account.

4. Pooled annuity fund evolution

With the means now at hand to calculate the income withdrawal rate from the pooled annuity fund with integrated bequest, turn now to the specification of the model for the numerical simulations. The evolution of the fund participants' accounts and the stochastic mortality model are detailed in this section, with some supplementary results relegated to the Appendix.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which all stochastic processes are defined. Time is measured in years.

4.1. Economic model

It is assumed that invested assets earn returns at a constant effective interest rate of i per annum.

4.2. Stochastic mortality model

In a deterministic mortality model, the assumption is that future mortality rates are known. The source of risk is solely idiosyncratic risk: the risk that the observed mortality rates – calculated from the observed deaths in the population – do not match the known mortality rates.

Under a stochastic mortality model, it is only the distribution of future mortality rates which is known. With an increasing population size, the distribution tends to a limiting distribution. However, the mortality rate experienced by the population remains random, being a random draw from the distribution. This is the undiversifiable, systematic longevity risk; it cannot be eliminated by adding more people to the population. With a finite population, idiosyncratic longevity risk also comes into play.

The stochastic mortality model used in the numerical simulations is the two-factor Cairns-Blake-Dowd model, as detailed in Cairns et al. (2006). This is well-studied model which has is shown to effectively model mortality at higher ages.

Let the term $\bar{x}$ denote the constant average age of the studied population. Define the inverse of the logit function $\operatorname{logit}^{-1}(y) := e^y/(1+e^y)$ for $y \in \mathbb{R}$ (Aitchison and Shen, 1980).

In the Cairns-Blake-Dowd model applied in discrete time, the annual probability of someone who is age x years at integer time $t \in \mathbb{N}_0$ dying by time $t+1$ is modelled by a stochastic process $q : \Omega \times \mathbb{N}_0 \times [0, \infty) \to [0, 1]$, with

$$q(t, x) := \operatorname{logit}^{-1}\left(K_1(t+1) + K_2(t+1)(x - \bar{x})\right).$$

The terms $K_1 : \Omega \times \mathbb{N}_0 \to \mathbb{R}$ and $K_2 : \Omega \times \mathbb{N}_0 \to \mathbb{R}$ represent random walks with drift. Specifically, defining $\mathbf{K}(t) = (K_1(t), K_2(t))'$, in which the prime denotes transpose, then

$$\mathbf{K}(t+1) = \mathbf{K}(t) + \boldsymbol{\nu} + C\,\mathbf{Z}(t), \tag{2}$$

in which $\boldsymbol{\nu}$ is a 2-dimensional vector and C is a 2×2-matrix, both of which are chosen by fitting the model to some chosen data set. The term $\mathbf{Z}(t) = (Z_1(t), Z_2(t))'$ is a standard two-dimensional normal bivariate random vector. The filtration is the natural one generated by the process $\mathbf{Z}$. Note in particular that $q(t,x) \in \mathcal{F}_{t+1}$; for someone age x at time t, the probability of dying between time t and $t+1$ is not known until time $t+1$.

The parameter $\boldsymbol{\nu}$ represents the level of mortality improvements, which push the value $K_1(t+1)+K_2(t+1)(x-\bar{x})$ towards the bottom tail of the inverse logit function. The matrix C represents the level of systematic longevity risk in the model. The systematic longevity risk cannot be removed by having large numbers of participants, which will be seen in the numerical simulations. It will cause the income distribution at each future time to widen as time increases.

4.2.1. Parameters of the fitted model

To use the model to project future mortality paths, the general idea is to first calculate $\boldsymbol{\nu}$ and C from an existing data-set. Here a modification of the fitted model in Cairns (2011) is used. Cairns (2011) fits the Cairns-Blake-Dowd model to England and Wales male data, from ages 60 to 89 in the calendar years 1981 to 2008. From the model fitting (Cairns, 2011, Equation (1)),

$$CC' = \begin{pmatrix} 0.000\,453\,8 & 0.000\,015\,85 \\ 0.000\,015\,85 & 0.000\,001\,256 \end{pmatrix}.$$

In our model, the time scales are fairly large. Longevity improvements also make a comparison of outcomes by age of cohorts more difficult, as each cohort lives longer than its predecessor. For example, in the presence of longevity improvements, the future mortality faced by a 70-year old today is not the same as that faced by a 70-year old in 10 years' time. Since removing longevity improvements does not change the results in any meaningful way for the objectives of this paper, the longevity improvement parameter is set to $\boldsymbol{\nu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in this paper. This decision to remove longevity improvements also side-steps the issue of perpetually improving life-spans, which arises in the presented model.

Time 0 corresponds to the end of calendar year 2008, the last year of the data used to fit the model. The starting vector value of the random walk, $\mathbf{K}(0)$, is estimated from the data set as part of the model-fitting process; it is the last value of $(K_1(t), K_2(t))'$ generated by the estimation procedure. Here, $\mathbf{K}(0) = (-3.2717, 0.1079)'$ (Cairns, 2011).

Under the chosen mortality model, with no allowance for future longevity increases, the expected future lifetime for a member age 65 is 17.3 years, with a standard deviation of 0.5 years. As an example, if 100 people joined the scheme at time 0 when they were 65 years old, then around 2 people survive to age 95 and less than 1 to age 105.

4.2.2. Life annuity calculation under the Cairns-Blake-Dowd model

Suppose an individual is age x at integer time $t \geq 0$. They are to receive a life annuity of amount 1 unit per annum payable in advance, starting at time t. Applying the expression (1) and noting that the calculation is done with the information available at the start of the payment stream, the expected value at time t of the life annuity is

$$\ddot{a}_x^{(\alpha)}(t) := \ddot{a}_x^{(\alpha)}(t, \mathbf{K}(t)) = 1 + \sum_{T=1}^{\infty} (1+i)^{-T}\, \mathbb{E}\left(\prod_{k=0}^{T-1} \frac{1 - q(t+k, x+k)}{(1-(1-\alpha)q(t+k, x+k))} \,\middle|\, \mathbf{K}(t) \right).$$

Define for $t \in \mathbb{N}_0$ and $T \in \mathbb{N}$,

$$p^{(\alpha)}(t, t+T, x-t, \mathbf{K}(t)) := \mathbb{E}\left(\prod_{k=0}^{T-1} \frac{1 - q(t+k, x+k)}{(1-(1-\alpha)q(t+k, x+k))} \,\middle|\, \mathbf{K}(t) \right). \tag{3}$$

Then

$$\ddot{a}_x^{(\alpha)}(t) = 1 + \sum_{T=1}^{\infty} (1+i)^{-T}\, p^{(\alpha)}(t, t+T, x-t, \mathbf{K}(t)).$$

With $\alpha = 1$, the annuity expression becomes the same one studied in Cairns (2011, Section 3.1). In that instance, the interpretation of $p^{(1)}(t, t+T, x-t, \mathbf{K}(t))$ is the expected probability of surviving from time t to time $t+T$, for an individual who is age x at time t (and hence age $x-t$ at time 0), conditional on the aggregate mortality at time t (as represented by $\mathbf{K}(t)$).

From the Markov and time-homogeneous nature of the random walk $\mathbf{K}$, it follows that $p^{(\alpha)}(t, t+T, x-t, \mathbf{k}) = p^{(\alpha)}(0, T, x, \mathbf{k})$. Thus, as noted in Cairns (2011) for the case $\alpha = 1$ but which applies to any $\alpha \in [0, 1]$,

$$\ddot{a}_x^{(\alpha)}(t, \mathbf{K}(t)) = 1 + \sum_{T=1}^{\infty} (1+i)^{-T}\; p^{(\alpha)}(0, T, x, \mathbf{K}(t)). \tag{4}$$

4.2.3. Approximation applied to the life annuity

An issue with applying a stochastic mortality model like the Cairns-Blake-Dowd model is the number of computations required to calculate $\ddot{a}_x^{(\alpha)}(t)$ exactly. To determine how much income to pay out at each time, it is necessary to value a life annuity at each time $t = 0, 1, 2 \ldots$ and along each sample path. However, the calculation of the life annuity value at each time t and in each sample path requires itself a numerical simulation; these are called nested simulations. The number of computations required is extremely large and very time-consuming.

To avoid waiting a long time, an approximation to the life annuity value (4) at each time t and in each sample path is used. As discussed and shown in Cairns (2011), the expression for $p^{(1)}(t, t+T, x-t, \mathbf{K}(t))$ can be approximated well using a quadratic approximation. The approximation is based on a second-order Taylor expansion of a function of the survival probabilities. It is shown for the more general function $p^{(\alpha)}(t, t+T, x-t, \mathbf{k})$ in Appendix A, applying the approach of Cairns (2011).

Having presented the economic and stochastic mortality model, turn to the membership profile and the dynamics of the participants' fund values.

4.3. Participants in the fund

Suppose there are $N+1$ cohorts who join the fund, with $N \in \mathbb{N}$. Cohort n joins at integer time n when they are integer age $x > 0$ years old, for $n = 0, 1, \ldots, N$.

The future lifetimes of the participants are independent random variables. The mortality of each survivor follows the Cairns-Blake-Dowd mortality model, detailed above. Thus the mortality distribution faced by each survivor depends on when they joined the fund, as well as their current age.

4.4. Participants' fund value and income payments

Each new participant brings the same amount of pension savings to the fund, of constant amount $f(0) > 0$ units, when they join the fund. A fraction α of each participant's pension savings, or fund value, is allocated to their tontine account, to give $\alpha f(0)$ units in their tontine account at time 0. The remainder, $(1-\alpha) f(0)$ units, is allocated to their bequest account.

The fund value changes over time as investment returns and longevity credits are earned and as income is paid out. Investment returns and longevity credits are added to the fund value at the end of each year, just before income payments are made. The fund value of a survivor in Cohort n is modelled by a stochastic process, which has value $F^{(n)}(t)$ at time $t \geq n$, for $n = 0, 1, \ldots, N$.

The amount of income paid out at time n to a participant in Cohort n, when they first join the fund, is $C^{(n)}(n) = f(0)/\ddot{a}_x^{(\alpha)}(n)$, leaving fund value $F^{(n)}(n) = f(0) - C^{(n)}(n)$. The income is withdrawn from both the tontine account and bequest account so that the $\alpha : (1-\alpha)$ ratio between them is maintained. Investment returns are earned at the effective rate i per annum on the entire fund value. The calculation of the longevity credit, $M^{(n)}(t)$, the amount earned at time $t \geq n$ is detailed in Section 4.5.1.

At integer time $t > n$, the income paid to a survivor of Cohort n, who is age $x + t - n$, is calculated as

$$C^{(n)}(t) = \frac{(1+i)\, F^{(n)}(t-1) + M^{(n)}(t)}{\ddot{a}_{x+t-n}^{(\alpha)}(t)}. \tag{5}$$

Then the fund value at time $t > n$ of each survivor in Cohort n is

$$F^{(n)}(t) = (1+i)\, F^{(n)}(t-1) + M^{(n)}(t) - C^{(n)}(t).$$

Each survivor's fund account is re-balanced at the end of year, after the payment of any longevity credit, to maintain the $\alpha : (1-\alpha)$ ratio between the tontine account and bequest account (this is the same assumption on which the expression for the life annuity value $\ddot{a}_{x+t-n}^{(\alpha)}(t)$ rests).

No income is paid out from integer time t to a participant who dies over the time period $(t-1, t]$. Instead, the amount in their bequest account at time t, $(1-\alpha)\,(1+i)\, F^{(n)}(t-1)$ is paid to their estate. The amount in their tontine account, $\alpha\,(1+i)\, F^{(n)}(t-1)$, is shared out among the surviving participants, as detailed in Section 4.5.1, as a longevity credit.

4.5. Longevity risk-sharing rules

The next step is to allocate the funds of those who have died among the survivors, in the form of a payment called a longevity credit. It is established that such survivor-only rules are only actuarially fair (at least, in a theoretical setting) if some restrictions are placed on the membership (Sabin, 2010). However, the actuarially-fair rule of Sabin (2010) is computationally challenging to implement. While dropping the requirement that only survivors receive a longevity credit leads to a straightforward rule to implement (Donnelly et al., 2014), it means paying some money to the estate of the recently deceased. Consequently, longevity credits are lower, meaning that survivors have lower fund values. The position in this paper is that it is more attractive to maximise the survivors' income than worry about actuarial fairness holding precisely.

The rule used here is that of Qiao and Sherris (2013). It has two main advantages: (i) it is straightforward and easy to calculate compared to that of Sabin (2010); and (ii) it avoids restrictions on the membership which are implicit in the application of the rule in Sabin (2010). While the rule of Qiao and Sherris (2013) is not perfectly actuarially fair, it is approximately so. It is shown in Donnelly (2015) that, for its antecedent rule detailed in Piggott et al. (2005), this only matters when the membership is either very small or highly heterogeneous. The argument for accepting only approximate actuarial fairness instead of laboratory perfection is that, it holds well enough when the fund is pooling longevity risk sufficiently well. If the fund is not pooling longevity risk well enough then other risk management procedures should be done, since the fund is not doing what it should be doing: pooling longevity risk effectively.

4.5.1. Longevity credits via Qiao and Sherris (2013) risk-sharing rule

The number of participants in each cohort at the time they join is a constant integer $\ell > 0$. The number of surviving participants in Cohort n at integer time $t \geq n$ is denoted by the random variable $L^{(n)}(t) \geq 0$, with $L^{(n)}(n) = \ell$, for $n = 0, 1, \ldots, N$.

As only the tontine account of a participant is at risk of loss, through the death of the participant, only the tontine account values are used in the calculation of the longevity credit. For example, each participant in Cohort k who is alive at time $t-1$, risks losing the value of their tontine account $(1+i) \times \alpha F^{(k)}(t-1)$ at time t, with the materlialisation of the risk arising on the event that they die in the time interval $(t-1, t]$.

Let $G(t)$ represent the total value of all the members' tontine accounts at integer time t, for the members who were alive and in the fund at time $t-1$ (i.e. excluding any new members who join at time t and those who died before time $t-1$), and just before any income is withdrawn or longevity credits allocated, i.e. for $t = 1, 2, \ldots$,

$$G(t) = \sum_{k=0}^{t-1} L^{(k)}(t-1) \times (1+i) \times \alpha F^{(k)}(t-1).$$

The value of $G(t)$ represents the total sum-at-risk over the time period $(t-1, t]$, for those who are alive at time $t-1$. The calculation of the longevity credit to each survivor at time t is a two-step calculation. First, the value of $G(t)$ is notionally divided among those who survive to time t, in line with the rule set out in Qiao and Sherris (2013, equation (4)). Each survivor is allocated a fraction of $G(t)$ which is proportional to the product of their own tontine account value and the inverse of their probability of surviving from time $t-1$ to time t. Second, the longevity credit paid to each survivor is obtained by deducting each survivor's tontine account value from this share of the total sum-at-risk. Formally, the longevity credit is credited to the survivor's tontine account.

For notational brevity, denote the survival probability of a member in Cohort n surviving from time $t-1$ to time t as

$$p_{t-1}^{(n)} := p^{(1)}(t-1, t, x-n, \mathbf{K}(t-1)), \qquad \text{for } n = 0, 1, \ldots, t-1 \text{ and } t = 1, 2, \ldots,$$

in which the right-hand side is calculated from (3) with $\alpha := 1$. The survival probability $p_{t-1}^{(n)}$ is calculated at time $t-1$ and is the prediction of the one-year survival probability from time $t-1$ to time t, for someone age $x-n$ at time 0. It does not involve α. It is interpreted as the probability of a participant in Cohort n, who is alive at time $t-1$, keeping their tontine account, by surviving for another year until time t.

Based on the risk-sharing rule Qiao and Sherris (2013, equation (4)), the longevity credit awarded to a surviving member of Cohort n at integer time $t \in \{n+1, n+2, \ldots\}$ is

$$M^{(n)}(t) = G(t) \times \frac{\frac{\alpha(1+i)F^{(n)}(t-1)}{p_{t-1}^{(n)}}}{\sum_{k=0}^{t-1} L^{(k)}(t) \frac{\alpha(1+i)F^{(k)}(t-1)}{p_{t-1}^{(k)}}} - \alpha(1+i)F^{(n)}(t-1).$$

Substituting for $G(t)$ and simplifying the expression, gives

$$M^{(n)}(t) = \alpha \times (1+i) \times \left(\sum_{k=0}^{t-1} L^{(k)}(t-1) \times F^{(k)}(t-1) \times \frac{\frac{F^{(n)}(t-1)}{p_{t-1}^{(n)}}}{\sum_{k=0}^{t-1} L^{(k)}(t) \frac{F^{(k)}(t-1)}{p_{t-1}^{(k)}}} - F^{(n)}(t-1) \right).$$

A higher amount of longevity credit $M^{(n)}(t)$ is obtained by a higher value of α, since that increases the amount of money in each participant's tontine account and hence increases the amount of money shared out upon any deaths. Further, the more money in the account of a participant from Cohort n, $F^{(n)}(t-1)$, and the more likely they were anticipated to die over the previous time period, $p_{t-1}^{(n)}$, the higher the longevity credit awarded to that participant. This illustrates the fundamental principle underlying the longevity risk-sharing mechanism: the more money a participant stands to lose when they die and the more likely they are to die, the higher longevity credit they receive upon survival.

The amount of longevity credit is volatile, as it depends on who has died in the pool and when they die. As the longevity credit is proportional to α, its volatility decreases with α. In turn, as the fund value is a linear function of the longevity credit,, by a factor of α, as can be seen from the expression for $M^{(n)}(t)$. This feeds into a lower volatility in the income paid out when $\alpha < 1$, which is observed in the simulations.

5. Numerical simulations

With the economic and mortality model specified and the details set out above of how the income paid out to each participant evolves, the results of a numerical simulation of the income payments to a surviving member in the pooled annuity fund are described in this section.

First, the distribution of the income at each fixed time is studied, for each cohort entering the fund. Second, the behaviour of the income along each sample path is analysed, again for each cohort.

5.1. Parameters for the simulations

The fund values of participants earn returns at a constant effective interest rate of $i = 0.02$ per annum. In the multi-cohort setting, it is assumed that 31 cohorts join every year for 30 years with the same number of members in each cohort. Thus Cohort m joins at time m, for $m = 0, 1, 2, \ldots, 30$. The fund is closed to new members after time 30, but the fund continues to pay out income to the surviving participants. One simulation is shown of a single cohort fund, which means that a cohort joins at time 0 and no other join after that. This is done purely to emphasize the importance of pooling across cohorts, which is discussed in Qiao and Sherris (2013), as it is such an important point.

Every participant is assumed to be age 65 years at the time they join the scheme and they each bring pension savings of $f(0) = 100$ units to the pooled annuity fund. Their mortality follows the Cairns-Blake-Dowd model, detailed in Section 4.2, and each participant's future mortality is independent of any other participant's. As the same Cairns-Blake-Dowd model is used to calculate the annuity values and sample the deaths occurring in the group of participants, there is no model error.

The income is calculated via equation (5), but using the approximation of the expected present value of the life annuity value which is described in Appendix A. For each plot, a total of 5 000 simulations were done.

5.2. Analysis of the income distribution at each fixed time

The first remark on the distribution of the income plotted against age is that it has the shape of a cone (for example, see Figure B.2). This is due to the stochastic mortality model, as mortality rates become more uncertain, the further into the future they are projected. The amount of income paid out over time reflects this uncertainty. Roughly speaking, if people live longer than expected, then too much income will have been paid out at earlier ages. This leads to lower fund values among the participants which are not able to sustain the expected payment stream. In turn, this depletion of the fund values leads to lower longevity credits; the fund values of those who have died is smaller than they should be. The longest-lived in the pooled annuity fund thus bear the consequences of people living longer than expected, which crystallises as a decline in their income payment in their old age. This example illustrates the effect of systematic longevity risk, although it may be that people live less than expected or sometimes less and sometimes more. Nonetheless, the longer that a participant survives, the more systematic longevity risk they will bear. Moreover, the later a cohort joins, the more systematic longevity risk they take on too. The consequences of other cohorts' mortality rate being different to that expected, manifests itself through the amount and frequency of the longevity credits received by all participants in the fund.

A second remark is that, due to the stochastic mortality model used, the income paid out to each cohort – except for the first cohort – is a distribution of possible values. In particular, the first income payment to each future cohort is a distribution. This is because the simulations are done from the perspective of knowing only what has happened up to time zero, and not knowing exactly what will happen after then. Conditioning on what has happened at the time each cohort joins will give an income distribution for each cohort which starts at a single value and which is narrower than the unconditioned distribution. The analysis of the behaviour of the income along each sample path does the conditioning, whereas the analysis of the income distribution at each fixed time does not.

General findings

Begin with the main findings from studying the income paid out in a pooled annuity fund. While these findings apply whether there is a bequest feature or not, they are discussed in the context of the fund without a bequest ($\alpha = 1$).

- The benefits of an open fund in terms of reducing idiosyncratic longevity risk at older ages, is seen with relatively small cohort sizes (Figure B.1). For example, only about 50 members are needed to join each year, to get an income distribution which is narrower from age 90 compared to a single cohort fund of 10 000, for the first cohort to join. A cohort size of 100 members gives a narrower distribution from about age 85. The income distribution is wider in the multi-cohort funds below these ages, but not significantly so compared to the reduction in the range of income values above the ages.

  The reason is due to the benefit of pooling the longevity risk of the older participants – who, in effect, have high levels of diversifiable longevity risk but are fewer in number by virtue of having died off – with the younger participants – who have low levels of diversifiable longevity risk but are more numerous by virtue of being less likely to die. Qiao and Sherris (2013) discuss this, but use a much higher number of members in each cohort to demonstrate that a multi-cohort fund is better than an single cohort fund in this regard. The results in this paper extend this demonstration, to show that only a relatively small cohort size is needed in the multi-cohort fund.

- The income distribution varies across cohorts, and is widest at the oldest ages of the last cohort (Figure B.2). The first cohorts have relatively few people with whom to pool their longevity risk. However, this is not a significant problem as their chance of dying is relatively low and so their expected longevity credit is small. The distribution of their income at younger ages is wider than for later cohorts, but not by much.

As the first cohorts become older and older, more and more cohorts have joined the fund. As they age, the first cohorts have plenty of people with whom to pool their longevity risks. Their income distribution is the narrowest.

However, from about time 14 onwards in this model fund, the income distribution of Cohort 15, and higher cohorts, begins widening at the oldest ages. The survivors of these later cohorts have fewer people with whom to pool longevity risk when they reach old age, as no new entrants join after time 30. The increase in idiosyncratic longevity risk leads to a wider range of possible income values.

The last cohort has the widest income distribution when they are old, for two reasons. First, there are few people alive with whom they can pool their longevity risk. Second, they bear the most systematic longevity risk of all the cohorts, as discussed above.

The difference in the income distribution, and hence the income volatility, across cohorts may have implications for studies in which, for example, a comparison is made of whether a pooled annuity fund is better (in some defined way, for example, gives a higher expected utility of lifetime income) than other income-generating alternatives, such as a guaranteed life annuity contract. The last cohorts may prefer a guaranteed life annuity contract, to avoid the high income volatility when they are old, whereas the earlier cohorts may prefer a pooled annuity fund, since they experience a much lower income volatility at old ages.

Results for the pooled annuity fund with integrated bequest

Now turn to the pooled annuity fund with integrated bequest, with half of each participant's fund value allocated to their tontine account (i.e. $\alpha = 0.5$). In this fund with a 50% bequest, half of each participant's fund value is paid to their estate upon their death. The remaining proportion is shared out among the living fund participants, as described in Section 4.5.1.

Figure B.3 shows the distribution of the amount of money in the bequest account, as participants in selected cohorts age. In this strucutre, the bequest account is simply a fixed proportion of a surviving participant's fund value.

In amount of bequest goes down over time, reflecting that fact that income is withdrawn over time from the bequest account, thus depleting the surviving participants' fund values. The declining bequest may be attractive to those who wish to have a higher income while they are alive and leave something to their estate, but they are not fussy about how much. The uncertainty about the future bequest amount increases over time, reflecting the uncertainty inherent in the stochastic mortality model used. Mirroring the income distribution, the last cohort faces the most variability in their bequest account in extreme old age.

The income paid out is lower in the fund with a bequest, as it costs more to provide the same level of income (Figure B.4 versus Figure B.2). However, the development of the income distribution in the fund with a 50% bequest for each cohort follows a similar pattern to that in a fund without a bequest. This shows that, as long as an appropriate life annuity value is used, the addition of the bequest does not result in different general conclusions. However, there is less variability in the income paid out in the fund with a bequest, as a consequence of receiving lower longevity credits. This is discussed in the next section, when the income paid out in each sample path is analysed.

5.2.1. Analysis of the distribution of each sample path

In Bernhardt and Donnelly (2020), the probability of the income sample paths between two bounds was used as a risk measure. It is applied here, to see for how long the income stays above a specified fraction of the initial income.

It is assumed that the initial income paid out to each cohort of participants in the fund is psychologically important. People are likely to think "that is what I will get for life" when they are told their initial income, even if they are warned that it may go down. This may be referred to as a cognitive bias called anchoring, or it may be a confidence in the manager of the pooled annuity fund, that they will deliver what is promised. Additionally, it is likely that this type of risk measure is communicated to participants, to indicate to them that the income paid from the fund may go down as well as up. For example, participants may be told that their income is unlikely to drop by more than, for example, 10% of their initial income most of the time.

Probability of the income staying above a lower bound

Recall that $C^{(n)}(t)$ is the amount of income paid out at integer time $t \geq n$ to each survivor in Cohort n. The probability of the income paid to each survivor in Cohort n staying above a constant fraction $\gamma \in (0, 1]$ of their initial income, $C^{(n)}(n)$, for an integer $T \geq 0$ years from time n is

$$r^{(n)}(T; \gamma) := \mathbb{P}\left[C^{(n)}(t) \geq \gamma C^{(n)}(n), \ \text{ for } t = n, n+1, \ldots, n+T\right], \quad \text{for } n = 0, 1, \ldots, M.$$

For example, if the lower bound is 90% of the initial income and $r^{(0)}(20; 0.9) = 98\%$ then there is a 98% chance that the income paid to Cohort 0 does not drop below 90% of the initial income for the first 20 years of payment. In the numerical study, $r^{(n)}(T; \gamma)$ is estimated as the proportion of the simulations in which the income stays at or above the amount $\gamma C^{(n)}(n)$, from time n, the time that Cohort n joins the fund, until time $n + T$.

Figure B.5 shows the estimated value of $r^{(n)}(T; \gamma)$ for Cohorts $n \in \{0, 15, 30\}$, as a function of T, for different values of γ. The first cohort to join the fund, Cohort 0 has generally a superior outcome, compared to Cohort 15 and Cohort 30, in terms of their income staying above the lower bound for longer. This emphasizes the importance of pooling longevity risk at higher ages. Cohort 0 is the cohort with the most number of other people in the fund when they are old. In contrast, being the cohort with the fewest number of people in the fund when they are young does not seem to affect their income stability in any significant way.

The left-hand plots of Figure B.5 show that there is a less than 15% chance of the income falling by more than 10% over a 20 year period since each cohort joins, in the fund without a bequest. There is an extremely low probability of falls of more than 20% in this 20-year time frame. Falls of more than 5% are more likely, with around a 60% to 70% chance, depending on the cohort.

The right-hand plots of Figure B.5 are for the fund with a 50% bequest. The income is less likely to experience falls below the lower bound in this fund, since only the tontine account earns the random, volatile longevity credits. Due to the lower volatility, the first time at which there is a non-zero chance of the income dropping below each bound is extended by about 5 years for the fund with a 50% bequest. Additionally, as this first time is pushed out farther, the impact of Cohort 0 of having few people with whom to pool their longevity risk when they are young is slightly reduced.

Number of participants required in each cohort

A different representation of the analysis of the simulations allows an easy view of how many participants are needed to join each year, to adequately pool longevity risk.

Figure B.6 plots the estimated values of $r^{(n)}(30; \gamma)$, the probability of falling below against the number of participants joining each year, against the number of participants in each new cohort. This is done for Cohorts $n \in \{0, 15, 30\}$ and for different values of γ. Similar plots are obtained over shorter or longer time frames, with the lines shifted upwards for shorter time periods and downwards over longer time periods.

Broadly, there are only small increases in the probability of the income falling below the given lower bound over 30 years, when the cohort size is around 50 to 100 members. However, this is not true for the last cohort, Cohort 30. For that cohort, since there are fewer participants in the entire fund when they are age 95 as the fund closes immediately after they enter it at age 65, they benefit from larger cohort sizes. Indeed, the slope of the probability lines gradually increases from Cohort 25 through to Cohort 30.

Figure B.6 also allows a different view of the systematic longevity risk in the model. Even with a cohort size of 10 000 members, there is a 40-50% chance of the income falling by more than 5%.

6. Conclusion

Pooled annuity funds offer an opportunity to give pension savers a lifetime income. As many retirees also value leaving a bequest to their heirs, a pooled annuity fund which combines the delivery of a lifetime income and a bequest is studied. The number of members needed to join each year and the stability of the income paid out are key indicators to support the practical implementation of these funds. In particular, these factors have not been studied for the pooled annuity fund with bequest.

The results indicate that only about 100 members are needed to join each year, in order to adequately pool longevity risk. This number is the same whether there is a bequest or not. The relatively lor number suggests that pooled annuity fund can be offered to a sufficiently large workforce rather than, for example, at a national level.

While the fund remains open, the income paid out is stable for decades, under the assumption of no model error. There is volatility in the income paid out and falls of more than 5% should be expected. However, the income is unlikely to drop by more than 10% and extremely unlikely to drop by more than 20%.

However, once the fund closes to new members, the cohorts left in the fund experience an increasing chance of their income falling as fewer people are left in the fund, leading to insufficient pooling of idiosyncratic longevity risk. Additionally, the last cohorts bear the cumulative effects of systematic longevity risk, as the result of too much or too little income being paid out over time to earlier cohorts. They have a much higher chance of higher falls in their income, and this chance increases the later than a cohort joins a closed fund.

This leads to questions for the later cohorts, such as the right balance between the benefits of pooling longevity risk versus bearing the cumulative effects of systematic longevity risk. For example, is there an optimal time of the fund to stay open to new members? What risk management should be put in place, so that the later cohorts are not disadvantaged compared to the earlier cohorts?

References

Aitchison, J. and Shen, S. M. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika*, 67(2):261–272.

Bernhardt, T. and Donnelly, C. (2019). Modern tontine with bequest: Innovation in pooled annuity products. *Insurance: Mathematics and Economics*, 86:168–188.

Bernhardt, T. and Donnelly, C. (2020). Quantifying the trade-off between income stability and the number of members in a pooled annuity fund. *ASTIN Bulletin*, 51:101–130.

Cairns, A. (2011). Modelling and management of longevity risk: Approximations to survivor functions and dynamic hedging. *Insurance: Mathematics and Economics*, 49:438–453.

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

Chen, A., Guillen, M., and Rach, M. (2021). Fees in tontines. *Insurance: Mathematics and Economics*, 100:89–106.

Chen, A. and Rach, M. (2022). Bequest-embedded annuities and tontines. *Asia-Pacific Journal of Risk and Insurance*, 16(1):1–46.

Chen, A., Rach, M., and Sehner, T. (2020). On the optimal combination of annuities and tontines. *ASTIN Bulletin*, 50(1):95–129.

Dagpunar, J. (2021). Closed-form solutions for an explicit modern ideal tontine with bequest motive. *Insurance: Mathematics and Economics*, 100:261–273.

Denuit, M. and Robert, C. (2021). From risk sharing to pure premium for a large number of heterogeneous losses. *Insurance: Mathematics and Economics*, 96:116–126.

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

Donnelly, C. and Bernhardt, T. (2021). Staying the course: how pooled annuity funds are proving an attractive alternative. *The Actuary - The Magazine of the Institute and Faculty of Actuaries*. `https://www.theactuary.com/2021/04/30/staying-course-how-pooled-annuity-funds-are-proving-attractive-alternative`.

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

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

Hanewald, K., Piggott, J., and Sherris, M. (2013). Individual post-retirement longevity risk management under systematic mortality risk. *Insurance: Mathematics and Economics*, 52(1):87–97.

Milevsky, M. and Salisbury, T. (2015). Optimal retirement income tontines. *Insurance: Mathematics and Economics*, 64:91–105.

Milevsky, M. A. and Salisbury, T. S. (2016). Equitable retirement income tontines: Mixing cohorts without discriminating. *ASTIN Bulletin*, 46(3):571–604.

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.

Qiao, C. and Sherris, M. (2013). Managing systematic mortality risk with group self-pooling and annuitization schemes. *Journal of Risk and Insurance*, 80(4):949–974.

Sabin, M. (2010). Fair tontine annuity. Available at SSRN. `http://ssrn.com/abstract=1579932`.

Stamos, M. Z. (2008). Optimal consumption and portfolio choice for pooled annuity funds. *Insurance: Mathematics and Economics*, 43:56–68.

Valdez, E. A., Piggott, J., and Wang, L. (2006). Demand and adverse selection in a pooled annuity fund. *Insurance: Mathematics and Economics*, 39(2):251–266.

Weinert, J. and Gründl, H. (2021). The modern tontine. *European Actuarial Journal*, 11(2):49–86.

Zhou, Y., Sherris, S., Ziveyi, J., and Xu, M. (2021). An innovative design of flexible, bequest-enhanced life annuity with natural hedging. *Scandinavian Actuarial Journal*, 0(0):1–22.

A. Taylor-series approximation to $p^{(\alpha)}(t, t+T, x-t, \mathbf{k})$

Life annuity value to calculate

Define for $t =\in \mathbb{N}_0$ and $T \in \mathbb{N}$, the random variable

$$p^{(\alpha)}(t, t+T, x-t, \mathbf{K}(t)) := \mathbb{E}\left(\prod_{k=0}^{T-1} \frac{1-q(t+k, x+k)}{(1-(1-\alpha)q(t+k, x+k))} \,\middle|\, \mathbf{K}(t) \right). \tag{6}$$

The goal is to calculate, at each time t and along each sample path, the annuity value

$$\ddot{a}_x^{(\alpha)}(t, \mathbf{K}(t)) = 1 + \sum_{T=1}^{\infty} (1+i)^{-T} \; p^{(\alpha)}(t, t+T, x-t, \mathbf{K}(t)).$$

From the Markov and time-homogeneous nature of the random walk $\mathbf{K}$, $p^{(\alpha)}(t, t+T, x-t, \mathbf{k}) = p^{(\alpha)}(0, T, x, \mathbf{k})$, as noted in Cairns (2011) for the case $\alpha = 1$. Thus

$$\ddot{a}_x^{(\alpha)}(t, \mathbf{K}(t)) = 1 + \sum_{T=1}^{\infty} (1+i)^{-T} \; p^{(\alpha)}(0, T, x, \mathbf{K}(t)). \tag{7}$$

Without using any approximations, determining the annuity value at each time t along each sample path requires the numerical calculation of $\{p^{(\alpha)}(0, T, x, \mathbf{K}(t)) : T \in \{0, 1, 2, \ldots, \}\}$. This means numerically calculating the right-hand-side of equation (6) via simulation using the salient value of the stochastic process $\mathbf{K}(t)$. To avoid this, a Taylor series approximation is made of $p^{(\alpha)}(t, t+T, x-t, \mathbf{K}(t))$.

Taylor series approximation to $p^{(\alpha)}(t, t+T, x-t, \mathbf{k})$

Fixing x and T, the idea of Cairns (2011) is to calculate only one value of $p^{(1)}(0, T, x, \mathbf{K}(t))$ via simulation at each time t, instead of at every sample value of $\mathbf{K}(t)$. For this single calculation, $\mathbf{K}(t)$ is set equal to its expected value. Then each $\{p^{(1)}\,(0, T, x, \mathbf{K}(t))\}$ is estimated using a Taylor series approximation about the expected value of $\mathbf{K}(t)$. The approximation avoids having to do a full numerical simulation in each sample path to calculate $p^{(1)}(0, T, x, \mathbf{K}(t))$, which is itself a conditional expectation. Here, the same technique is applied to $p^{(\alpha)}(0, T, x, \mathbf{K}(t))$.

Define $f(T, x, \mathbf{k}) = \Phi^{-1}\left(p^{(\alpha)}(0, T, x, \mathbf{k})\right)$, where Φ^{-1} is the inverse of the standard normal distribution function. Define

$$\hat{\mathbf{k}} = (\hat{k}_1, \hat{k}_2)' = \mathbb{E}\left(\mathbf{K}(T)\right) = \mathbf{K}(0) + T \cdot \boldsymbol{\nu}.$$

Then the second-order approximation used here is

$$f(T, x, \mathbf{k}) \approx D_0(T, x) + \mathbf{D_1}'(T, x)\left(\mathbf{k} - \hat{\mathbf{k}}\right) + \frac{1}{2}\left(\mathbf{k} - \hat{\mathbf{k}}\right)' \mathbf{D_2}(T, x)\left(\mathbf{k} - \hat{\mathbf{k}}\right),$$

in which the derivatives are

$$D_0(T, x) = f(T, x, \hat{\mathbf{k}}),$$

$$D_{1,i}(T, x) = \frac{\partial f}{\partial k_i}(T, x, \mathbf{k})\bigg|_{\mathbf{k}=\hat{\mathbf{k}}}, \quad \text{for } i = 1, 2, \quad \mathbf{D_1}(T, x) = \begin{pmatrix} D_{1,1}(T, x) & D_{1,2}(T, x) \end{pmatrix}$$

$$D_{2,ij}(T, x) = \frac{\partial^2 f}{\partial k_i \partial \hat{k}_j}(T, x, \mathbf{k})\bigg|_{\mathbf{k}=\hat{\mathbf{k}}}, \quad \text{for } i = 1, 2 \text{ and } j = 1, 2, \quad \mathbf{D_2}(T, x) = \begin{pmatrix} D_{2,11}(T, x) & D_{2,12}(T, x) \\ D_{2,21}(T, x) & D_{2,22}(T, x) \end{pmatrix}.$$

To determine the derivatives, the first step is to calculate $p^{(\alpha)}(0, T, x, \mathbf{k})$ via simulation. This means starting the random walk with dynamics as shown in equation (2) but with starting value $\mathbf{K}(0) = \hat{\mathbf{k}}$. 2 000 sample paths of the random walk were simulated, for each starting value $\mathbf{K}(0) = \mathbb{E}(\mathbf{K}(t))$, $t = 0, 1, \ldots$. For perturbations of the starting value, the same simulations of the random walk can be re-used but with perturbation added onto each sample of the random walk.

Numerical calculation method of derivatives

The derivatives are calculated numerically. This is done by defining $\mathbf{h_1} = \begin{pmatrix} 0.01 & 0 \end{pmatrix}'$ and $\mathbf{h_2} = \begin{pmatrix} 0 & 0.01 \end{pmatrix}'$. Then for the first derivatives,

$$D_{1,1}(T,x) = \frac{f(T,x,\hat{\mathbf{k}}+\mathbf{h_1}) - f(T,x,\hat{\mathbf{k}})}{|\mathbf{h_1}|}$$

and

$$D_{1,1}(T,x) = \frac{f(T,x,\hat{\mathbf{k}}+\mathbf{h_2}) - f(T,x,\hat{\mathbf{k}})}{|\mathbf{h_2}|}.$$

Similarly for the second derivatives,

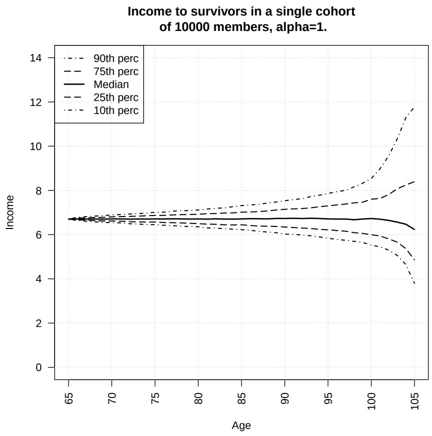
$$D_{2,11}(T,x) = \frac{f(T,x,\hat{\mathbf{k}}+\mathbf{h_1}) - 2f(T,x,\hat{\mathbf{k}}) + f(T,x,\hat{\mathbf{k}}-\mathbf{h_1})}{|\mathbf{h_1}|^2},$$

$$D_{2,22}(T,x) = \frac{f(T,x,\hat{\mathbf{k}}+\mathbf{h_2}) - 2f(T,x,\hat{\mathbf{k}}) + f(T,x,\hat{\mathbf{k}}-\mathbf{h_2})}{|\mathbf{h_2}|^2}$$

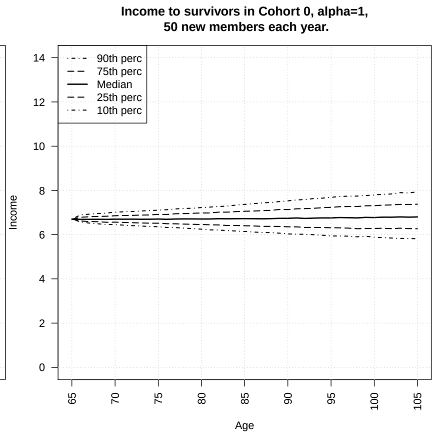
$$D_{2,12}(T,x) = \frac{f(T,x,\hat{\mathbf{k}}+\mathbf{h_1}+\mathbf{h_2}) - f(T,x,\hat{\mathbf{k}}-\mathbf{h_1}+\mathbf{h_2}) - f(T,x,\hat{\mathbf{k}}+\mathbf{h_1}-\mathbf{h_2}) + f(T,x,\hat{\mathbf{k}}-\mathbf{h_1}-\mathbf{h_2})}{4|\mathbf{h_1}||\mathbf{h_2}|}$$

and $D_{2,12}(T,x) = D_{2,21}(T,x)$. The values of the derivatives can be stored and used as required for simulations of the income paid out from the pooled annuity fund with integrated bequest.

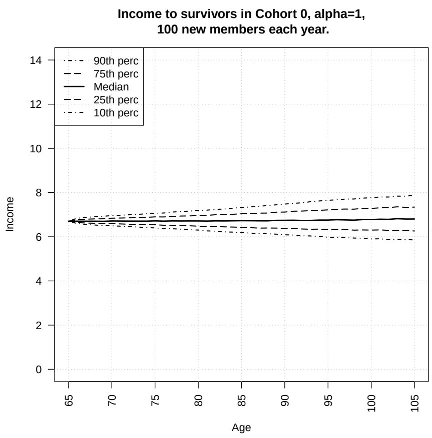
B. Figures



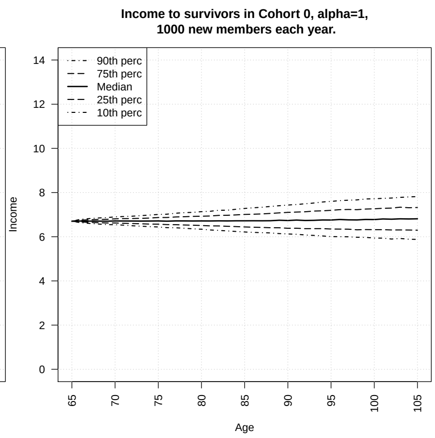
(a) Single cohort of 10 000 members



(b) 50 new members each year, Cohort 0

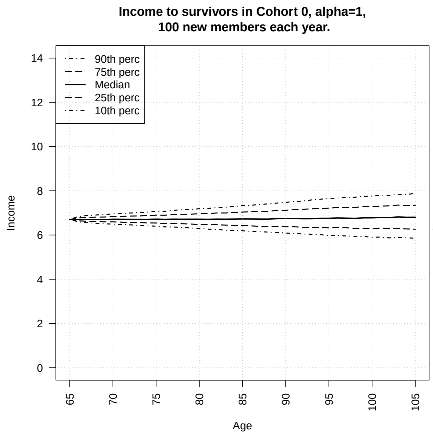


(c) 100 new members each year, Cohort 0

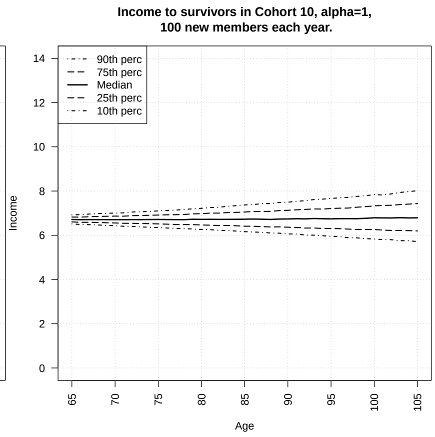


(d) 1 000 new members each year, Cohort 0

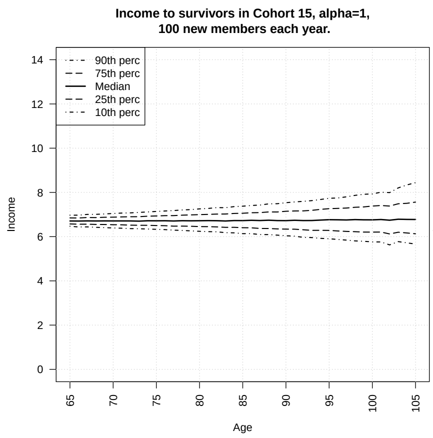
Figure B.1: Benefit payments to each surviving member of the first cohort, plotted against age, in a pooled annuity fund with no bequest feature ($\alpha = 1$). The top-left plot is for a single cohort of 10 000 participants and the remaining plots relate to fund in which a cohort joins every year for 30 years, with each cohort the same size, as indicated in the caption below each chart. All members of each cohort join at age 65 with $f(0) = 100$ units.



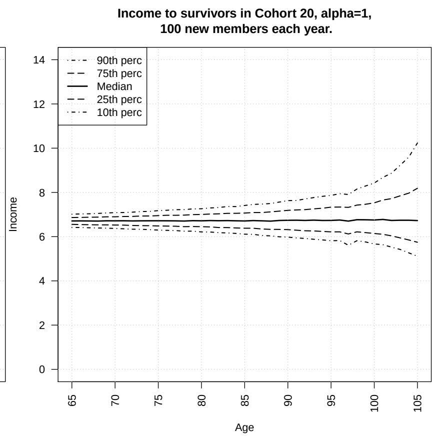
(a) 100 new members each year, Cohort 0



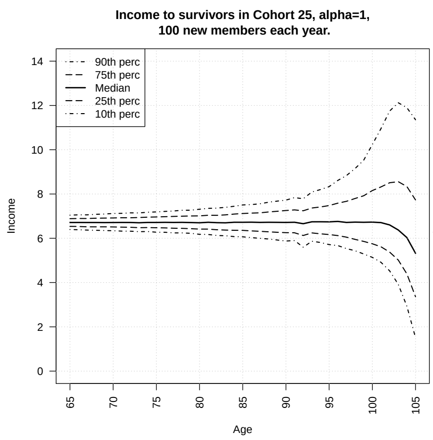
(b) 100 new members each year, Cohort 10



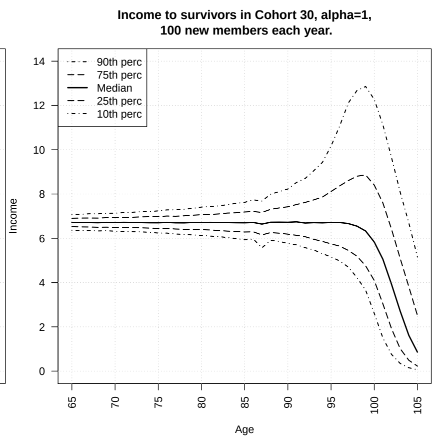
(c) 100 new members each year, Cohort 15



(d) 100 new members each year, Cohort 20

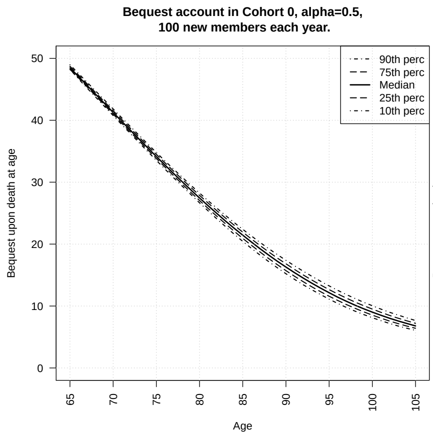


(e) 100 new members each year, Cohort 25

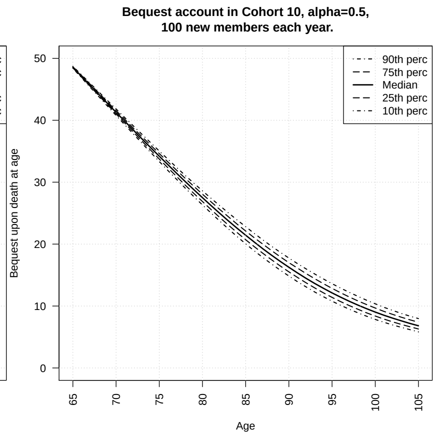


(f) 100 new members each year, Cohort 30

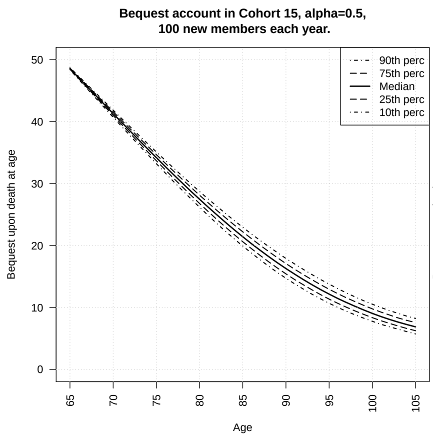
Figure B.2: Benefit payments at each age from a pooled annuity fund with no bequest feature ($\alpha = 1$), plotted against age. A cohort joins the fund every year for 30 years, with each cohort the same size, as indicated in the caption below each chart, and all members of each cohort joining at age 65 with 100 units.



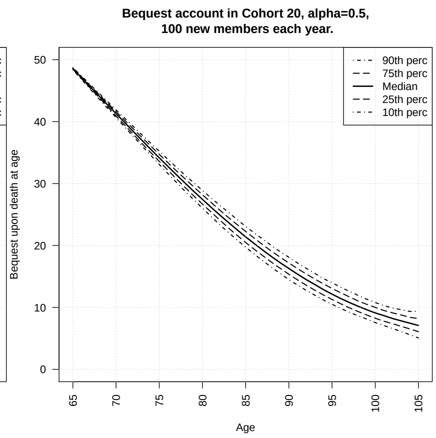
(a) 100 new members each year
Cohort 0



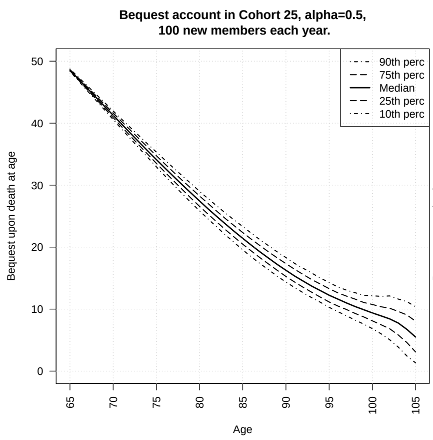
(b) 100 new members each year
Cohort 10



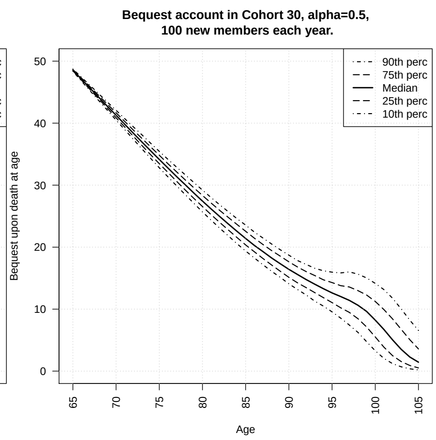
(c) 100 new members each year
Cohort 15



(d) 100 new members each year
Cohort 20

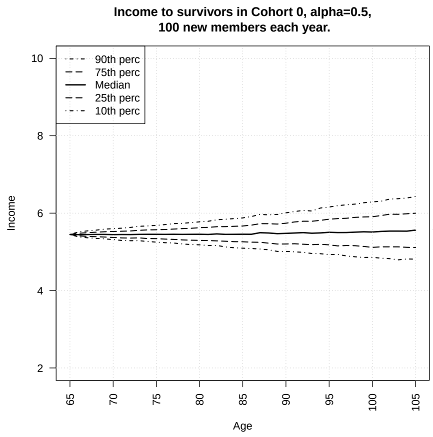


(e) 100 new members each year
Cohort 25

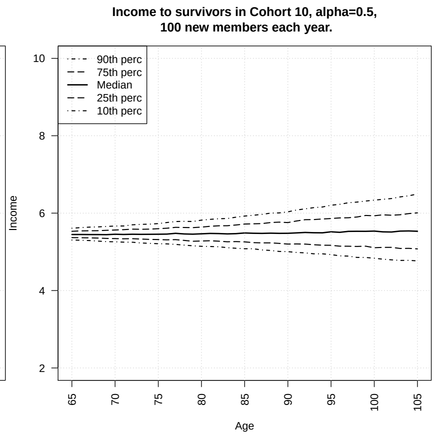


(f) 100 new members each year
Cohort 30

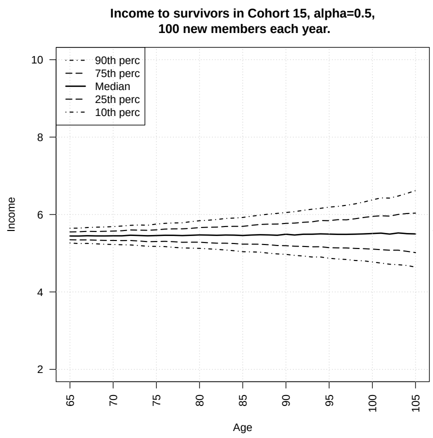
Figure B.3: Bequest account value at each age in a pooled annuity fund with integrated bequest feature, with a $\alpha = 50\%$ allocation to the tontine account, plotted against the age. A cohort joins the fund every year for 30 years, with each cohort of size 100, and all members of each cohort joining at age 65 with 100 units.



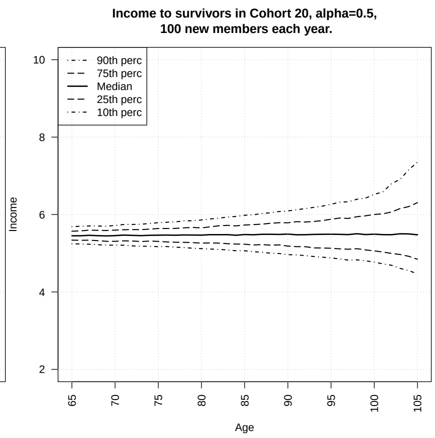
(a) 100 new members each year
Cohort 0



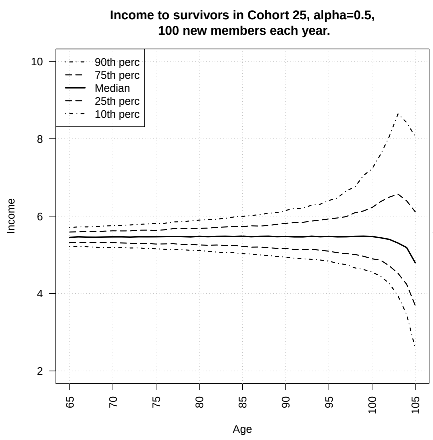
(b) 100 new members each year
Cohort 10



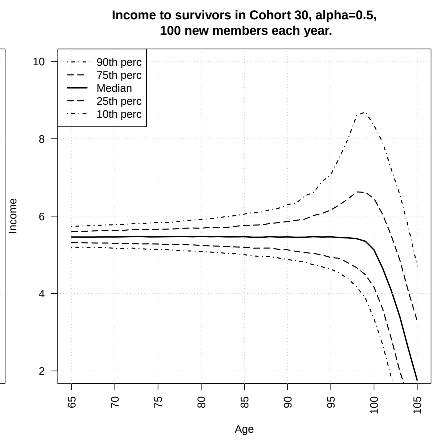
(c) 100 new members each year
Cohort 15



(d) 100 new members each year
Cohort 20

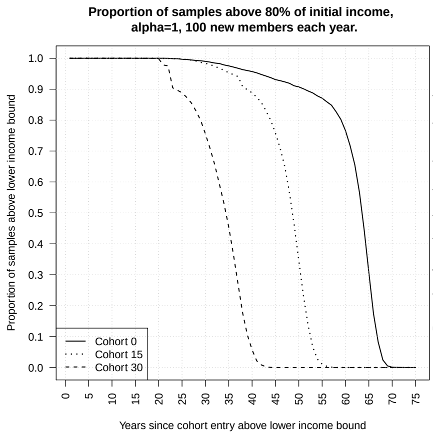


(e) 100 new members each year
Cohort 25

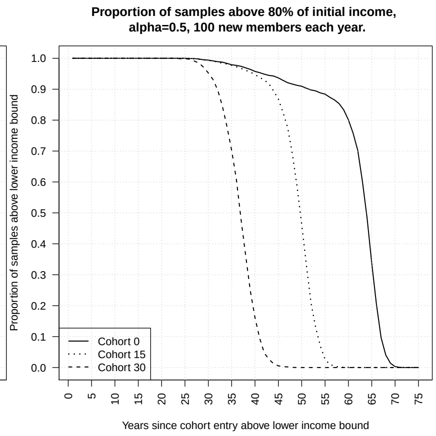


(f) 100 new members each year
Cohort 30

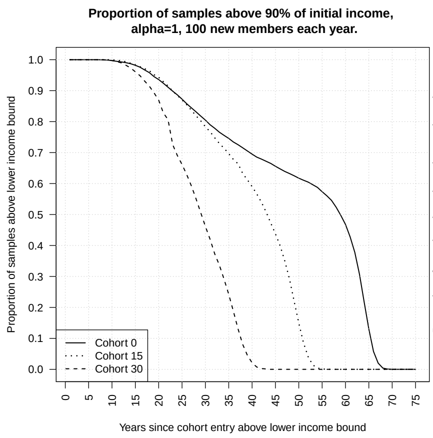
Figure B.4: Benefit payments at each age from a pooled annuity fund with integrated bequest feature, with a $\alpha = 50\%$ allocation to the tontine account, plotted against age. A cohort joins the fund every year for 30 years, with each cohort of size 100, and all members of each cohort joining at age 65 with 100 units.



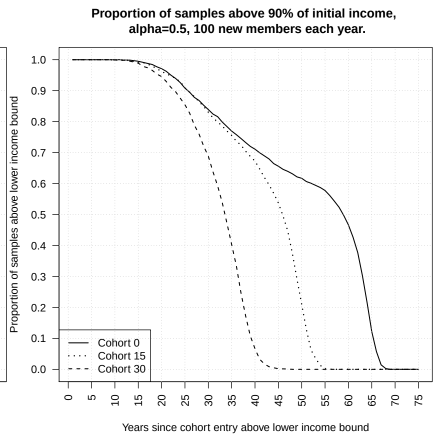
(a) 100 new members each year
$\alpha = 1$, $\gamma = 0.8$



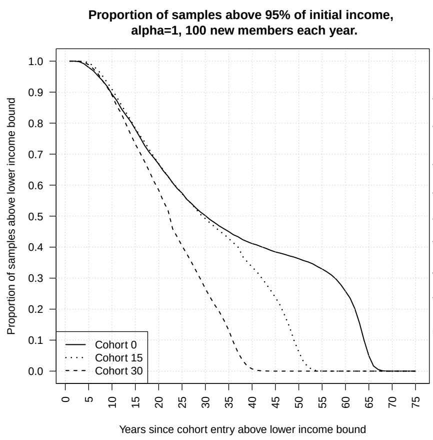
(b) 100 new members each year
$\alpha = 0.5$, $\gamma = 0.8$



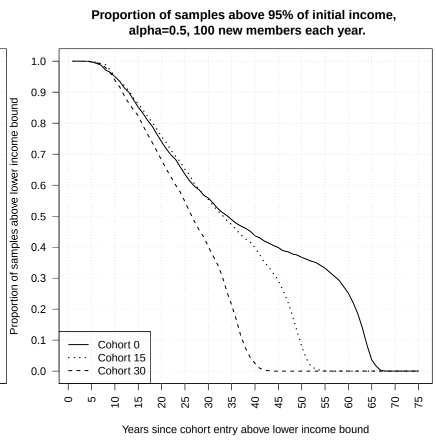
(c) 100 new members each year
$\alpha = 1$, $\gamma = 0.9$



(d) 100 new members each year
$\alpha = 0.5$, $\gamma = 0.9$

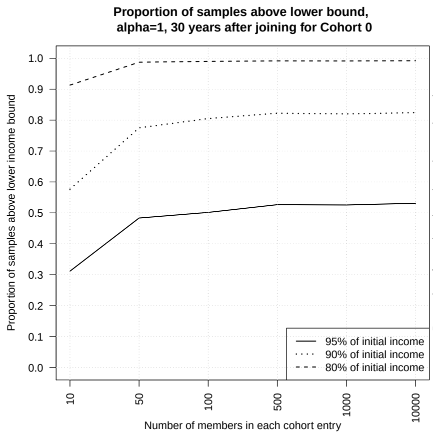


(e) 100 new members each year
$\alpha = 1$, $\gamma = 0.95$

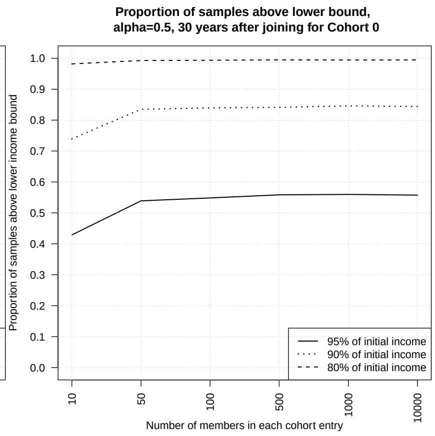


(f) 100 new members each year
$\alpha = 0.5$, $\gamma = 0.95$

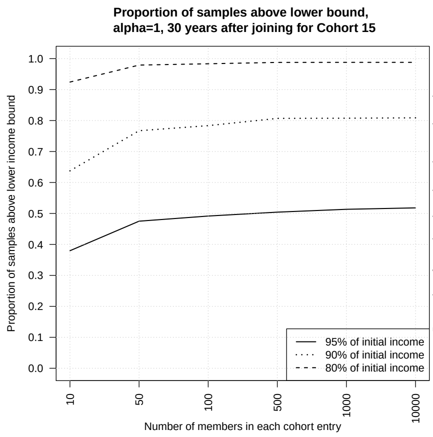
Figure B.5: The plots show the proportion of simulations in which the income paid out is above a specified fraction γ of the initial income for T years, plotted against T, for three different cohorts. The left-hand plots are within a fund with no bequest ($\alpha = 1$) and the right-hand plots are within a fund with a 50% bequest ($\alpha = 0.5$). 100 members join each year for 30 years.



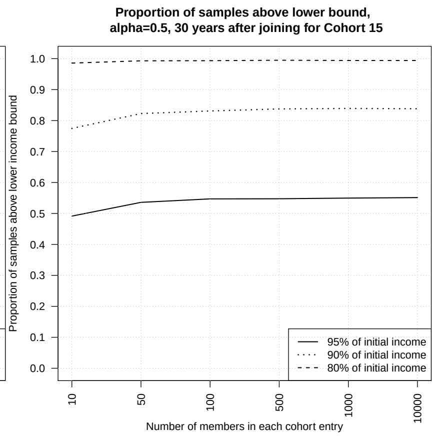
(a) Cohort 0, $\alpha = 1$



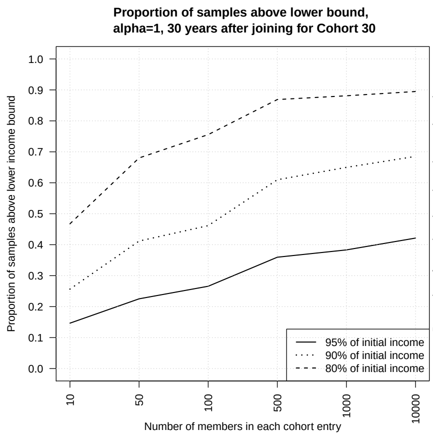
(b) Cohort 0, $\alpha = 0.5$



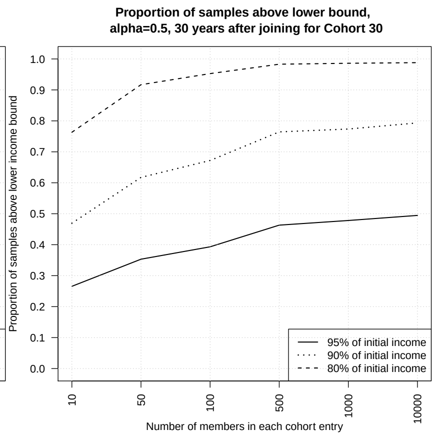
(c) Cohort 15, $\alpha = 1$



(d) Cohort 15, $\alpha = 0.5$



(e) Cohort 30, $\alpha = 1$



(f) Cohort 30, $\alpha = 0.5$

Figure B.6: The proportion of samples above the stated lower income bound, plotted against the number of members, 30 years after joining. The x-axis is not to scale and, as in all the plots, points are joined with lines as a visual aid.