Income Stability of Pooled Annuity

Ge Qu

Under the supervision of Professor Thomas Berhardt[1]

[1]Department of Mathematics – University of Michigan(UM)

quge@umich.edu

Abstract. *The study focus on the income stability of pooled annuity fund members. We derive a mathematical expression to find the number of people who can receive lifelong stable income with a certain probability in a given pooled annuity fund. To solely study the idiosyncratic risk, we fixed the investment return and ignore the systematic risk. In the homogeneous case, each fund members are identical to each other, the trade-off between group size and income stability is studied. Then, we extend the method to the non-homogeneous case. In the non-homogeneous case, fund members are differ in initial account values. And the focus is on quantifying the effect of the composition of fund members with different initial account values on income stability.*

1. Background

There has been a transition from defined benefit pension plan(DB) to defined contribution pension plan(DC) in the US pension system under the increasing labor mobility and regulation changes. Under the DC plan, the risks are transferred to sponsors to employees since they are no longer guarantee a fixed, lifelong retirement income.

Defined benefit pension plan(DB) becomes less popular under the increasing labor mobility. The number of mobile workers in the US has been increasing in the past decades. DB plan is less attractive to mobile workers. On the one hand, the income calculation of DB is partially based on the length of service, which means long-tenured workers can benefit more than mobile workers. On the other hand, DB plans are not portable between different sponsors, which can cause accrual loses to mobile workers when they are changing employers.

Regulation changes triggered the transition from DB to DC. The U.S. Employee Retirement Income Security Act of 1974 (ERISA) issued a series of regulatory restrictions on firms to protect workers' benefits and reduce discrimination. The new regulation increases the administration cost and limits firms' flexibility of allocating benefits, which reduced their incentive to sponsor such traditional pension. The shift from DB to DC happened over decades after ERISA in the private sector. Through the transition, the portion of labors in the private industry covered by DC plan experienced a rapid increase and by 2012, exceeded 40%(Bureau of Labor Statistics, 2012, Table 2). Moreover, since the financial crisis in 2008, public DB fund liabilities start to exceed fund assets, which brought the pension reform into the public sector.

The shift from DB to DC transfers the risk from sponsors to households. In the DB plan, the payout to employees after retirement is predetermined with a formula including years of service and historical salary. Therefore, the sponsors bear the risk since they guarantee employees a fixed retirement income. However, for DC plan, the sponsors are no longer obligated to pay a fixed retirement income. The sponsors are only responsible to make regular deposits to the employees' retirement account.This addresses the risk to employees since they will decide how to invest the money on their retirement accounts to ensure financial security after retirement.

2. Introduction

For risk-averse investors who want to ensure a stable lifetime income especially after retirement, tontine is an attractive option. Tontine is an investment plan that provides a solution to longevity risk(the risk form uncertainty to future life time). In tontine plan, participants submit a lump-sum payment and receive a periodic payout afterward. And the dead participants' payout entitlements are distributed to the alive members, which allows members to pool the longevity risk.

Longevity risk is an important topic in the study of tontine, it affects the income stability of participants. Longevity risk management is widely discussed in the academic world.

Longevity risk pooling is discussed in Piggott et.al(2005)'s study. The article introduces group-self annuitization(GSA) plan where participants bear systematic longevity risk(the risk from choosing the wrong mortality model) but share idiosyncratic longevity risk(the risk that a finite number of observations cannot recapture the continuous ideal mortality distribution). The study suggests that the variation between mortality observations and mortality expectations results in payout fluctuations, which can be mitigated by sharing idiosyncratic risk among fund members.

Extending Piggott et.al's (2013) study on GSA, Qiao, and Sherris investigate the systematic longevity risk management. The study shows the threaten of systematic risk(risk of using the wrong mortality model and the uncertainty of future investment return) to the effectiveness of idiosyncratic risk pooling. Moreover, the study suggests that by adopting a mortality model that allows mortality rates to evolve randomly with expected changes, the negative effect of systematic risk can be mitigated.

Bernhardt and Donnelly (2020) further investigate on the idiosyncratic risk pooling. In their study, a mathematical expression that holds for any mortality model is derived to evaluate the income stability of a homogeneous pooled annuity group. The study generates numerical results on income stability with different group sizes. By quantifying the effect of group sizes on income stability, the study suggests how to effectively diversify idiosyncratic risk by adjusting group size.

In this study, we adopt the same assumptions and methods in Bernhardt and Donnelly's article but study the heterogeneous group. To isolate the idiosyncratic risk, we ignore the systematic risk and fix the investment returns. We derive a theoretical result on calculating the number of people receiving lifelong stable income in a pooled annuity group. Then, we conduct simulation on samples to obtain income streams of different groups. In our analysis on numerical results, we investigate strategies on arranging people with different account values in a group so that obtaining desired income stability.

3. Overview on Homogeneous-Case

Bernhardt and Donnelly (2020) studied on the income stability of homogeneous pooled annuity fund group. The study derives a way of quantifying income stability, that is, find the number of people who can receive life long stable income with a certain probability in a given homogeneous pooled annuity group. In the results, the study measures how changing 16 group size affects the income stability of a given homogeneous pooled annuity group. The study considers a pooled annuity fund where the following assumptions are made:

- Every member in the pooled annuity fund has the same age.
- Every member joins the pooled annuity fund at the same time.
- No member can leave the pool except through death.
- Each member have the same initial account value.
- The dead members' money on fund accounts will be distributed to the rest of the alive member in the next time period after the death time.

3.1. Variable Definitions

- N represents the initial number of members in the pooled annuity fund, where $N \geq 2$.

- x represents the age of each members when the pooled annuity fund start. That is, every fund members receive their first income payment at the age x+1, where $x \geq 0$.
- T_i, for $i \in \{1, 2...N\}$, is random variable representing ith members' future life time after age x. $T_1, T_2...T_N$ are independent and identically distributed random variables and are defined on probability space$(\Omega, \mathcal{F}, P)$.
- $T_{(i)}$, for $i \in \{1, 2...N\}$, is generated from rearranging T_i into an increasing order. That is, $T_{(1)} \leq T_{(2)} \leq ... \leq T_{(N)}$.
- L_{x+t} is the Number of alive members at the age of x+t which can be expressed as following:

$$L_{x+t} = \sum_{i=1}^{N} \mathbb{1}_{[T_i > t]}$$

where $\mathbb{1}$ is the indicator function of the set $T_i > t$

3.2. Survival Probability

3.2.1. Empirical Survival Probability

The empirical survival probability of a member to be alive at the age of x+t conditional on being alive at the age of x is expressed as:

$$_t\hat{p}_x = L_{x+t}/L_x, t \geq 0$$

3.2.2. Assumed True Survival Probability

The real survival probability that assumed to be true of a member to be alive at the age of x+t conditional on being alive at the age of x of is:

$$_tp_x = P[T_i > x + t | T_i > x]$$

3.3. Income fluctuation

In the homogeneous case, since the investment return is fixed and systematic longevity risk is ignored, we can solely investigate on the idiosyncratic longevity risk (the risk that comes from the difference between the observed empirical survival probability and assumed true survival probability). In the study, income of each fund member at time t can be expressed as:

$$C(t) = C(0)\frac{_tp_x}{_t\hat{p}_x}, \quad \text{if} \quad L_{x+t} > 0$$

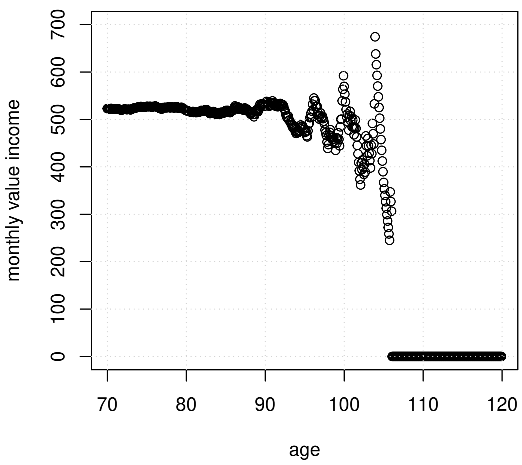

Figure 1. Income flow of an 1000 people group

Where C(t) is the income withdrawn at time t, $t \geq 0$.This expression reveals that the income fluctuations are only depend on the ratio of empirical survival probability and true survival probability. Recall that the systematic longevity risk comes from choosing the wrong future lifetime distribution which affects the value of ${}_tp_x$ here. The idiosyncratic longevity risk is the risk that future lifetime distribution is not observed perfectly. In the study, the investment return is fixed and the systematic longevity risk is ignored so that the idiosyncratic longevity risk can be studied in isolation. If the idiosyncratic longevity risk is perfect diversified, the income would stays the same as the initial income $(C(0) = C(t))$ for all $t \geq 0$. While the study focuses on non-perfect idiosyncratic longevity risk pooling, that is, the focus is on how close the future income stays to the initial income C(0).

3.4. Main Theorem of Bernhardt and Donnelly (2020)

The goal of pooled annuity funds is to ensure members a long-term stable income with a certain large probability. Therefore, the study set $(1 - \varepsilon)C(0)$ as the lower income threshold, $(1 + \varepsilon_2)C(0)$ as the upper income threshold, where $\varepsilon_1 \in (0, 1)$ and $\varepsilon_2 \ > 0$. Then, say the income that stays within the thresholds as stable. Thus, the study has the following expression:

$$P[\,(1 + \varepsilon_2)C(0) \geq C(s) \geq (1 - \varepsilon_1)C(0) \quad \text{for all} \quad s \in \{1, 2, ..., \lfloor T_{(k)} \rfloor\}] \geq \beta, \qquad (1)$$

where $\lfloor T_{(k)} \rfloor$ is the integer part of $T_{(k)}$, the time when k members have dead in the pool. β is the probability for the statement to hold. Note that s is in the unit of months, it represents the time points when payments are made in each month. The reason for using month as unit of time is to ensure the income in each month stays within the thresholds. Moreover, introducing $T_{(k)}$ into the formula is to study the number of people,k, who can receive lifelong stable income. β is the smallest probability that k out of N members in the group can receive lifelong stable income.

The expression (1) allows us to quantify the income stability by finding the maximum of k which is denoted as k_C with given ε_1, ε_2, β. That is, K_C is the number of members in the pool can enjoy a life long stable income. However, calculating K_C directly from expression (1) can be complicated since it requires to choose the distribution of T and the results vary for different mortality distribution. So instead of directly calculating K_C from (1), the study calculates the lower bound of K_C, using theorem 2.1, which largely simplifies the simulation process.

Theorem 3.1. Let $U_{(1)}, U_{(2)}, ..., U_{(N)}$ be the ordered statistics of N independent and standard uniformly distributed random variables $U_1, U_2, ..., U_N$. Fix constants $\varepsilon_1 \in (0, 1)$, $\varepsilon_2 > 0$ and $K \in \{1, 2, ...N\}$.

$$P[\,(1+\varepsilon_2)C(0) \geq C(s) \geq (1-\varepsilon_1)C(0) \quad \text{for all} \quad s \in \{1, 2, ..., \lfloor T_k \rfloor\}] \geq$$

$$P[\,(1-\varepsilon_1)\frac{i-1}{N} + \varepsilon_1 \geq U_{(i)} \geq (1+\varepsilon_2)\frac{\min\{i, N-1\}}{N} \text{for all} \quad i \in \{1, 2, ...N\}]\,.$$

Note the order statistics $U_{(1)}, U_{(2)}, ..., U_{(N)}$ is generated by ranking $U_1, U_2, ..., U_N$ from the lowest to highest. That is, $U_{(1)} \leq U_{(2)} \leq ... \leq U_{(N)}$. Same reasoning, $T_{(i)}$ is the order statistic of increasing order generated from T_i.

3.5. Approximation of K_C

Recall that due to the complexity of calculating K_C, the study approximate K_C to its lower bound, K_U. Based on Theorem 2.1, the K_U is the maximum of i. This approximation largely improves the efficiency of calculation: 1.it is independent from the mortality distribution T_i 2. It only involves random variable simulation and gets rid of stochastic process. Monte Carlo Simulation is used to obtain K_U, which is because that the probability of the expression including $U_{(i)}$ in Theorem 2.1 is challenging to compute. Moreover, the study also examines the accuracy of the approximation by calculating and comparing K_C and K_U. The result shows that the approximation is accurate and the accuracy increases with the value of ε.

3.6. Results and Conclusion

The study derive an efficient way of quantifying the income stability of a given homogeneous pooled annuity group. That is, calculating the number of people in a pooled annuity group who can receive life long income with certain probability. By conducting Monte Carlo simulation on homogeneous pooled annuity group of different size, the study find out the condition when the fund members can obtain desirable income stability. That is, when the group size is in thousands, the income stability can be ensured with a desirable large probability.

Figure 2 gives us an intuitive way of understanding the result. It shows one sample from the simulations of the income process for 1000 member groups. The blue lines are the income thresholds within which the income is defined to be stable. The plot shows that in the 1000 member group, the income stays stable until 725 members died. In the simulation results, there are 90% of samples maintain stable income until the death of the 725th member.

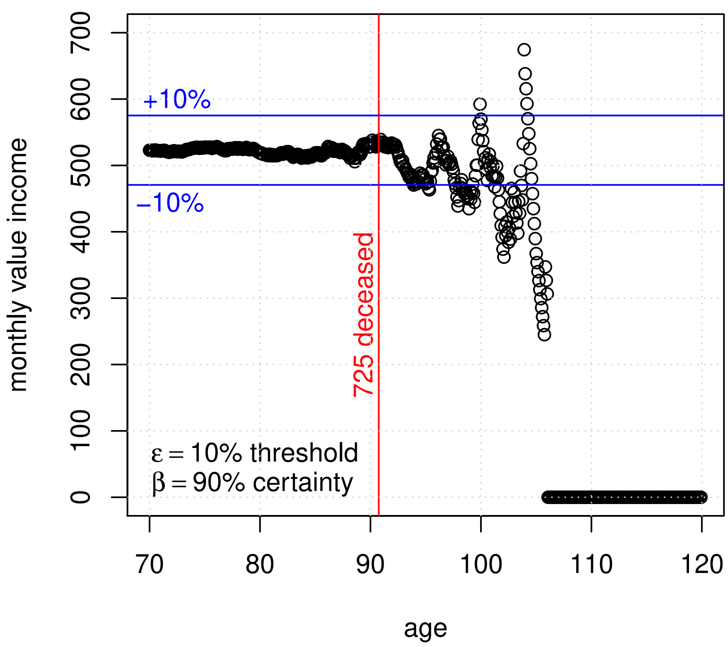

Figure 2. Income flow of 1000 member group

4. Non-homogeneous Case

We assume the participants of pooled annuity fund are composed of two groups of people: Group H, where members hold the same relatively high initial account value, and Group L, where members hold the same relatively low initial account value. Extending the method in the homogeneous case, we derive an expression to calculate the number of members in each group who can receive the life long stable income in a given non-homogeneous group. Therefore we can quantify how changing the ratio of the size of two groups, or the ratio of initial account values of two groups affect the income stability in the pooled annuity fund. We study the group where we make the following assumptions:

4.1. The Operation of the Pooled Annuity Fund

We study pooled annuity fund that follows the assumptions below:

- There are more than one participant in Group L and in Group H.
- Every member in the pooled annuity fund is of the same age.
- Every member join the pooled annuity fund at the same time at age x.
- No member can leave the pool except for death.
- Each member in Group L and in Group H have the same initial account value.
- The dead members' money on fund accounts will be distributed to the rest of the alive member in the next time period after the death time.

4.1.1. Future Life time Random Variables

Recall that there are N fund members initially, each fund member join the at age x, or equivalently at time 0. T_i, for $i \in \{1, 2...N\}$, is random variable representing ith members' future life time after age $x > 0$. $T_1, T_2...T_N$ are independent and identically distributed random variables an are defined on probability space$(\Omega, \mathcal{F}, P)$. $T_{(i)}$, for $i \in \{1, 2...N\}$, is order statistics of increasing order. That is, $T_{(1)} \leq T_{(2)} \leq ... \leq T_{(N)}$.

The number of alive members in Group L at age x+t is:

$$L_{x+t}^{l} = \sum_{i=1}^{N} \mathbb{1}_{[T_i > t| \text{ ith member belongs to Group L}]}$$

where $\mathbb{1}$ is the indicator function of the set $T_i > t$

The number of alive members in Group H at age x+t is:

$$L_{x+t}^{h} = \sum_{i=1}^{N} \mathbb{1}_{[T_i > t| \text{ ith member belongs to Group H}]}$$

where $\mathbb{1}$ is the indicator function of the set $T_i > t$

The number of alive members on the pooled annuity group at age x+t is:

$$L_{x+t} = L^l_{x+t} + L^h_{x+t}$$

4.1.2. Income Calculation

In the non-homogeneous pooled annuity group, each fund member in Group H hold the same account value, $W^h(t) \geq 0$, at time t. And each fund member in Group L hold the same account value, $W^l(t) \geq 0$, at time t.

The fund money is invested and yield investment returns with investment rate R, $R > -1$.

Each alive fund member withdraw income from their fund account periodically. And the value of income is the same as the payout of a fair life annuity if purchases with the current fund value. There for the income at time $t \in \{1, 2, ...\}$ is expressed as:

$$C^l(t) = W^l(t)/\ddot{a}_{x+t} \tag{2}$$

$$C^h(t) = W^h(t)/\ddot{a}_{x+t} \tag{3}$$

where

$$\ddot{a}_{x+t} = 1 + \sum_{j=1}^{\infty} (1+R)^{-j} jP_{x+t}$$

4.1.3. Adjusted Survival Probability

Let constant $a \geq 1$ be the ratio between initial account values of Group H and Group L:

$$W^h(0) = a * W^l(0)$$

And given (10), (11), it is easy to prove by induction that:

$$W^h(t) = a * W^l(t), t > 0 \tag{4}$$

Since the two subgroups, Group L and Group G, are two homogeneous groups, we follow the definition of empirical survival probability in the homogeneous case to define the empirical survival probability to survive at age x + t + s conditional on being alive at age x + t given the member belongs to Group L as:

$$ {}_s\hat{p}^l_{x+t} = L^l_{x+t+s}/L^l_{x+t} \tag{5}$$

given the member belongs to Group L as:

$$ {}_s\hat{p}^h_{x+t} = L^h_{x+t+s}/L^h_{x+t} \tag{6}$$

As the account value of fund members in Group H is always *a* times of the account value of members in group L and the fund group is homogeneous when ignoring the difference in group members' account value. We consider each member in Group H as exactly a copies of member in Group L. Therefore, we express the empirical survival probability to survive at age x + t + s conditional on being alive at age x + t, for $s, t > 0$, is:

$$ {}_s\hat{p}_{x+t} = (L^l_{x+t+s} + a * L^h_{x+t+s})/(L^l_{x+t} + a * L^h_{x+t}) \tag{7}$$

We can also write ${}_s\hat{p}_{x+t}$ as the weighted average of the empirical survival probability of the two homogeneous subgroups, Group H and Group L.

$$\begin{aligned} {}_s\hat{p}_{x+t} &= (L^l_{x+t+s} + a * L^h_{x+t+s})/(L^l_{x+t} + a * L^h_{x+t}) \\ &= \frac{L^l_{x+t}}{L^l_{x+t} + a * L^h_{x+t}} * \frac{L^l_{x+t+s}}{L^l_{x+t}} + \frac{a * L^h_{x+t}}{L^l_{x+t} + a * L^h_{x+t}} * \frac{L^h_{x+t+s}}{L^h_{x+t}} \\ &= \frac{L^l_{x+t}}{L^l_{x+t} + a * L^h_{x+t}} *_s \hat{p}^l_{x+t} + \frac{a * L^h_{x+t}}{L^l_{x+t} + a * L^h_{x+t}} *_s \hat{p}^h_{x+t} \end{aligned} \tag{8}$$

The real survival probability that assumed to be true of a member to be alive at the age of x+t+s conditional on being alive at the age of x+t of is:

$$ {}_sp_{x+t} = P[T_i > x + t + s | T_i > x + t]$$

4.1.4. The Longevity Credit Calculation

In the pooled annuity fund, the longevity credit is the money that alive members received from the dead members. Specifically, the account value of member who dead during time $(t, t+1]$ will be distributed to the account of survival fund member at time t+1. Recall that we consider the account of Group H member as *a* copies of Group L member's account. Therefore, we express the longevity credit to Group L member from time t to time t+1 as:

$$ M^l(t+1) = \frac{(W^l(t) - C^l(t))(1+R)(L^l_{x+t} - L^l_{x+t+1} + m * L^h_{x+t} - m * L^h_{x+t+1})}{L^l_{x+t+1} + m * L^h_{x+t+1}}. \tag{9}$$

Where $L^l_{x+t+1} + m * L^h_{x+t+1} > 0$. In the equation (7), $(W^l(t) - C^l(t))(1 + R)$ represents the account value of dead member in Group L at time t+1, $(L^l_{x+t} - L^l_{x+t+1} +$

$m * L^h_{x+t} - m * L^h_{x+t+1}$) represents the number of copies of Group L member converted from all dead members. Then the sum of dead members' account value at time t+1 is evenly distributed to each copy of Group L account. If $L^l_{x+t+1} + m * L^h_{x+t+1} = 0$, the pooled annuity fund terminates.

Since we find that the longevity credit to Group H members is *a* times of $M^l(t+1)$:

$$M^h(t+1) = a * M^l(t+1) \tag{10}$$

After the payment of longevity credit at time t+1, the account value of fund members in Group L at time t+1 becomes:

$$W^l(t+1) = (W^l(t) - C^l(t))(1+R) + M^l(t+1), T > 0 \tag{11}$$

4.2. Income Fluctuation

By manipulating the formulas we derived, we can find the factor that cause the income fluctuations. First, substitute $w^l(t)$ and $w^l(t+1)$ in (9) using (2). Then, use the property $\ddot{a}_{x+t} - 1 = p_{x+t}\ddot{a}_{x+t+1}/(1+R)$. Finally substitute $M^l(t+1)$ in (9) using (7). We then obtain:

$$C^l(t+1) = C^l(t)p_{x+t}/\hat{p}_{x+t} \tag{12}$$

Same reasoning,

$$C^h(t+1) = C^h(t)p_{x+t}/\hat{p}_{x+t} \tag{13}$$

This gives us the same conclusion on what causes the income fluctuation as in the homogeneous case. That is, under our assumptions of fixed investment return and zero systematic longevity risk, the income fluctuations only come from the difference between the survival probability observed and the survival probability assumed to be true which is the source of idiosyncratic longevity risk. If the idiosyncratic longevity risk is perfectly pooled, the empirical survival probability is observed equal to the true survival probability ($p_{x+t}/\hat{p}_{x+t} = 1$), the income fluctuation will become zero.

plugging equation (4) into equation (3), we obtain that:

$$C^h(t) = a * C^l(t), t \geq 0 \tag{14}$$

4.3. The Main Theorem

We use the same method to study income stability as in the homogeneous case. Fix $\varepsilon_1 \in (0,1)$ as the lower threshold parameter. Fix $\varepsilon_2 > 0$ as the upper threshold parameter.

Fix β as the smallest probability for a scenario of future income stream to happen. Income $C^l(t)$ at time $t > 0$ is considered as stable if it stays close enough to the initial income $C^l(0)$. Therefore, we define $C^l(t)$ to be stable if $C^l(t) \in [C^l(0)(1-\varepsilon_1), C^l(0)(1+\varepsilon_2)]$, where $C^l(0)(1-\varepsilon_1)$ is the lower income threshold and $C^l(0)(1+\varepsilon_2)$ is the upper income threshold. Same reasoning, $C^h(t) \in [C^h(0)(1-\varepsilon_1), C^h(0)(1+\varepsilon_2)]$ is defined as stable income. Using (10), we can prove that if $C^l(t)$ stays stable, then $C^h(t)$ also stays stable and vise versa. Therefore, we can simply observe the time period where $C^l(t)$ stays stable to determine the time length of stable income for the whole group.

To find the number of members named K_C who can receive stable income for life with at least probability β, we have the following expression:

$$P[\,(1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}] \geq \beta$$

where$\lfloor T_k \rfloor$ represents the integer part of $T_{(k)}$, the time point when k members have dead in the pooled annuity group. The largest k that satisfy the expression above is denoted as K_C. Recall that due to the complexity of calculating K_C, we calculate its lower bound K_U based on the theory below.

Theorem 4.1. *Let $U_{(1)}, U_{(2)}, ..., U_{(N)}$ be the ordered statistics of N independent and standard uniformly distributed random variables $U_1, U_2, ..., U_N$. Fix constants $\varepsilon_1 \in (0,1)$, $\varepsilon_2 > 0$ and $K \in \{1,2,...N\}$.*

$$P[\,(1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}]$$
$$\geq P\left[\left[U_{(i)} \leq \varepsilon_1 + \frac{(1-\varepsilon_1)(L^l(T_{(i-1)}) + a * L^h(T_{(i-1)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1,2...k\}\right]\right.$$
$$\left.\cap \left[U_{(i)} \geq -\varepsilon_2 + \frac{(1+\varepsilon_2)(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1,2...k\} \setminus \{N\}\right]\right]$$

Proof. Fix $t \in N$ and assume that at least one person is alive at age x + t + 1. With equation (10), it can be proved by induction that:

$$C^l(t+1) = C^l(0) \prod_{j=0}^{t} \frac{p_{x+j}}{\hat{p}_{x+j}} = C^l(0) \frac{{}_{t+1}p_x}{{}_{t+1}\hat{p}_x}$$

Noting ${}_s\hat{p}_x > 0 \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}$, and as the joint distribution of $T_1, T_2, ..., T_n$ is continuous, the set $[\lfloor T_k \rfloor \leq T_{(k)}, \text{for } k = 1,2,...,N]$ has measure one.

$$[\,(1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}]$$

$$= [\, 1+\varepsilon_2 \geq \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1 \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}]$$

$$= \left[\inf_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1 \right] \cap \left[\sup_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2 \right]$$

Since $\{1,2,..,\lfloor T_k \rfloor\} \subset [\, 0, T_k) \subset [\, 0, T_{k+1})$

$$\inf_{s\in[\,0,T_k)} \frac{{}_s\hat{p}_x}{{}_s p_x} \quad \leq \quad \inf_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s\hat{p}_x}{{}_s p_x}$$

$$\sup_{s\in[\,0,T_{k+1})} \frac{{}_s p_x}{{}_s\hat{p}_x} \quad \geq \quad \sup_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x}$$

And,

$$\inf_{s\in[\,0,T_k)} \frac{{}_s\hat{p}_x}{{}_s p_x} \geq 1-\varepsilon_1 \Rightarrow \inf_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1$$

$$\sup_{s\in[\,0,T_{k+1})} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2 \Rightarrow \sup_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2$$

Therefore,

$$\left[\inf_{s\in[\,0,T_k)} \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1 \right] \cap \left[\sup_{s\in[\,0,T_{k+1})} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2 \right]$$

$$\subset \left[\inf_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1 \right] \cap \left[\sup_{s\in\{1,2,...,\lfloor T_k \rfloor\}} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2 \right]$$

In summary,

$$\left[\inf_{s\in[\,0,T_k)} \frac{{}_s p_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1 \right] \cap \left[\sup_{s\in[\,0,T_{k+1})} \frac{{}_s p_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2 \right]$$

$$\subset [\, (1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1,2,...,\lfloor T_k \rfloor\}] \tag{15}$$

The Empirical distribution function of death times of Group L members $T_1^l, T_2^l, ..., T_{N^l}^l$ is defined at $t > 0$:

$$\hat{F}_{N^l}(t) := \frac{1}{N^l}\sum_{i=1}^{N^l} \mathbb{1}_{\{T_i^l \leq t\}}$$

The Empirical distribution function of death times of Group H members $T_1^h, T_2^h, ..., T_{N^h}^h$ is defined at $t > 0$:

$$\hat{F}_{N^h}(t) := \frac{1}{N^h}\sum_{i=1}^{N^h} \mathbb{1}_{\{T_j^h \leq t\}}$$

The distribution function of all fund members' death time $T_1, T_2, ..., T_N$ is defined at $t > 0$ as $F(t)$. And ${}_sp_x = 1 - F(s)$.

Then, from equation (5) and (6):

$${}_s\hat{p}_x^l = 1 - \hat{F}_{N^l}(s)$$

$${}_s\hat{p}_x^h = 1 - \hat{F}_{N^h}(s)$$

Therefore,

$$\left[\inf_{s\in[0,T_k)} \frac{{}_sp_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1\right] \cap \left[\sup_{s\in[0,T_{(k+1)})} \frac{{}_sp_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2\right]$$

$$= \left[\inf_{s\in[0,T_k)} \frac{1-F(s)}{(\frac{N^l}{N^l+aN^h}) * (1-\hat{F}_{N^l}(s)) + (\frac{aN^h}{N^l+aN^h}) * (1-\hat{F}_{N^h}(s))} \geq 1-\varepsilon_1\right]$$

$$\cap \left[\sup_{s\in[0,T_{(k+1)})} \frac{1-F(s)}{(\frac{N^l}{N^l+aN^h}) * (1-\hat{F}_{N^l}(s)) + (\frac{aN^h}{N^l+aN^h}) * (1-\hat{F}_{N^h}(s))} \leq 1+\varepsilon_2\right]$$

To simplify the writing, we let:

$$\mu = \frac{N^l}{N^l + aN^h}$$

$$\nu = \frac{aN^h}{N^l + aN^h}$$

Therefore,

$$\left[\inf_{s\in[0,T_k)} \frac{{}_sp_x}{{}_s\hat{p}_x} \geq 1-\varepsilon_1\right] \cap \left[\sup_{s\in[0,T_{(k+1)})} \frac{{}_sp_x}{{}_s\hat{p}_x} \leq 1+\varepsilon_2\right]$$

$$= \left[\inf_{s\in[0,T_k)} \frac{1-F(s)}{\mu * (1-\hat{F}_{N^l}(s)) + \nu * (1-\hat{F}_{N^h}(s))} \geq 1-\varepsilon_1 \right] \tag{16}$$

$$\cap \left[\sup_{s\in[0,T_{(k+1)})} \frac{1-F(s)}{\mu * (1-\hat{F}_{N^l}(s)) + \nu * (1-\hat{F}_{N^h}(s))} \leq 1+\varepsilon_2 \right]$$

Where $\frac{1-F(s)}{\mu*(1-\hat{F}_{N^l}(s))+\nu*(1-\hat{F}_{N^h}(s))} := 1$

When $\mu * (1-\hat{F}_{N^l}(s)) + \nu * (1-\hat{F}_{N^h}(s)) = 0$. Let $T_{(0)} := 0$. As the joint distribution of $T_1, T_3, ..., T_N$ is continuous, the set $[T_{(i-1)} < T_{(i)}]$, for $K = 1, 2, ..., N+1]$ has measure one. In the following, we work on this set only.

Consider a random time interval $[\,T_{(i-1)}, T_{(i)})$. According to the definition, $\hat{F}^l_N(S)$ $(\hat{F}^h_N(S))$is an increasing function which increases by one at each Group L (Group H) member's death time.Therefore the value of $\hat{F}^l_N(S)$ $(\hat{F}^h_N(S))$ in $[\,T_{(i-1)}, T_{(i)})$ is a fixed number $\hat{F}^l_N(T_{(i-)}) = \hat{F}^l_N(T_{(i-1)})$ $(\hat{F}^l_N(T_{(i-)}) = \hat{F}^l_N(T_{(i-1)}))$.

And as $F(S)$ is an increasing function, $1-F(S)$ is an decreasing and takes minimum in $[\,T_{(i-1)}, T_{(i)})$ at $T_{(i-)}$. By continuity of F, $T_{(i-)} = T_{(i)}$.

Therefore,

$$\inf_{s\in[T_{(i-1)},T_{(i)})} \frac{1-F(s)}{\mu * (1-\hat{F}_{N^l}(s)) + \nu * (1-\hat{F}_{N^h}(s))}$$

$$= \frac{1-F(T_{(i-)})}{\mu * (1-\hat{F}_{N^l}(T_{(i-)})) + \nu * (1-\hat{F}_{N^h}(T_{(i-)}))}$$

$$= \frac{1-F(T_{(i)})}{\mu * (1-\hat{F}_{N^l}(T_{(i-1)})) + \nu * (1-\hat{F}_{N^h}(T_{(i-1)}))}$$

Same reasoning, F(S) takes maximum in $[\,T_{(i-1)}, T_{(i)})$ at $T_{(i-1)}$.

Therefore,

$$\sup_{s\in[T_{(i-1)},T_{(i)})} \frac{1-F(s)}{\mu * (1-\hat{F}_{N^l}(s)) + \nu * (1-\hat{F}_{N^h}(s))}$$

$$= \frac{1-F(T_{(i-1)})}{\mu * (1-\hat{F}_{N^l}(T_{(i-1)})) + \nu * (1-\hat{F}_{N^h}(T_{(i-1)}))}$$

Therefore,

$$\left[\inf_{s\in[0,T_k)} \frac{1-F(s)}{\mu*(1-\hat{F}_{N^l}(s))+\nu*(1-\hat{F}_{N^h}(s))} \geq 1-\varepsilon_1\right]$$

$$= \left[\inf_{i\in\{1,2...k\}} \frac{1-F(T_{(i)})}{\mu*(1-\hat{F}_{N^l}(T_{(i-1)}))+\nu*(1-\hat{F}_{N^h}(T_{(i-1)}))} \geq 1-\varepsilon_1\right] \tag{17}$$

$$\left[\sup_{s\in[0,T_{k+1})} \frac{1-F(s)}{\mu*(1-\hat{F}_{N^l}(s))+\nu*(1-\hat{F}_{N^h}(s))} \leq 1+\varepsilon_2\right]$$

$$= \left[\sup_{i\in\{1,2...k+1\}} \frac{1-F(T_{(i-1)})}{\mu*(1-\hat{F}_{N^l}(T_{(i-1)}))+\nu*(1-\hat{F}_{N^h}(T_{(i-1)}))} \leq 1+\varepsilon_2\right]$$

$$= \left[\sup_{i\in\{1,2...k+1\}} \frac{1-F(T_{(i-1)})}{\mu*(1-\hat{F}_{N^l}(T_{(i-1)}))+\nu*(1-\hat{F}_{N^h}(T_{(i-1)}))} \leq 1+\varepsilon_2\right] \tag{18}$$

As F is continuous, the random variables $U_i = F(T_i)$ are independent and standard uniformly distributed. Therefore, we have:

$$U_{(i)} = F(T_{(i)}) \text{ for } i = 1, 2, ..., N$$

Then, according to the definition of $\hat{F}_{N^l}$ and $\hat{F}_{N^h}$, we have:

$$\left[\inf_{i\in\{1,2...k\}} \frac{1-F(T_{(i)})}{\mu*(1-\hat{F}_{N^l}(T_{(i-1)}))+\nu*(1-\hat{F}_{N^h}(T_{(i-1)}))} \geq 1-\varepsilon_1\right]$$

$$\cap\left[\sup_{i\in\{1,2...k\}} \frac{1-F(T_{(i)})}{\mu*(1-\hat{F}_{N^l}(T_{(i)}))+\nu*(1-\hat{F}_{N^h}(T_{(i)}))} \leq 1+\varepsilon_2\right]$$

$$= \left[\inf_{i\in\{1,2...k\}} \frac{1-U(i)}{\frac{N^l}{N^l+a*N^h}*(1-\frac{L^l(T_{(i-1)})}{N^l})+\frac{a*N^h}{N^l+a*N^h}*(1-\frac{L^h(T_{(i-1)})}{N^h})} \geq 1-\varepsilon_1\right]$$

$$\cap\left[\sup_{i\in\{1,2...k\}} \frac{1-U(i)}{\frac{N^l}{N^l+a*N^h}*(1-\frac{L^l(T_{(i)})}{N^l})+\frac{a*N^h}{N^l+a*N^h}*(1-\frac{L^h(T_{(i)})}{N^h})} \leq 1+\varepsilon_2\right]$$

$$= \left[U_{(i)} \leq \varepsilon_1 + \frac{(1-\varepsilon_1)(L^l(T_{(i-1)}) + a * L^h(T_{(i-1)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1, 2...k\}\right]$$

$$\cap \left[U_{(i)} \geq -\varepsilon_2 + \frac{(1+\varepsilon_2)(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1, 2...k\} \setminus \{N\}\right]$$

Combine (15)-(18) and take the probability, we obtain the desired result. □

Corollary 4.2. *Suppose that for $K \in N$ and $\varepsilon \in (0,1)$,*

$$P\left[\left[U_{(i)} \leq \varepsilon + \frac{(1-\varepsilon)(L^l(T_{(i-1)}) + a * L^h(T_{(i-1)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1, 2...k\}\right]\right.$$
$$\left.\cap \left[U_{(i)} \geq -\varepsilon + \frac{(1+\varepsilon)(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1, 2...k\} \setminus \{N\}\right]\right] \geq \beta \tag{19}$$

Then,

$$P[\,(1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1, 2, ..., \lfloor T_k \rfloor\}] \geq \beta \tag{20}$$

Proof. Apply theorem 3.1 with $\varepsilon_1 = \varepsilon_2 = \varepsilon$ □

Corollary 4.3. *Suppose that for $K \in N$ and $\varepsilon \in (0,1)$,*

$$P\left[\left[U_{(i)} \leq \varepsilon + \frac{(1-\varepsilon)(L^l(T_{(i-1)}) + a * L^h(T_{(i-1)}))}{N^l + a * N^h} \quad \text{for all} \quad i \in \{1, 2...k\}\right]\right] \geq \beta \tag{21}$$

Then,

$$P[\,(1+\varepsilon_2)C^l(0) \geq C^l(s) \geq (1-\varepsilon_1)C^l(0) \quad \text{for all} \quad s \in \{1, 2, ..., \lfloor T_k \rfloor\}] \geq \beta \tag{22}$$

Proof. Apply theorem 3.1 with $\varepsilon_1 = \varepsilon, \varepsilon_2 \uparrow \infty$,

$$-\varepsilon_2 + \frac{(1+\varepsilon_2)(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h}$$
$$= \frac{(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h} + \varepsilon_2\left(\frac{(L^l(T_{(i)}) + a * L^h(T_{(i)}))}{N^l + a * N^h} - 1\right) \to -\infty$$
$$\text{for all } i \in \{1, 2, ..., N-1\}$$

□

4.4. Calculation of K_U

Recall that our goal is to find the number of people who can receive stable life-long income with at least probability β. But due to the complexity of calculating K_C directly from (20), we calculate its lower bound K_U from (19). To calculate K_U, we use Monte Carlo simulation.

Let N as the initial number of people in the pooled annuity fund group. Then we let N^l be the initial number of people in Group L, and N^h be the initial number of people in Group H. Thus we have $N = N^h + N^l$. Then we fix β, and ε. Let $\varepsilon_1 = \varepsilon_2 = \varepsilon$ thus income thresholds are systematic about initial incomes. Let positive integer M be the number of repentance of the simulation process. In our study, we fix M as 100000.

Moreover, it is important for our study to know, to which group a person belongs, when the person dies. Considering an empty vector of length N, we sample N^l positions from the vector and mark them as deaths of Group L members, and the rest of the positions are marked as deaths of N^h members.

We generate M sample vectors of the uniform ordered statistics $(U_{(i)})_{i=1}^{N}$. We call mth sample vector as $(U_{(i)}^{(m)})_{i=1}^{N}$ for $m \in \{1, 2, ..., M\}$. Then we conduct simulation and find the first integer $i(m) \in \{1, 2, ..., N\}$ that fails equation (19). Then record $K(m) = i(m) - 1$. If i(m) does not exist, record k(m) = N.

After M times of simulation, we obtain a vector of K(m) which we called $K(m)_{i=1}^{M}$. Then we take the $\beta - th$ quantile of $K(m)_{i=1}^{M}$ as K_U.

We use a shortcut which avoid sorting when generating ordered standard uniform distribution $(U_{(i)}^{(m)})_{i=1}^{N}$. This method is according to Devroye's study, where Devroye concludes that uniform ordered statistics are ratios of sums of exponential random variables.

4.5. Numerical Results

Pooled annuity group of 1000 members is studied and the systematic income threshold ε is fixed as 0.1, desired probability β is fixed as 0.1. To simplify the simulation process, a lower bound is considered. And each K_U is calculated from 100000 times of simulation. The goal is to see how does K_U responds to the changes in account value ratio and the size of Group L. We choose five account ratio to observe: $W^l(t)/W^h(t) = 0.1, 0.2, 0.3, 0.5, 0.7$. With each fixed account ratio, we choose N^l to be 10,20,30,...,1000. Then, we compute K_U for each possible combination of parameters and generate a data set of 500 K_U values. To visualize the result, we generate figure 4. The solid lines shows the relationship between K_U and N^lThe dashed lines are reference lines that generated from homogeneous groups. The long dashed line shows the number of members in Group H who can receive lifetime stable income when isolating Group H as a homogeneous group. The dotted dashed line shows the number of members in Group L who can receive lifetime stable income when isolating Group L as a homogeneous group. The short dash line shows the number of members who can receive lifetime stable income in a 1000 member homogeneous group. By observing figure 4, we have several findings:

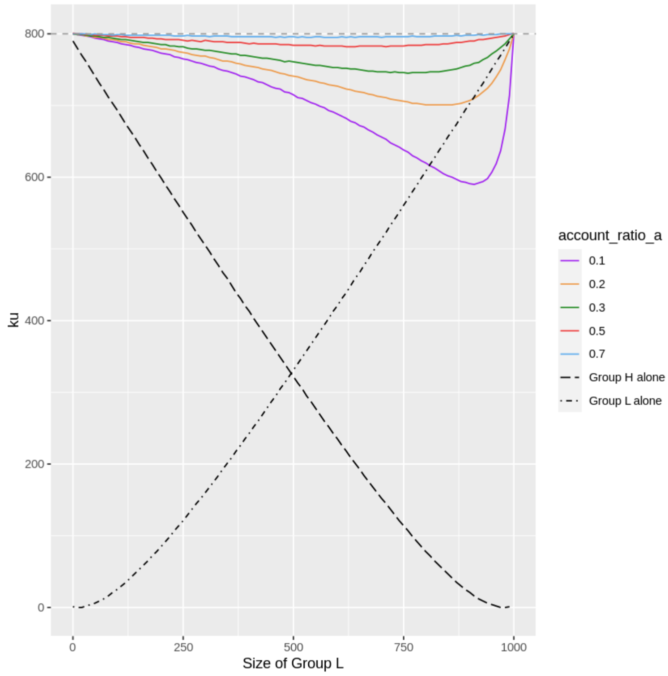

Figure 3. Simulation Result

- The number of members who can receive lifelong stable income (K_U) depends on the initial sub-group sizes and account value ratio. In Figure 5, at the same size of Group L, the group with a higher account value ratio has more people receiving lifelong stable income. When the initial account value is fixed, K_U changes with the size of Group L.
- When the size of Group L is fixed, non-homogeneous group is more likely to obtain better income stability. For example when the size of Group L is 500, the smallest K_U of non-homogeneous group in the graph is 715. However, when separating Group L and Group G as two homogeneous groups, the total number of people who can receive lifetime stable income in Group L and Group G is 662.
- When the account ratio becomes smaller enough and the size of Group L becomes large enough, the advantage of the non-homogeneous group in income stability

starts to vanish and Group L starts loosing benefit. There are parts of solid curves staying below the dotted dashed curve, where the K_U of Group L in mixed group is less than that of isolated Group L.

- The income stability of an 1000 member homogeneous group seems to be always better than the income stability of an 1000 member non-homogeneous group. In the graph, the horizontal curve at $K_U = 800$ is always above the solid lines, which shows K_U of the 1000 member homogeneous group is always larger than that of 1000 member non-homogeneous groups.

4.6. Conclusions

The income stability of a non-homogeneous pooled annuity group is studied. A mathematical expression used to calculate the number of members who can receive lifelong stable income is derived. Through Monte Carlo simulation, numerical results are generated for groups with different sub-group size and account ratio.

The numerical result suggests that given the group size, whether the non-homogeneous group can obtain better income stability than separated homogeneous groups depends on the choice of group size and account ratio. By choosing an appropriate combination of account ratio, and sub-group size, the mixed non-homogeneous group can have more people receiving a lifelong stable income than the separated homogeneous group. However, when the account ratio becomes sufficiently small and the portion of members with lower account value becomes sufficiently large, the members with lower account value can no longer have higher income stability in the non-homogeneous group.

In the future, the study will focus on quantifying the trade-off between account ratio, sub-group size and income stability. Specifically, knowing homogeneous groups and their income stability, the study will derive a model on predicting the income stability of the mixed non-homogeneous group. In application, the study will provides methods on finding the way of maximizing income stability when designing the group formation.

References

[1] Bernhardt,T., Donnelly. (2020). *Quantifying the trade-off between income stability and the number of members in pooled annuity funds* ASTIN Bulletin(forthcoming).

[2] Bureau of Labor Statistics 2012. (2012). *National compensation survey: employee benefits in the United States.*

[3] Butrica, B. A., Iams, H. M., Smith, K. E., Toder, E. J. (2009). *The disappearing defined benefit pension and its potential impact on the retirement incomes of baby boomers.* Soc. Sec. Bull, 69, 1.

[4] Devroye, L. (2009). *Non-Uniform Random Variate Generation.* Springer-Verlag.

[5] Piggott, J., Valdez, E. A., Detzel, B. (2005). *The simple analytics of a pooled annuity fund.* Journal of Risk and Insurance, 72(3), 497-520.

[6] Qiao, C., Sherris, M. (2013). *Managing Systematic Mortality Risk With Group Self-Pooling and Annuitization Schemes.* Journal of Risk and Insurance, 80(4), 949-974.

[7] The US government. (1974). *The U.S. EmployeeRetirement Income Security Act.*

[8] Webber, D. H. (2019). *Reforming Pensions While Retaining Shareholder Voice.* BUL Rev., 99, 1001.