UNIVERSITA' DEGLI STUDI DI TORINO

Dipartimento di Scienze Economico-Sociali e Matematico-Statistiche

Corso di Laurea in Quantitative Finance and Insurance

TESI DI LAUREA MAGISTRALE

Pension Decumulation Strategies: Analysis of Modern Tontines and Simulation of a GSA plan

Relatore: Prof.ssa Elena Vigna
Correlatore: Prof. Bertrand Lods

Candidato: Silvia Salemme

Anno Accademico 2020/2021

Contents

Introduction

Many studies have shown that pay-as-you-go pension systems are no longer sustainable: most European countries are facing major pension challenges and are no more able to guarantee stable incomes for retirees. The demographic trend is very dramatic: on the one hand, mortality rates are decreasing, leading to an increasing elderly population; on the other hand, birth rates are falling, leading to a decreasing young population. Taking demographic development into account, it is therefore evident that people are increasingly been called on to take responsibility for their future retirement income. People need to supplement the first public pension pillar with additional plans in order to be sure that enough resources will be available in old age.
Moreover, in most public pension schemes, retirees are forced to convert their accumulated funds into life annuities. The issue, however, is that the annuities available on the insurance market are not sufficiently attractive for investors: premiums charged are too high, and retirees prefer to benefit from a greater degree of flexibility and an higher level of freedom.

In this work, I decide to examine some of the actions performed by the United Kingdom as an exemplary way to overcome the previously presented issues. The UK has, in fact, introduced two important reforms in recent years that really try to accommodate consumer needs:

- The Automatic Enrollment in 2012 [25], which requires all employers to enroll their eligible workers into a workplace pension plan. This choice was designed to increase the amount of retirement savings for the entire population.

- The Pension Freedom in 2015 [22], which allows people who have reached the age of 55 to decide how, when and whether to access their pension pot. The UK government tried to provide pensioners with greater flexibility and to give them different options at their disposal. It enabled people not to be forced to buy a typical annuity product with their accumulated fund, but to opt for other alternatives. Under the careful guidance of qualified experts, retirees can therefore decide which decumulation strategy is most suitable for them. Individuals can decide to leave the pension pot untouched, buy an annuity, cash in the whole pocket or drawdown their accumulated fund.

Hence, in this study I analyze in great detail the decumulation strategies proposed in the literature. In particular, I decide to examine extensively a peculiar product, seen as a valuable alternative to life annuities, called tontine. Tontines were special forms of investment, dating back to the 17th century, in which the investor paid a lump sum of money and received annual payments of 'dividends' until his death. The special feature of this product was that when an investor died, his shares were divided among the surviving members of the tontine. The tontine was thus seen as a group annuity in which the investor living longer would get larger annual payments. Today, a growing number of financial advisors, academics, and Fintech firms think that it might be time to take a second look at these financial arrangements. Many authors have examined whether an historical insurance concept such as the tontine has sufficient innovative potential to extend and improve on prevailing private pension solutions in a modern way. A recent stream of literature proposes and analyzes modern versions of tontines. The main core of this thesis is therefore devoted to the analysis and explanation of modern tontines proposed in the literature. In order to better understand the actual importance, innovation and attractiveness of these products, in most of the analyses presented I introduce comparisons with annuities. Tontines, unlike annuities, have no guarantee of income, exhibit more volatile payments, but offer higher expected returns and lower costs. Depending on the risk aversion preferences of individuals, tontines can thus be seen as a better and more cost-effective way of restoring pensions. The message I would like to leave out is that tontines can certainly be interesting and valuable ways for people to fund their later years and thus solve some of the pension challenges that many countries are facing.

The thesis is organized as follows: Chapter 1 is mainly devoted to the analysis and description of decumulation strategies proposed in the literature; it is made up of 4 sections. The first one is dedicated to the introduction of pension schemes in general; the second one is devoted to the analysis of a survey conducted by Insurance Europe [15], regarding the importance of supplementary pension plans; the third one is reserved to the presentation of reforms introduced in the UK; finally in the fourth section decumulation strategies are explained. I will cover the analysis of many approaches, including: Utility theory-based methods, such as Expected Utility Theory and Cumulative Prospect Theory; techniques based on the minimization of the difference between current consumption and a desired consumption; probabilistic methods, such the minimization of the probability of ruin; habit formation approaches. Chapter 2 focuses completely on the explanation of modern tontines; each section is dedicated to a specific tontine proposed in the literature. Consistently, I submit the following proposals: the Group Self-Annuitization plan of Piggott et. al(2016) [26], the Optimal Retirement Tontine of Milevsky and Salisbury (2016) [20], the Fair Tontine Annuity of Sabin (2010) [27], the Annuity Overlay Fund of Donnelly et. al (2014) [9] and the Pooled Annuity Fund of Stamos (2008) [30]. Chapter 3 is entirely devoted to the simulation of a Group Self-Annuitization plan. Finally, Chapter 4 concludes.

Chapter 1

Pension Decumulation Strategies

1.1 Introduction to pensions schemes

" Pension scheme means a contract, an agreement, a trust deed or rules stipulating which retirement benefits are granted and under which conditions"

— Article 6 of Directive 2016/2341/EU - IORP II Directive

A pension scheme is organised in two phases: an accumulation phase in which people set aside sums of money during their working life, and a decumulation phase in which persons receive payments after retirement.
Consequently, the most important components to account for in a pension scheme are the contributions made during the working life of a person and the benefit that person will receive during retirement. We can therefore distinguish two different types of monetary flows: before retirement people set money aside, in order to receive payments in later life.
The main objectives of pension schemes are to protect against the risk of poverty in old ages and to transfer resources from work to retirement, in order to provide smooth consumption during the entire lifespan of people.

Considering the relationship between contributors and retirees there are two kinds of pension systems:

- Pay as you go (PAYG): PAYG systems imply that contributions paid by workers are directly transferred to retirees to pay their pension benefit. Contributions are not accrued or invested in the market, there is not a personal bank account, thereby there is not a direct link between contribution paid and benefit received. Due to the key relationship between the number of workers and the number of pensioners, PAYG systems must ensure sustainability and long-term transaction periods that are determined by long longevity and low fertility trends. It is indeed important to take into account demographic changes to ensure the sustainability of the system.

- Funded : in founded systems the contribution that each person makes during his working life is transferred to a personal account, invested and used to finance their own future pension. The performance of the pension will depend on the market, and the returns of the underlying assets are uncertain and variable. Such a system is not affected by changing demographics, but it bears a much higher level of financial risk.

Depending on whether benefits or contributions are fixed, we can distinguish:

- Defined benefit pensions (DB): The pension provider promises a defined benefit at retirement to the pensioner, based on a predefined formula which depends on the earning history of the individual, his age ,working life and possibly gender. Independently on the investment returns the sponsor must honour his obligations and provide the fixed return agreed in advance. In this type of pension system the risks are borne by the pension provider: the retiree does not hold any type of risk, in fact he is guaranteed a fixed amount per month forever.[1]
- Defined contribution pensions (DC): In DC plans contributions are made on a regular basis and are paid by employers and employees as a fixed percentage of the salary. The contributions paid during the working life are set up into accounts and benefits received are proportional to the amount contributed plus investment earnings. Benefits of DC plans are not stable and fluctuate depending on investment returns: the pension provider has not obligation to provide fixed returns. There are no guarantees in terms of final resources, and the retiree bears all pension risks.

Over the past few decades there has been a gradual shift from defined benefit to defined contribution plans. The shift towards DC pension plans has been a response to the under-funding of pensions, declining long-term interest rates, increasing regulatory burden and uncertainty, recognition of the effects of increased longevity on plan costs and increasing desire of the employer to get rid of financial risk. There has thus been a progressive shift of risks from employers to employees, combined with a significant increase in the accountability of individuals.

[1] In defined benefit pension schemes the employee pays a fixed percentage of his or her salary as a pension contribution. By the other side the employer ensures a fixed benefit at maturity: in order to guarantee such a coverage, considering that markets are uncertain and fluctuating, he has to pay an additional variable contribution. Actuaries in fact calculate each year the adjusted contribution rate that the employer has to settle in order to be able to guarantee the promised return at retirement. There also exists the so-called "contribution holiday": when the market performance is so good that the employer does not have to contribute to the employee's pension at all, in essence when the adjusted contribution rate calculated by actuaries is zero.

Finally, pension provisions are usually classified into pillars.
Pillar I corresponds to the Public Pension provided by the government, usually under a PAYG approach; Pillar II corresponds to occupational pensions, where employers provide a pension designed as DB or DC plan, usually under a funded in advance approach; and Pillar III refers to personal pensions and covers individual saving plans typically set up on voluntary basis by households.

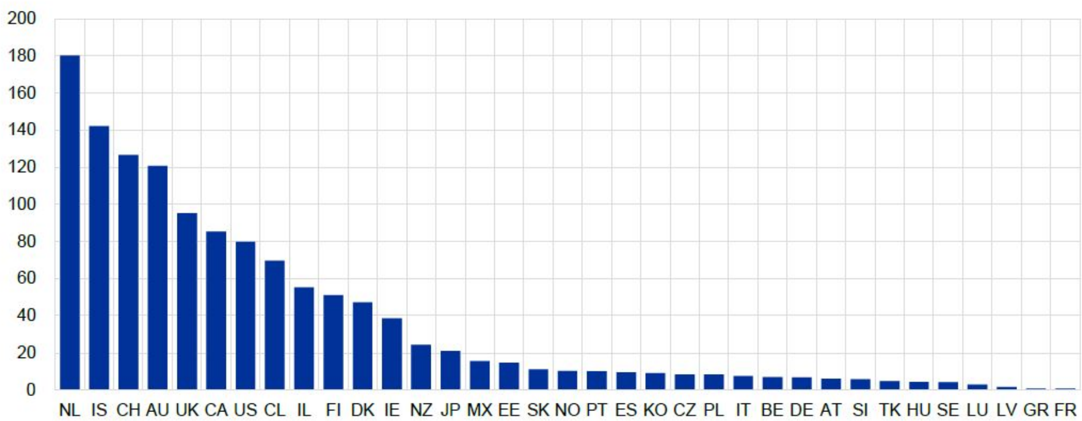
First Pillar pension schemes are typically publicly managed mandatory PAYG systems historically based on the defined benefit principle. Now, as a result of the many reforms adopted that have led to a gradual decay of defined benefit plans, the first pillar consists mainly of notional defined contribution plans. Pillar I can be seen as a basis for all the retirees, it should guarantee an adequate pension for the whole population and its aim is to avoid poverty in old ages. In general, it is organized as a pay-as-you-go system, where benefits at retirement are determined by earnings, the number of contribution years, the accrual rate and the indexation method. The typical final payout is an annuity during retirement and it is very often indexed to inflation.
In order to reduce public liabilities, most reforms in Europe aim at reshaping statutory public pension schemes: increasing life expectancy and falling birth rates have led to an increase in the proportion of population dependent on public pension schemes. For this reason, some countries implemented reforms to ensure a dynamic adjustment to life expectancy such as increasing contribution rates, raising the retirement age and changing the eligibility criteria for early retirement schemes.

Pension schemes in Pillar II are occupational schemes that cover private employment related pensions plans. The purpose of the second pillar is to provide a reasonable replacement ratio - i.e. ratio of pension income over the last wage - to enable individuals to maintain consumption smoothing. Pillar II consists of private pension schemes that come in the form of defined benefit or defined contribution. Defined contribution plans are often collective schemes organized along occupation lines, are more flexible and allow individuals to choose their own portfolio composition and risk profile. On the other hand, in defined benefit plans the employer bears all the risks since benefits do not depend on financial market returns.
Depending on countries and workplace the enrollment in Pillar II could be mandatory: in the UK , starting from October 2012, the government made it compulsory for employers to automatically enrol eligible workers into a qualifying workplace pension scheme. [25]

Pillar III Pensions schemes are voluntary, private and marked-based plans based on the defined contribution principle. Pillar III makes some individualisation of the replacement ratio possible and is available to everyone who wishes to supplement the retirement income provided by the first two pillars. It consists of private contract between individuals and financial institutions, in which the state does not intervene, except in the case of tax incentives. These type of pension schemes can be seen as simple individual ways of saving for retirement, where all the risks are borne by the individual. One particular way to reduce costs and lower the risk exposure of Pillar III is to allow for voluntary contribution to be paid into an already existing collective occupational defined contribution scheme in the second pillar.

Several studies have found that in countries where Pillar I schemes are predominant, private savings are lower. People know that the first pillar will be able to guarantee them an adequate replacement rate and therefore do not save further by integrating the other pillars into their pension.

Figure 1.1: Autonomous Pension Funds as a percentage of GDP



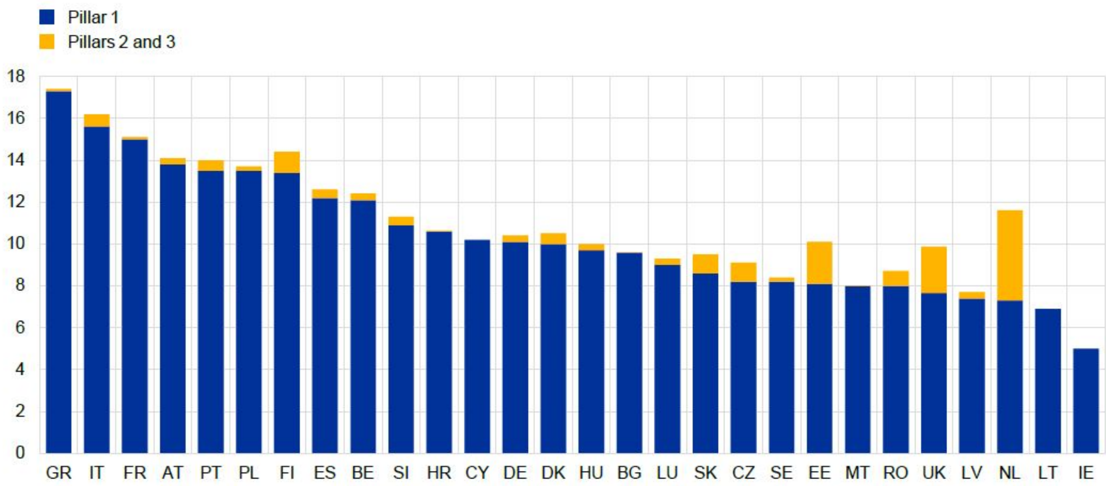
Source: OECD [29]
Note: Autonomous pensions funds represents all the assets bought with the contribution of pension plans for the purpose of financing pension benefits, thus excluding Pillar I schemes

In Figure 1.1 we can see that countries such as the Netherlands, Iceland, Switzerland and the UK (which in historical terms used Beveridgean[2] plans) invest an higher amount in autonomous pension funds, which means that these countries are characterised by a larger amount of savings in the second and third pillars.
On the other hand, countries such as Italy, Greece and France (that traditionally exploited Bismarkian[3] systems) invest very little in autonomous pension funds, meaning that the first Pillar is predominant and people decide not to supplement their savings using the other pillars, as can be seen in Figure 1.2.

[2]Under the Beveridgean system, social security benefits ensure for each citizen a basic income, a flat-rate pension independently of occupation and earnings during active employment. Typically the size of the public plan is small and it is integrated with other pillars.

[3]In the Bismarckian system benefits are earnings-related and profession-related. These programs therefore rely on a strong link between individual contributions from earnings and individual pensions. Typically the size of the public plan is large and it is not integrated with other pillars.

Figure 1.2: Size of contribution to Pillar 1,2 and 3 as a percentage of GDP



Source: OECD[29]

1.2 The importance of supplementary pensions

Demographic trends are constantly changing the face of the European continent: the proportion of the population over 65 has reached 19% and the number of people aged 80 or over is expected to more than double by 2100.[2]
Taking population developments into account, it is clear that most European countries will not be able to sustain a pay-as-you-go social security system and that Europe will face major pensions challenges.
Accepting the consequences of ageing trends is an unavoidable requirement: the decline in state pension provision can only be prevented by a steady increase in the retirement age or by supplementary pension provision.
Individuals are therefore increasingly being called on to take responsibility for their future retirement income through additional pension schemes, in order to improve the estimated lower public pensions. Consumers must be primarily responsible for their own destiny, and the second and third pillars of social security must be integrated with the first public pillar.
However, according to the results of a survey conducted by Insurance Europe, most European citizens do not supplement their pensions with additional plans and do not set aside enough resources for old age.

Insurance Europe, the federation of continental insurers, surveyed more than 10 000 people in 10 countries in February 2020 to find out the retirement preferences of people and their attitudes to saving.[15]
The first worrying conclusion of the survey concerns the inadequate number of Europeans allocating savings flows to instruments designed to supplement their basic pensions.

In the 10 countries investigated (Austria, France, Germany, Italy, Luxembourg, Poland, Portugal, Spain, Switzerland and Hungary) as many as 43% of the respondents are not saving for their future supplementary pensions. Almost half of the interviewed are not saving for retirement.

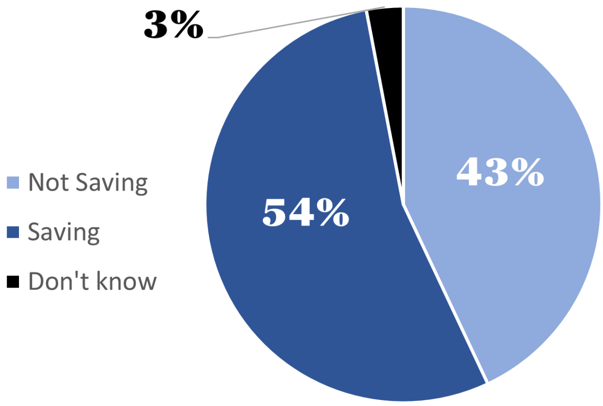
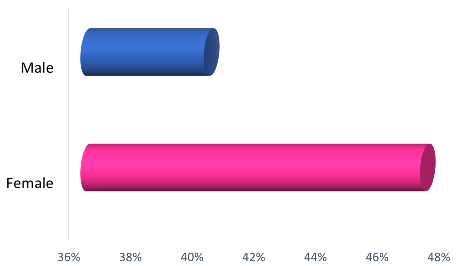


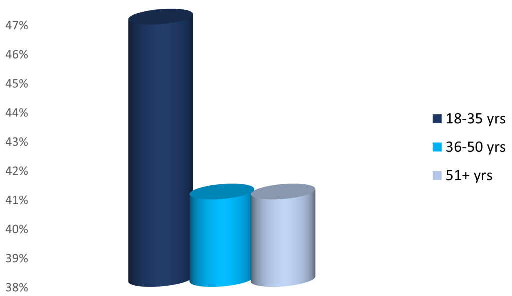
Figure 1.3: Savers for retirement through a supplementary pension

Responses are also influenced by personal circumstances such as age, gender and employment. Among women, the share of those who do not save rises to 47% and the same is true for young people aged 18-35. Level of education also influences savings choices and reduces the behavioural bias in financial choices as shown in Figure 1.4.

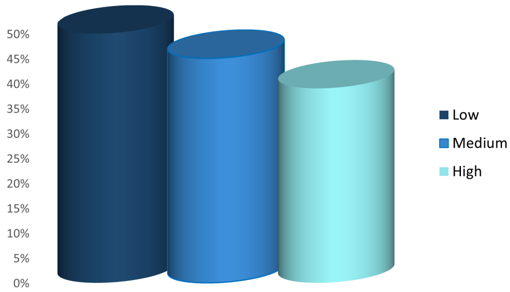
Figure 1.4: Respondents not saving



(a) By Gender



(b) By Age



(c) By education level

The survey shows that the main reasons why European citizens do not save for their retirement are as follows:

- 42% of "non-savers" can not afford it economically. In fact, given the large amount of compulsory contribution rates in the basic social security system (33% of the salary for Italian employees), there are not enough resources available to finance complementary social security.
- 15% of citizens who are not setting aside sufficient resources for old age plan to do so in the future. The delay in starting an additional pension plan is a sign of a trend well known from the behavioural finance literature: individuals systematically procrastinate on saving decisions and prefer instead immediate consumption.
- 28% of people who do not integrate their first pension pillar are not interested in the topic. The absence of awareness of their own future and the lack of information needed to make the right decisions lead people to neglect supplementary pension provision and not save enough.

Among the respondents to the European survey, by far the highest priority when saving for retirement is the security of the money invested: the guarantee is one of the most important requirements that people look for in their retirement choices.
For 60% of respondents, safety of investments is a priority. This explains why most European investors prefer insurance-based pension products. Obviously, the more uncertain and volatile the benefits offered by the first pillar become, the more savers demand greater certainty and protection from the second and third pillars. In order to plan their future efficiently, families need stability: in the past, it was the state that provided insurance for its citizens; today, individuals are asked to be more responsible for their own future, but to do so they need safeguards.

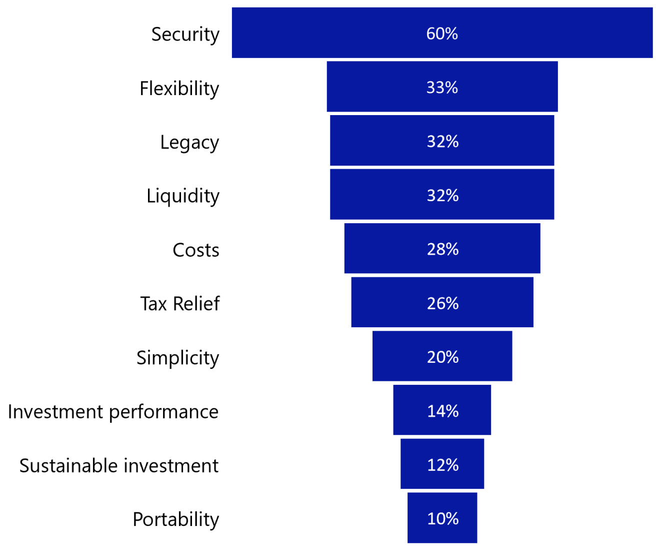


Figure 1.5: Pension saving priorities

Looking at graph 1.5, we can clearly see that flexibility is the second priority in people's social security preferences: 33% of respondents consider it essential to explore personalised and tailored solutions. Flexibility requires proactive behaviour from households: in the first pension pillar, citizens' freedom is strongly limited by rigid requirements and public finance constraints; on the other hand in complementary pension systems the choices are left entirely to the retiree. For this reason, interviewees demand flexibility at different stages of the investment: in the accumulation phase, in the allocation strategy, in the amount of contributions and also in the decumulation phase. Individuals want to be free to choose what to do with the money they have accumulated during their working life.

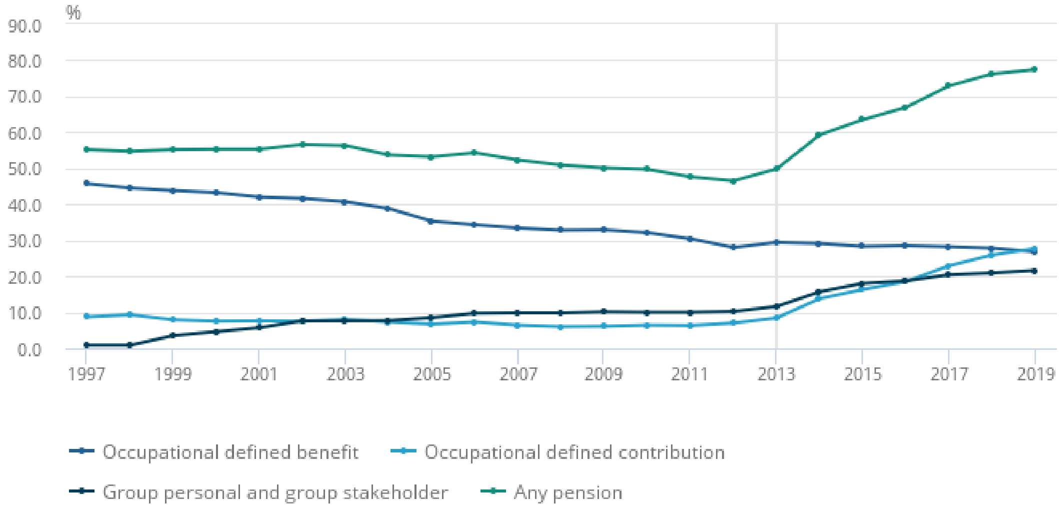
Finally, another surprising insight from the Insurance Europe survey concerns pay-out preferences. Responses to the survey show that life annuities are by far the preferred benefits for savers, chosen by 46% of respondents. In contrast, 30% of people favour flexible withdrawals, scheduled at regular intervals but for a defined period of time, and finally 24% of people chose to receive the capital set aside in a lump sum.

Choosing an annuity as the preferred form of benefit has the advantage of hedging against longevity risk by ensuring a gradual consumption of savings. However, the results of the survey should be interpreted and compared with the actual behaviour of savers: once they are faced with the real choice between a capital of 50 000€ and a life annuity of 2 500€ per year for their whole life, there is an equal split between those choosing annuities and those choosing lump sums. It is therefore not always the case that retirees prefer annuities to lump sums: it is thus necessary to give them the possibility to choose their preferred pay-out method.
In the UK a wide range of reforms have been made to the pension system that seek to meet the needs of pensioners, as shown in the European survey just analysed. In the next section I will focus on these.

1.3 Pension freedom

In order to overcome the issue discussed in the previous section, i.e. the problem that people are not saving enough for their pensions, in 2012 the UK government made important changes to how workplace pensions work.
With the introduction of the Automatic Enrolment all employers must in fact automatically enrol their eligible workers into a workplace pension scheme unless the worker chooses to opt out. There is a minimum total amount that has to be contributed by workers and employers in the form of tax relief. As a result, many more people have been able to accumulate larger savings for their retirement. Overall, occupational pension membership increased by 30 percentage points since the inception of automatic enrolment from 47% in 2012 to 77% in 2019.[25]
As can be clearly seen in Figure 1.6 after 2013 participation in workplace pensions in the UK increased a lot, mainly caused by a growth in membership of occupational defined contribution plans. With this choice, the UK government has successfully found a way to increase the savings of the population and solve a major problem.

Figure 1.6: Proportion of employees with workplace pensions in the UK



Source: Office for National Statistics[25]
Note: The "group personal and group stakeholder" category includes group personal pensions, group stakeholder pensions and group self-invested personal pensions

From 6 April 2015, another major change has been introduced in the UK: the Pension Freedoms reform. [22]
Pension freedoms only apply to defined contribution pension funds and allow people who have reached the age of 55 to decide how, when and whether to access their pension pot. Whereas previously, people almost always had to buy an annuity with their pension fund, they now have more options at their disposal. Once again in the UK the government has tried to meet the needs of pensioners, by providing them with greater flexibility during the phase of decumulation of investments.
There are indeed now six retirement alternatives that pensioners should consider:

- Leave the pension pot untouched and invested in the market. There are no requirements for the pensioner to take his money immediately; he can decide to keep the pension fund intact and take his money later on.

- Get a guaranteed income. Retirees can use their entire pension pot to buy an annuity, which provides them with a regular and guaranteed income for life. There are many different types of annuity and the benefits received will depend on the specific characteristics of individuals and their lifestyle.

- Seek an adjustable income called flexi-access drawdown. With this option retired people can take 25% of their pot as a single, tax-free lump sum; the remaining 75% of the fund stays invested in the market to provide them with a regular taxable income. The value of the fund that remains untouched can go up or down as it is invested in the stock market; therefore the amount of financial risk involved in this option is high, and some advice on how to manage investments may be needed.

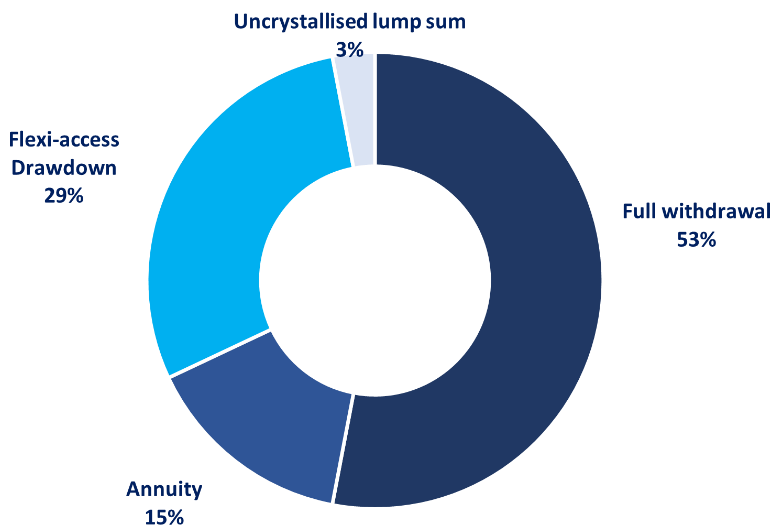
- Take money in chunks. It is another form of drowdown option where the pensioner can decide how much and when to take money until the pension pot runs out. Each time individuals take chunks of money, 25% is tax free and the rest is taxable. This option is known as Uncrystallised Funds Pension Lump Sum (UFPLS) and once again, it is important to rely on financial advice in order to make the fund last as long as possible. There are many drawdown rules proposed by financial literature that I will describe in the next section.

- Cash in the whole pot in a lump sum. Individuals can withdraw the entire pension fund in one single amount, but usually it is subject to taxes. If people opt for this option, they should be cautious to avoid spending all their money in one go.

- Mix the above options. One of the most important things of the Freedom Reforms is that retirees can really have freedom about what to choose. They are not forced to select just one option, they can mix them over time or on their total pot.

In the following graph we can see how pensioners reacted to pension freedom and which were their decumulation preferences: many consumers withdrew their pension pot in full or opted for drawdown products.[10]

Figure 1.7: Consumer's choices after pension freedom



Source: Financial Conduct Authority[10]

However, freedom to choose is not enough, people must have freedom to make informed choices.To support consumers in making their retirement choices, the government has indeed introduced Pension Wise[22], a national pension guidance service offering free and impartial guidance for people aged 50 and over. The service operates online, via telephone or face to face and provides

guidance on pension options by looking for the most suitable alternative for the individual and by planning long term strategies. But evidence suggests that a considerable proportion of consumers, particularly those with a low accumulated pension pot, are making decumulation decisions without taking advantage of the service offered by the government and without seeking professional financial advice.
The Financial Conduct Authority (FCA) Retirement Outcomes Review found that the proportion of consumers purchasing drawdown products without advice grew from 5% before the Pension Freedoms legislation to 30% in 2016.[10] A large number of people started buying drowdown products, which are characterized by high financial risk and can ruin out, without fully understanding the consequences of their choice. These products may offer greater returns but are riskier. Precisely for this reason, a high degree of flexibility in choices should be combined with knowledge and consciousness.

In the next section I will review and analyse in detail some of the decumulation strategies proposed in the literature, especially focusing on the drowdawn environment, in order to provide greater disclosure on this issue.

1.4 Decumulation Strategies

As I have extensively discussed in the previous section, decumulation may be an immediate decision, taken at the time of retirement, or it could be a series of decisions taken throughout the pensioner's life. The retiree can either make a unique choice - i.e. use all his accumulated wealth to buy a life annuity, or withdraw the entire fund in one go - or he could adopt and update a decumulation strategy for his entire lifespan.
The main problems are faced in the second case, as costumers have to consider both investment and withdrawal strategies. Individuals should indeed decide how much to withdraw periodically from the accumulated fund, considering that the lifespan of anyone is unknown, and choose the optimal investment strategy to follow for the remaining pension savings. These two issues must be assessed together: it is not enough to choose an optimal investment strategy if people badly decumulate their fund. Pensioners should therefore be driven by advisors who manage their portfolios efficiently, to ensure them a sustainable income for a lifetime. Certainly this is a difficult problem to settle, but I will review in the following subsections some proposed approaches in the literature. I will cover the analysis of many approaches, including: Drawdown rules derived from experience, Utility theory-based methods, such as Expected Utility Theory and Cumulative Prospect Theory; techniques based on the minimization of the difference between current consumption and a desired consumption; probabilistic methods, such the minimization of the probability of ruin; habit formation approaches.

1.4.1 Drawdown Rules derived from experience

In this section I outline the main approaches widely used in practice in a drawdown environment.

Whereas life expectancy is unknown, common practical advices on drawdown are often based on the simplified assumption that there exists a fixed lifespan which is considered unlikely to be exceeded. Based on historical series or simulated data, consultants makes for instance the assumption that the maximum number of years left to live after retirement are 30, taking into account a retirement age of 65 years.

One possible drawdown strategy could therefore be the 1/30 rule.[4]
The idea is to split the fund accumulated into thirty equal proportions and annually pay out one part to the pensioner. In order to maintain the long-term value of each segment, the fund is invested in the financial market: the amount of the withdrawal will therefore fluctuate directly with investment returns. The main problem with this type of strategy is however that if the pensioner lives more than 30 years, he outlives its resources, exhausts his fund too early and ends up with no savings before he dies.

Similarly another widely used rule is the safe withdrawal rate (SWR) , sometimes known as the "4% rule" when the SWR is set equal to 4%.[4]
The retiree invests in a constant-mix strategy of bonds and stocks throughout a 30 years retirement period and he withdraws an inflation-linked income each year. This method tries to prevent worst case scenarios, related to the uncertainty of life expectancy, market returns and future expenses, by encouraging pensioners to withdraw only a small percentage of their fund each year, typically 3% to 4%. There is intended to be a 90% probability of success for this type of strategy, meaning that in 90% of cases the income is expected to last for at least 30 years. The numerical value of the safe withdrawal rate is the first year withdrawal income divided by the pension fund at the time of retirement.
For instance if the accumulated savings are $800 000 and the retiree withdraws $35 000 as income in the first year, the SWR would be 4.3% (35 000 / 800 000). Assuming that the inflation is 5% over the first year, the second year's withdrawal amount would be $36 750 ($35 000 · 1.05)
The idea is to maintain the real value of the first year's income for the entire lifespan. This strategy pays a stable income in real value terms to the retiree, irrespective of market returns, but also in this case the fund may no last for 30 years.
SWR has been criticised on two further reasons :

- For the stable real value income: Milevsky and Huang[17] illustrate a life-cycle problem, underlying how the longevity risk aversion affects the retirement spending rates. They find that only very high longevity risk aversion leads to stable withdrawal rates.

- For the constant mix investment strategy: Scott et al. [28] point out the inefficiencies of the SWR approach by showing that options or dynamic investment strategies could ensure better outcomes for the same cost or same outcomes at lower cost. Therefore they shown that a retiree using a 4% rule follows a wasteful strategy in many possible scenarios, as there are wasted surpluses when risky investments outperform, which leads to an excess of asset at the end of 30 years.

Hence, another drawdown rule is to make the income withdrawal dependent on investment returns. This rule results in a dynamic income strategy that adjusts the withdrawal income according to what investment returns and inflation have been experienced. The main idea is that the retiree cash in part of his investment gain when returns have been particularly good, while the income is reduced when returns have been negative. The investment strategy remains constant, a proportion of the fund is always invested in equity, but according to the market, the income is adjusted.
Guyton and Klinger[14], based on their expert judgment, established principles of dynamic income strategies and showed that their rules lead to higher safe initial withdrawal rates compared to previously published results, which do not take into account income adjustments. They defined the probability of success as the chance that at least $1 remains in the retiree's account after 40 years of withdrawal. Then they tested their results considering three different types of equity allocation: 50%, 65% and 80%. They found that when portfolios contains at least 50% shares, initial withdrawal rates of 4.6% are sustainable at a 99% confidence level. Whereas portfolios composed of a larger amount of equity allocation provide more purchasing power - i.e higher level of initial withdrawal rate - while maintaining a slightly lower probability of success.

Another possible rule to be adopted in the drawdown environment is a dynamic investment strategy. Now withdrawal rates vary during the lifespan, according to the asset allocation strategy. Many studies shown that rising equity glide-paths in retirement – where the portfolio starts out conservative and becomes more aggressive through the retirement time horizon – have the potential to actually reduce both the probability and the magnitude of failure for client portfolios. Fullmer[12] argues that this strategy is strictly related to the longevity risk: longevity risk tends to decrease with age, given that the probability of outliving resources decreases with time; so by contrast the investor can afford to take more investment risk and to increase his exposure to equity as time passes.

1.4.2 Expected Utility approach

Another common decumulation strategy is to model explicitly consumer's preferences through utility functions.
According to the expected utility theory, the way to determine the optimal amount of consumption, and the optimal amount of wealth invested in risky

assets, is to maximise, with respect to some decision variables such as consumption and/or investment strategies, the expected discounted utility of consumption:

$$\mathbb{E}\left[\int_0^T e^{-\rho s}u(c(s)) + e^{-\rho T}v(X(T))\right] \tag{1.1}$$

where:

- u represents the utility of consumption and reflects preferences and risk aversion behaviour of the investor.
- $c(s)$ represents consumption, i.e. the amount which must be withdrawn from the pension fund at each point in time s.
- $e^{-\rho s}$ is a discount factor, that accounts for the time preferences of individuals, since more distant consumption is less liked. ρ is called the time-preferences rate.
- T is the terminal date: usually it is a random variable since it represents the life expectancy.
- v is the terminal condition. It reflects the preferences of the retirees concerning the importance of remaining savings at the terminal date T. It could reflect the intention to leave a bequest or to ensure that as little as possible is left.
- $X(T)$ represents the pension fund at time T. Usually the dynamic of the fund is modeled according to the Black and Scholes rule: the market consists of only one riskless asset earning interest at a constant rate and one risky asset whose price follows geometric Brownian motion.

$$dX(t) = [rX(t) + (\mu - r)\alpha(t)X(t) - c(t)]dt + \alpha(t)\sigma X(t)dW(t) \tag{1.2}$$

$$\text{X(0)} = \text{Initial Fund}$$

where μ is the mean rate of return of the risky stock, r is the risk free rate, σ is the volatility of the risky stock and W is a standard Brownian motion. $\alpha(t)$ represents the proportion of wealth invested in risky stock at time t, and usually it is a decision variable. The idea is that the fund evolves according to this rule: at each point in time it is equal to the fund at the instant before, plus earnings from risky and riskless investments, minus withdrawals.

Some Examples

A large number of authors have tried to implement an expected utility maximisation problem in order to find the optimal level of withdrawals that a pensioner should follow. There are many standard choices for utility functions, but some shapes are more suitable to obtain close form solutions.
One of the most well know and widely used class of utility function is the CRRA, which stands for constant relative risk aversion.

$$u(c) = \frac{c^{1-\gamma}}{(1-\gamma)}$$

where γ reflects the level of risk aversion of individuals.

As mentioned in the previous section, Milevsky and Huang[17] criticised the SWR rule because of the smoothness of withdrawal rates, and in doing so, they solved an expected utility maximisation problem. The authors deliberately ignored market risk to focus on the role of longevity risk aversion in determining optimal consumption (or spending) during a period of stochastic length. They thus assumed that a pensioner faces random lifetime, but that investment returns are known and unvarying: only life spans are stochastic. Moreover, an additional fixed income is expected to be obtained as state pension or annuity and a CRRA utility function is considered.

Computationally they solved the following problem:

$$\max_c \left[\mathbb{E} \int_0^T e^{-\rho s} u(c(s)) ds \right] = \max_c \int_0^T {}_s p_x \cdot e^{-\rho s} \cdot u(c(s)) ds \tag{1.3}$$

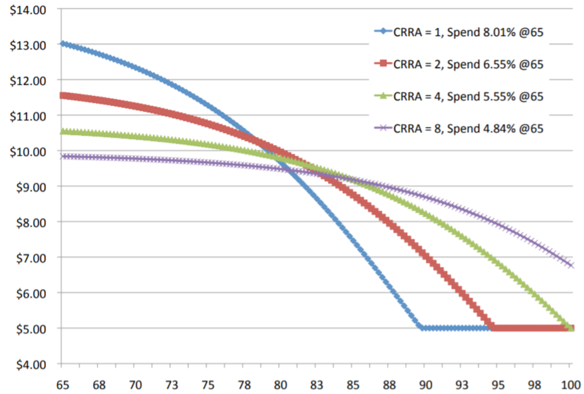
where ${}_s p_x$ is the survival probability for an x-year old to live up to $x + s$ years, parameterized using Gompertz-Makeham law of mortality.
The pension fund's wealth X now obey:

$$dX(t) = [rX(t) - c(t) + \pi]dt \tag{1.4}$$

where π is the (constant) pension income.
They solved the problem by considering different levels of longevity risk aversion γ and some important results are shown in Figure 1.8

Milevsky and Huang found the optimal amount of money that should be withdrawn at any given point in time, for different levels of risk aversion. Assuming a 65-year-old with a (standardized) \$100 initial accumulated fund, we can see that the optimal consumption path is decreasing with age, and that only high level of risk aversion lead to smoother withdrawal.
They point out that only under a very limited set of implausible preference parameters, life-cycle consumption are smoothing and for this reason they criticized the SWR rule.

Figure 1.8: Optimal consumption path with $\pi = 5\$$, r=2.5% and $X(0) = 100$



Note: Consumption at each age is equal to the sum of the constant pension income and the optimal withdrawal amount. For instance the consumer with a CRRA of 1 (i.e., very low aversion to longevity risk and $u(x) = ln(x)$ will start retirement by withdrawing \$8 from his fund and by receiving a pension income of \$5. Its total initial consumption is therefore equal to \$13.

Another important study that seeks to find optimal decision rules by maximizing the expected discounted utility function was developed by Andésson et al.[1]
It is an Australian study, in which the pensioner receives a fixed income during retirement through a state pension, and he should decide how much to withdraw from his accumulated fund and how much to invest in risky assets in each point in time. In addition, the model extends the setting described by other studies by allowing the retiree to choose the proportion of wealth H allocated to housing at retirement. It is a more complicated study that takes into account the randomness of stock's returns, that allows the proportion of fund invested in equity to vary over time, and also involves bequest in the utility function. The authors solved the following maximization problem:

$$\max_{c,\,\alpha,\,H} \mathbb{E}\left[\int_o^T e^{-ps}u(c(s),H) + e^{-pT}v(X(T))\right] \tag{1.5}$$

Considering that the dynamic of the pension fund now satisfies:

$$dX(t) = [rX(t) + (\mu - r)\alpha(t)X(t) - c(t) + P(t)]dt + \alpha(t)X(t)\sigma dW(t)^4 \tag{1.6}$$

Subject to:

$$\text{X(0) = Initial Fund - H}$$

They calibrated the model with Australian empirical data for consumption and housing, and estimated suitable parameters via the maximum likelihood method. They were therefore able to find the optimal housing level H at retirement, optimal consumption c(t) and optimal risky asset allocation $\alpha(t)$ for each time, considering different levels of initial wealth.

[4] P(t) is the means-tested Australian pension

1.4.3 Minimize distance from a Target

Optimal solutions obtained by maximising the expected utility function usually lead to constantly varying income during retirement. To tackle this problem, another possible decumulation approach could be the introduction of a reference consumption level to be met at each age.
The idea is to set desired levels of intermediate consumption at each age and to fix a final desired level of wealth. By minimising the distance between actual consumption and desired consumption over the lifetime of the retiree, it is possible to estimate optimal investment and consumption strategies.
A quadratic function is usually applied to measure the distance from withdrawal income and fixed target. Thus the problem consists in minimizing:

$$\mathbb{E}\left[\int_0^{T_D} e^{-\rho s}(c(s)-b(s))^2 ds + \theta e^{-\rho T_D}(X(T_D)-F(T_D))^2\right] \tag{1.7}$$

where

- $c(s)$ represents actual consumption at time s.
- $b(s)$ represents the desired level of consumption at time s.
- $X(T_D)$ is the actual level of final wealth, and it follows a dynamic according to the Black and Scholes model.
- $F(T_D)$ represents the desired level of final wealth at time T_D.
- T_D is the final date, which could be random since it represents the individual's life expectancy.
- $e^{-\rho s}$ is a discount factor, that accounts for the time preferences of individuals.
- θ is a constant which reflects the importance of final wealth target compared to intermediate consumption levels.

Some Examples

A significant study by Gerrard et al.[13] on defined contribution pension plans employed the technique of minimising the distance from a target to find optimal withdrawal and investment strategies.
In this study an important assumption is made: the pensioner chooses to defer the annuity, meanwhile consuming some income withdrawn from the fund and investing the remaining fund. The authors hypothesized that the reasons for the retiree to choose the option of deferring an annuity are either the hope of being able to buy a higher annuity in the future than the pension income provided by the immediate annuity at retirement or the possibility of bequeathing wealth in the event of death before annuitization. They assumed that the pensioner has three different types of goals during the decumulation

phase: a target for the size of the fund, a desired level of income to be consumed and a desired level of annuity. The idea is that a disutility is experienced whenever deviations from targets occur, and a utility is experienced in the event of death before full annuitisation, due to the bequest motive.
They described:

- A disutility function due to deviations (from desired consumption and fund targets) before annuitization:

$$L(t, x) = e^{-\rho t}[\mu(F(t) - x)^2 + v(b_0 - b(t))^2] \tag{1.8}$$

  Where F(t) is the running target for the level of the fund at age t, x is the actual value of the fund at time t which evolves according to (1.2) , b_0 is the target level of income periodically withdrawn from the fund during the drawdown phase and b(t) is the actual amount of consumption at time t. μ and v are weights given to the desire to monitor the fund and the daily consumption and ρ is the usual discount factor.

- A disutility function due to deviations (from desired annuity) at the time of annuitization T:

$$K(T, x(T)) = e^{-\rho T}[(b_1 - k \cdot x)^2] \tag{1.9}$$

  Where b_1 represents the target value of income from the annuity purchased at age T and k can be seen as the amount of annuity provided by the insurance company at age T for one unit of capital. Compulsory annuitization occurs at time T.

- A utility function due to bequest in case of death before annuitization, at time T_D

$$M(T_D, x(T_D)) = e^{-\rho T_D} x \tag{1.10}$$

  Where x is the actual value of the fund at time T_D and T_D represents the time of death of the individual.

Optimal investment and consumption strategies can be estimated by minimising the following function, which brings together all the concepts just described.

$$\mathbb{E}\left[\int_0^{min(T,T_D)} L(s, x(s)) + \theta K(T, x(T)\mathbb{1}_{\{T_D>T\}} - \eta M(T_D, x(T_D))\mathbb{1}_{\{T_D<T\}} ds\right] \tag{1.11}$$

Where θ and η are weights given to the importance of reaching the final annuity level b_1 and to the ability to leave a bequest.
By minimizing (1.11) the authors were able to find optimal value of running consumption b(t) and optimal equity allocation $\alpha(t)$.
An important conclusion of the study seems to be that weights given to the level of running consumption v, and weights given to the realization of final

annuity θ play an important role. Only a good trade off between the two could lead to meet all the target levels desired by the pensioner.

The previous study is not able to guarantee optimal solutions which avoid ruin fully. To address this problem Di Giacinto et al.[7] added some constraints on wealth and on investment strategies. Therefore they considered the same setting of Gerrard[13], but they imposed short selling constraints and required that the final fund cannot be lower than a certain pre-determined level S ≥ 0. The idea is that they solved the usual minimization problem (1.11), without considering the bequest motive, and assuming the following constraints:

$$\alpha(t) \geq 0 \quad and \quad x(T) \geq S$$

1.4.4 Habit Formation

A more plausible representation of preferences is to allow for habit formation, meaning that the utility of a given level of current consumption is also a function of the level of past consumption.
The idea is to model a new type of utility function, which depends both on current consumption c_t and on the standard living of an individual h_t. Thus utility is not time separable but exhibits habit persistence.

$$U = u(c_t, h_t) = u(c_t - h_t)$$

where

$$h_t = e^{-at}h_0 + b\int_0^t e^{a(s-t)}c(s)ds \tag{1.12}$$
$$with \;\; h_0, b, a \geq 0$$

As can be seen from (1.12) h_t is a measure of the habit level of consumption, and it is a weighted average of past consumption rates. The weights are exponentially decreasing so that most recent consumption rates are more important. The constant h_0 is the initial habit level, a is a persistence parameter and b is a scaling parameter.

The idea is therefore the following: in order to have a positive utility the client require consumption above the habit level. The higher is the habit standard living of a client, the more he would like to consume. Moreover, the stronger is the habit persistence, the more averse is the consumer to a fall in consumption. It is an excess in current consumption over and above the habit which increases current utility.
Optimal investment and withdrawal strategies must ensure that future consumption rates are as high as past consumption rates, in order to maintain the habit formation.
Thus the "new" problem consists in maximizing:

$$\mathbb{E}\left[\int_0^T e^{-\rho t}u(c(t) - h(t))dt\right] \tag{1.13}$$

Some Examples

Munk[23] derived optimal consumption and investment policies for investors with habit persistence in preferences for consumption when the financial market provides time-varying investment opportunities. In line with most of the literature he assumed, both for tractability and for easy comparison, that the instantaneous utility function $u(c_t, h_t)$ is power-linear, i.e. that

$$u(c_t, h_t) = \frac{(c_t - h_t)^{1-\gamma}}{(1-\gamma)}$$

where γ is the usual risk preference parameter.

The author assumed that the fund evolves according with the usual dynamics proposed by Black and Scholes, and he further required that the consumption strategy is financeable, i.e. its present value cannot exceed the fund of the investor.
By maximizing (1.13) he was able to derive optimal strategies in concrete setting, such as in a stochastic interest rate model. As previously mentioned, due to habit formation, the optimal investment strategy must ensure that the habit consumption level is reached. The consumption rate is required to exceed the habit level : the habit level plays the role of a minimum or subsistence consumption rate determined by past consumption rates. The main result of the study is therefore the following: low-risk assets are better suited than higher-risk assets. The main effect on asset allocation comes from the fact that some assets (bonds and cash) are better investment objects than others (shares) when the aim is to ensure that future consumption does not fall below the habit level.

Another important study by Bruhn and Steffensen[5] allow for habit formation. Their setting is a combination of habit formation and minimization from a target. They indeed included the habit formation in their utility function but at the same time they tried to control the rate of change of consumption. Instead of allowing consumption to be rapidly adjusted, they minimized a quadratic objective function over a fixed time horizon, allowing also for a bequest motive. The general idea is to minimize the distance between changes in consumption rate and a target. The control variable is not consumption but the acceleration of consumption over time. The motivation is that people are very focused on the increase in welfare, regardless of their current level of consumption.
Their model, roughly speaking, consists in minimize (1.8) but instead of control the distance between optimal consumption $c(s)$ and a desired level of withdrawal, they minimize the gap between $[c(s) - h(s)]$ and a target level of increase in wealth.

1.4.5 Cumulative Prospect Theory - based objective function

A substantial body of evidence shows that decision-makers systematically violate the Expected Utility Theory and deviate from full rationality. For this reason, new theories have been developed in response to empirical challenges and the Cumulative Prospect Theory is one of the alternative models proposed.

The Cumulative Prospect Theory is a descriptive theory as it is based on the real behaviour of individuals and is able to model real needs and attitudes of consumers. The Expected Utility Theory, on the other hand, attempted to give common and general rules, by assuming a full rationality of investors and without taking into account behavioral biases.
Tversky and Kahneman[31], based on experimental evidence, proposed the CPT and modeled the expectation of an event as:

$$\mathbb{E}\left[v(x)\right] = \sum_{i=-m}^{n} \pi_i v(x_i) \tag{1.14}$$

where

- x_i represents gain or loss with respect to a reference point. If for instance consumption is the relevant object and we set the reference level as c^0, then every outcome below c^0 is defined as a loss and represented with a negative number $c_i - c^0 = x_i < 0$. Similarly outcomes above c^0 are defined as gains and represented with a positive number $c_i - c^0 = x_i > 0$. The m losses and n gains are moreover ordered such that

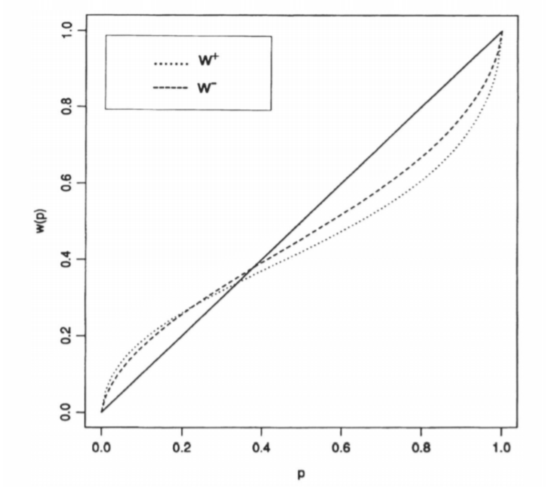
$$x_{-m} \leq x_{-m+1} \leq ... \leq x_{-1} \leq x_0 \leq x_1 \leq ... \leq x_{n-1} \leq x_n$$

- π_i are not the usual probabilities of an event, but are probability weights, calculated through probability weighted functions. Typically $\pi_i = w(p, x)$ where $w(\cdot)$ is the probability weighted function and usually it is not a probability measure; p is the usual probability associated to an event and x represents the gain/loss. There are two different weighted function for losses and gains. For instance Tversky and Kahneman proposed

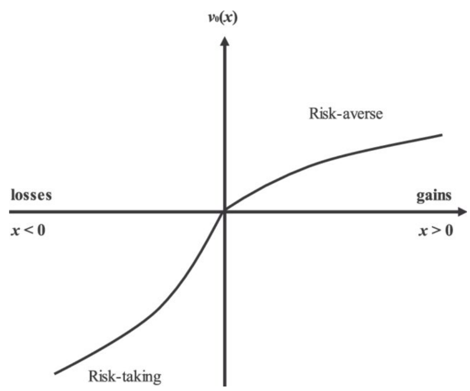
$$w(p, gain) = w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$$

$$w(p, loss) = w^-(p) = \frac{p^\lambda}{(p^\lambda + (1-p)^\lambda)^{\frac{1}{\lambda}}}$$

  estimating the parameters $\gamma = 0.61$ and $\lambda = 0.69$.
  This type of decision weights are able to capture an important behavioural aspect of investors: they tend to overweight events that have a low probability and to underweight events that have moderate or high probability. As can be clearly seen from the figure below, $w(p) > p$ for small p, while $w(p) < p$ for high p, both in the case of gains and losses.



- $v(\cdot)$ is a value function that assigns numerical values to gains and losses x_i. The value function reflects investors' preferences and their attitude to risk. Unlike the standard concave utility function under the Expected Utility Theory, here the value function is S-shaped, i.e. it is concave in the gain domain, and convex in the loss domain. The idea is that the consumer is risk-averse above the reference point, preferring to lock into a gain rather than gamble the possibility of a larger gain; while the investor is risk-lover below the reference point, as he prefers to risk a large loss - with the possibility of moving into the gain domain - rather than lock into a certain one.
  Moreover the value function is able to capture another important behaviour: people are loss averse. They suffer more from a loss of an amount of money than they enjoy when they earn the same amount. A financial gain of \$500 does not compensate the investors for a financial loss of \$500. For this reason the value function is steeper in the loss domain, while it is flatter for gains.
  An example of value function proposed by Tversky and Kahneman is the following:
  $$v(x) = \begin{cases} x^{\alpha} & \text{if x}\geq 0 \\ \lambda(-x)^{\beta} & \text{if x}<0 \end{cases}$$
  where α and β define the curvatures and λ represents the degree of loss aversion.
  Here the graph of the value function proposed by Tversky and Kahneman.



It is technically much more difficult to derive analytical solutions for an investor who behaves in line with Cumulative Prospect Theory compared to Expected Utility Theory. The first issue to overcome is the research of the reference point against which gains and losses are defined. Furthermore the value function is partly concave but partly convex, which makes it more difficult to obtain close form solutions. In addition decision weights, that replace probabilities, add further complexity to the problem.

Some Examples

Van Bilsen et al.[32] explored and derived optimal portfolio and consumption choices of investors who behave in line with CPT. They maximized the expected value function of future consumption over a dynamic reference consumption level. Indeed, the consumption reference level changes over time according with the habit formation based approach, and it is endogenously updated over time. The problem consists in maximizing:

$$\mathbb{E}\left[\int_0^T e^{-\rho t} v(c_t, \theta_t) dt\right] \tag{1.15}$$

where

- The expectation is calculated through standard probabilities rather than with decision weights.

- $$v(c_t, \theta_t) = \begin{cases} (c_t - \theta_t)^{\gamma_G} & \text{if } c_t > \theta_t \\ -k(\theta_t - c_t)^{\gamma_L} & \text{if } c_t < \theta_t \end{cases}$$

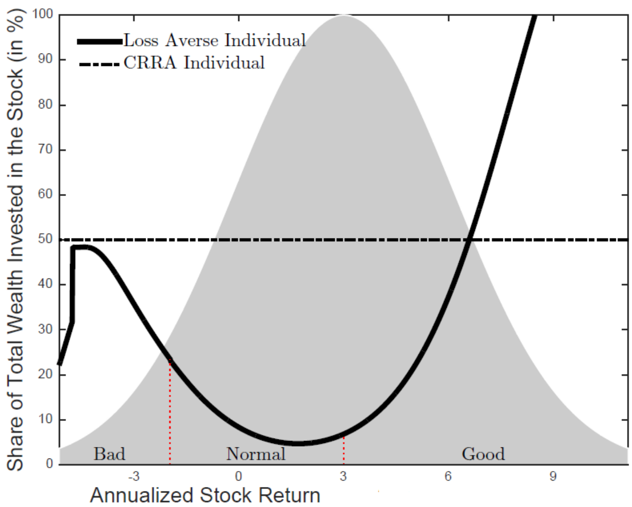
  It represents the usual value function under Cumulative Prospect Theory which takes into account loss aversion of investors.
  A standard assumption of conventional life-cycle models is that relative risk aversion is constant. Here the relative risk aversion depends on how close current consumption is to the individual's reference level.
  The authors showed that consumer can fall into three category of states of nature: normal states in which consumption remains at (or slightly above) the reference level; good states in which consumption is substantially above the reference level; and bad states in which consumption falls below the reference level. This leads to a state-dependent relative risk aversion: a loss-averse individual chooses optimal consumption and portfolio strategies depending on the state in which he ends up.

- $$\theta_t = e^{-\alpha t}\theta_0 + \beta \int_0^t e^{-\alpha(t-s)} c_s ds$$

  It represents the reference level of consumption constantly updated on the basis of past habits: it is a weighted average of past consumption with the last five years contributing at least 80% of the weight.

The authors solved for the optimal dynamic consumption and investment strategy.[5] They found that loss averse investors protect current consumption after a decrease in their pension savings, postponing reductions in current consumption following a wealth drop.
Furthermore, the individual's optimal portfolio weight in the risky stock is not constant across states of nature as is predicted by the conventional CRRA model, but rather depends on the state of nature. The individual implements a conservative investment strategy in normal states and typically a more aggressive strategy in good and bad states, as can be clearly seen in figure 1.9.

Figure 1.9: Optimal Portfolio choice of loss-averse individuals



Note: The figure shows the optimal investment strategy of a 65-year-old loss-averse individual as a function of the annualized stock return. The dash-dotted lines show the behavior of a CRRA individual with relative risk aversion equal to 2. The gray areas represent the probability density function of the annualized stock return.

We can notice that in normal states, a loss-averse individual adopts a (very) conservative portfolio strategy to prevent that consumption falls below the reference level. The optimal portfolio strategy is much less conservative in good states of nature. Indeed, in good states, a relatively aggressive portfolio strategy most likely does not lead to a loss with respect to the reference level in the near future. In bad states, a relatively aggressive portfolio strategy will increase the chance of realizing a future gain with respect to the reference level (but can eventually deplete consumption).

[5] They converted the individual's optimization problem with endogenous updating of the reference level into an equivalent dual problem without endogenous updating of the reference level. They solved the dual problem by using the martingale approach and then they transformed the optimal solutions of the dual problem back into the optimal solutions of the individual's original problem.

Van Bilsen et al.[32] also computed the welfare loss (in terms of the relative decline in certainty equivalent consumption) associated with the conventional CRRA strategy and analyzed the roles played by the reference parameter. When the reference consumption is fixed at the level θ_0 the loss averse individual suffers a minimum welfare loss of 3%, whereas when the individual has an endogenous reference level constantly updated , the minimum welfare loss is likely to exceed 10%.
The significant welfare losses incurred by the implementation of less suboptimal conventional policies, such as the CRRA, highlight the importance of an appropriate model capable of describing consumer behavioural aspects.

1.4.6 Minimizing the Probability of Ruin

Another way of determining optimal investment and withdrawal strategies is by using probabilistic methods. A number of papers have considered the problem of minimizing the probability that a retiree outlives his wealth, also known as the probability of ruin in retirement. The approach is based on minimising the probability that a retiree will end up with no savings before he or she dies: the idea is to avoid depletion of resources before death. The arguments in favour of this approach are:

- It is a more general criteria which does not take into account investors' risk aversion. Contrary to the expected utility theory approach in which the goal of maximizing expected discounted utility of consumption and bequest depends on a subjective utility function - i.e on subject risk and time preferences - here the problem appears more objective.
- Communicating the approach of minimizing the probability of being ruined is easier than communicating the idea of lifetime utility to a consumer.

Mathematically the problem consists in minimizing the probability of ruin before death, starting at time t with wealth x:

$$\mathbb{P}\left[\tau_a < \tau_d | X(t) = x, t_d > t\right] \tag{1.16}$$

where:

- τ_d represents the random time of death of an individual.
- τ_a represents the first time at which the fund value X hits the ruin level w_a, which is the minimum fund level that can be reached.

$$\tau_a = \inf\{t \leq 0 : X(t) = w_a\} \tag{1.17}$$

- X is a stochastic process which represents the investor's fund value and it evolves according to the usual formula (1.2)

Closed-form solution are difficult to obtain for ruin problems and often simplified assumption must be made in order to obtain analytical results. For instance, the retiree's force of mortality is usually assumed deterministic or the withdrawal amount is taken as constant.

Some Examples

The first paper to determine analytically the optimal investment strategy that minimizes the probability of ruin is by Young[33]. She considered two possible withdrawal strategies: (1) the individual consumes a constant amount c, (2) the retiree consumes a fixed percentage of his wealth pX. She assumed a constant force of mortality and a financial market consisting only of a risky asset (which follows a Geometric Brownian Motion) and a risk-free asset: as already mentioned simplified assumptions are essential to obtain a tractable problem. She was able to determine the optimal investment rule that tells individuals how much money to invest in the risky asset for a given wealth level and to calculate the probability that individuals outlive their wealth if they follow the optimal investment strategy proposed - i.e. the probability of ruin.
She also determined the sensitivity to various parameters of the optimal investment strategy and the main conclusions reached are:

- The higher the mortality rate, the lower the proportion of wealth invested in the risky asset. If the probability of dying is higher, the individual is less likely to outlive his or her savings and therefore does not need to invest as aggressively in the risky asset to ensure the consumption rate for so many years. The idea is as follows: since the individual has a shorter expected lifespan, he can afford to make less risky investments (which guarantee higher returns in expected value) and to invest more conservatively.

- The higher the volatility σ^2 of the risky asset, the less consumers invest in the risky asset. Consumers would like to ensure that their wealth does not reach the level of ruin, so they prefer to remain more conservative when the risky investment becomes too dangerous and returns too volatile.

- The higher is the consumption rate, the more people invest in the risky asset. In order to support an adequate and high level of consumption, consumers must try to obtain higher returns: only investing in riskier positions can ensure this goal.

- The optimal investment amount in the risky asset is a positive, decreasing and linear function of wealth. Wealthier consumers can afford to be more conservative and sustain their ideal level of consumption with their own savings, even without investing aggressively in risky activities.

In general, for the most part of the paper, results are realistic and reliable: changes in the ruin probability and the asset allocation are consistent with financial intuition and empirical evidence. However, when consumption is constant, Young[33] found that for wealth near zero, the optimal strategy is a heavily-leveraged position in the risky asset: in an attempt to escape ruin, the investor tries to achieve higher returns by borrowing money and investing them in the risky asset. Although the objective of minimizing the probability of lifetime ruin is intuitively appealing, this leveraging at low wealth is not.

To avoid this leveraged situation Bayraktar and Young[3] considered two model changes: one completely eliminates the leveraging by imposing borrowing and short selling constraints on the investor; and one reduces the leveraged situation by allowing the consumer to borrow money but only at a rate that is higher than the rate earned on the risk-free asset. The authors used the same set-up as Young, considering again the two forms of the consumption function and assuming a constant force of mortality. By minimizing the probability of ruin they were able to find optimal investment strategies.
Under the no-borrowing constraint, when the consumption function is constant, the optimal investment strategy results a truncated version of the optimal investment strategy in the unconstrained case: in fact, when the investor's wealth is above a reference level, the investment strategy turns out to be the same as it is under the unconstrained case; whereas when the investor is close to the ruin level he invests all his wealth in the risky asset.

1.4.7 Annuitization

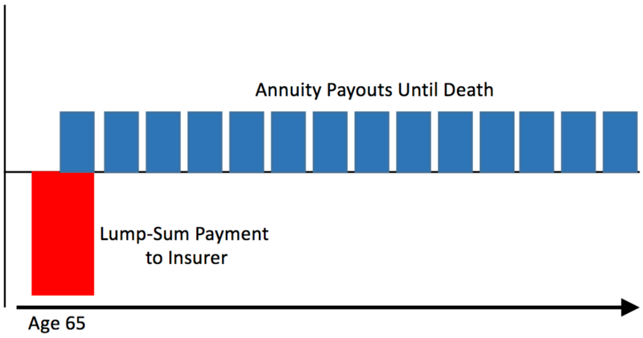
An alternative strategy able to significantly reduce or to completely eliminate the chance of running out of money before death is the annuitization. People can convert their accumulated fund into a specified income stream for a certain period of time, or for their whole lifespan. Annuities are thus products that are capable of eliminating longevity risk for the consumer and transferring it to the insurance company. Economic theory suggests that people should incorporate an annuity into their decumulation strategy, particularly when there is not bequest motive or when the longevity risk aversion prevails. Yet there are just few countries where people buy them voluntary. As mentioned

Figure 1.10: Composition of product purchases (by pot size) (%)

Source: Financial Conduct Authority[11] in the previous subsection, in the UK since 2015 consumers can freely choose their optimal decumulation strategy. However, if we look at the figure above, which shows the percentage of products purchased between March 2017 and October 2018 (for different pot sizes), we can clearly see that the amount of annuities purchased voluntarily is actually very low.

There are many different types of annuity product available on the market: in this subsection I will describe just a few.

- Immediate Lifetime Annuity paid in advance: it is a contract between an individual and an insurance company that pays the annuitant a guaranteed income for life starting almost immediately. Individuals typically buy immediate lifetime annuities by paying a lump sum of money to the insurer at time 0, i.e. when the consumer decides to annuitize. The insurance company, in turn, promises to pay the annuitant a regular income, according to the terms of the contract. Regular payments start to be paid at time 0 and continue to be paid at the beginning of each time-interval until the death of the insurer. Here a graphic representation of the policy, if we assume an insured aged 65.



  The amount of annuity payouts is calculated by the insurer, based on factors such as the pensioner's age, prevailing interest rates, and how long the payments are to continue, i.e mortality rates.
  Ideally, the expected present value of the payment stream matches the initial premium paid by the retiree, making it a fair annuity. Assuming annual and unitary payments, the present value of future benefits is a random variable Y with distribution:

$$Y = \begin{cases} 1 & q_x \\ 1+v & {}_{1|1}q_x \\ 1+v+v^2 & {}_{2|1}q_x \\ ... & ... \\ 1+v+v^2+...+v^m & {}_{m|1}q_x \\ ... & ... \end{cases}$$

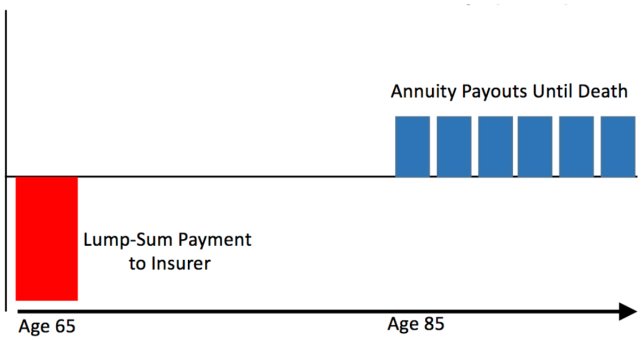
  v represents the discount factor equal to $1/(1+i)$, where i is the constant interest rate of the market.

${}_{m|1}q_x$ represents the probability that a retiree aged x dies at an age between $x + m$ and $x + m + 1$. The general idea is the following: if the pensioner dies during the first year his benefit would be only 1; if he dies between $x + 1$ and $x + 2$ the present value of his benefit will be $1 + v$; if he dies between $x + 2$ and $x + 3$ it will be $1 + v + v^2$ and so on...
The expected present value of Y, as well as the single premium paid at time zero for an immediate lifetime annuity, is:

$$\mathbb{E}[Y] = \sum_{h=0}^{+\infty}(1 + v + v^2 + ... + v^h)\ {}_{h|1}q_x$$
$$= \sum_{h=0}^{+\infty} v^h \cdot {}_hp_x \overset{def}{=} \sum_{h=0}^{+\infty} {}_hE_x \overset{def}{=} \ddot{a}_x$$

${}_hE_x$ is a mathematical notation denoting the premium that must be paid at age x in order to receive a unit payment in h years.
$\ddot{a}_x$ is an actuarial notation defining the expected present value of an immediate lifetime annuity paid annually in advance to a policyholder aged x.

- Deferred Lifetime Annuity paid in advance: it is a contract very similar to the Immediate Lifetime Annuity, but instead of providing benefits from time 0, the insurance company defers payments. The pensioner pays the single premium at time 0 but regular incomes start to be paid out at time $m > 0$ and continue to be paid at the beginning of each time interval until the retiree dies. Here a graphic representation of the policy, if we assume an insured aged 65 and deferral time of 20 years.



In this case, considering annual and unitary payments, the distribution of the random variable Y of the present value of future benefits is:

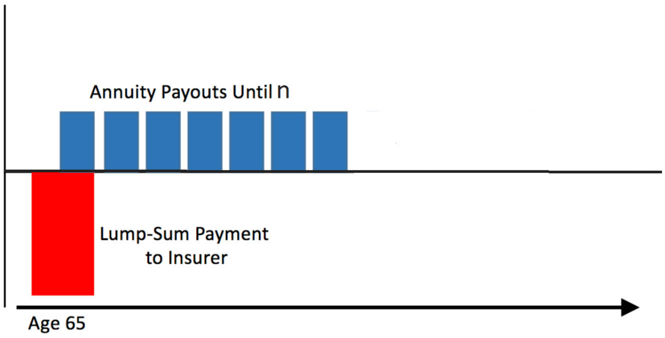
$$Y = \begin{cases} 0 & {}_mq_x \\ v^m & {}_{m|1}q_x \\ v^m + v^{m+1} & {}_{m+1|1}q_x \\ ... & ... \\ v^m + v^{m+1} + v^{m+2} + ... + v^{m+h} & {}_{m+h|1}q_x \\ ... & ... \end{cases}$$

If the pensioner dies before the deferral time he will receive nothing; if he dies at an age between $x+m$ and $x+m+1$ the present value of future benefits will be only v^m.
Again, if a fair annuity is assumed, the expected present value of future benefits is equal to the value of the single premium paid by the pensioner and results now to be:

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{h=0}^{+\infty}(v^m + v^{m+1} + v^{m+2} + ... + v^{m+h})\ {}_{m+h|1}q_x \\ &= {}_mE_x \cdot \sum_{h=0}^{+\infty} {}_hE_{x+m} \overset{def}{=} {}_mE_x \cdot \ddot{a}_{x+m} \overset{def}{=} {}_{m|}\ddot{a}_x\end{aligned}$$

${}_{m|}\ddot{a}_x$ is an actuarial notation denoting the price of a deferred lifetime annuity issued to a pensioner aged x and that starts payments after a deferral time of m years.

- Immediate Temporary Annuity for n years paid in advance: the regular payments provided by this type of policy start to be paid at time 0, when the pensioner decides to annuitize his savings, and continue to be paid at most for n years. If the insured dies before n years he is covered against longevity risk, whereas if he survives up to $x+n-1$ years he receives his last annuity income at time $n-1$ and then nothing. Here a graphical representation of the annuity.



Assuming once again annual and unitary payments, the present value of future benefits is a random variable Y with distribution:

$$Y = \begin{cases} 1 & q_x \\ 1+v & {}_{1|1}q_x \\ 1+v+v^2 & {}_{2|1}q_x \\ ... & ... \\ 1+v+v^2+...+v^{n-1} & {}_{n-1}p_x \end{cases}$$

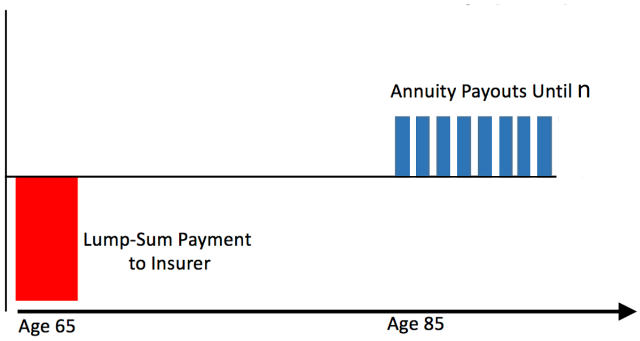
The payments of benefit continue until the death of the policy holder or until time $n-1$, whichever occurs first.
As before, in order to compute the value of the single premium paid at

time 0 by the pensioner, the expected present value of future benefits is derived:

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{h=0}^{n-1}(1+v+v^2+...+v^h)\ {}_{h|1}q_x \\ &= \sum_{h=0}^{n-1} v^h \cdot {}_hp_x \overset{def}{=} \sum_{h=0}^{n-1} {}_hE_x \overset{def}{=} \ddot{a}_{x:\overline{n}|}\end{aligned}$$

$\ddot{a}_{x:\overline{n}|}$ is another actuarial notation representing the expected present value of an immediate temporary annuity paid annually in advance. A pensioner aged x pays $\ddot{a}_{x:\overline{n}|}$ at time 0 in order to receive unit payments until death or up to $n-1$, depending on which event comes first.

- Deferred Temporary Annuity paid in advance: it is a contract very similar to the Immediate Temporary Annuity, but instead of providing benefits from time 0, the insurance company defers payments. The pensioner pays the single premium at time 0 but regular incomes start to be paid out at time $m > 0$ and continue to be paid at the beginning of each time interval until whichever occurs first between the death of the retiree and $m+n-1$. Here a graphic representation of the policy, if we assume an insured aged 65, deferral time of 20 years and n=8 years.



Assuming as always a fair annuity it is possible to compute the single premium as the expected present value of future benefits:

$$\mathbb{E}[Y] = \sum_{h=0}^{m+n-1} {}_hE_x = {}_mE_x \sum_{h=0}^{n-1} {}_hE_{m+x} = {}_mE_x \cdot \ddot{a}_{x+m:\overline{n}|} = {}_{m|}\ddot{a}_{x:\overline{n}|}$$

${}_{m|}\ddot{a}_{x:\overline{n}|}$ is the actuarial notation denoting the single premium paid in advance by an insured aged x, in order to receive, starting from time m, unit payments until death or up to $m+n-1$, whichever occurs first. It can be notice that ${}_{m|}\ddot{a}_{x:\overline{n}|}$ corresponds to an immediate temporary annuity started at time $x+m$, discounted using ${}_mE_x$.

All the products presented so far involve payments in advance; there are also products similar to those described above, but characterised by payments in arrears, i.e. at the end of the time interval.

Another assumption of the annuities presented so far is that the single premium paid by the policyholder is invested in a risk-free asset yielding an interest rate equal to i. However, there are more complicated products on the market that require careful financial management and elaborate actuarial calculations. For example, some policies may require a portion of the premium to be invested in risky assets, shares, financial options or to be managed in a dedicated way by life office managers.
Moreover, not all annuities are fair: in most cases the expected profit for the insurance company is not zero, i.e. the value of the single premium paid by the retiree (the net premium) does not correspond to the expected present value of future benefits. In order to have a solvency margin and to be sure of covering all expenses, the insurer must apply a safety loading to the net premium, increase the cost of the policy and charge the client at a gross premium (higher than the net one).

The decision for the retiree is therefore not easy: considering the large number of products available on the market, the retiree has to choose the most suitable for him or her, the one that best reflects his or her preferences. In addition, the pensioner is faced with another important choice: he has to determine the optimal time to annuitise. Indeed, depending on the degree of risk aversion of the insured, on market trends, on the life expectancy of the pensioner and on the fairness of the annuity, the optimal annuitization time varies widely among investors. And further, another decision to which the retiree is subject is: should I partially or totally annuitise? What portion of the fund should I allocate to an annuity?
Many papers tackled with this problems: some authors tried to determine the optimal time to annuitize, others determined the optimal annuity strategy considering different types of markets, others calculated the loss suffered by the retiree when forced to immediately annuitize at retirement, and yet others proposed new typology of annuity products.

Further, in many papers annuities were benchmarked against an other important product which is catching on around the world: the tontine. Indeed, the tontine is an innovative and attractive alternative pension decumulation strategy. I will analyze in detail its characteristics, history, and proposed literature versions in the following chapter.

Chapter 2

Modern Tontine

Tontines were special forms of investment, dating back to the 17th century, in which the investor paid a lump sum of money and received annual payments of 'dividends' until his death. The special feature of this product was that when an investor died, his shares were divided among the surviving members of the tontine. The tontine was thus seen as a group annuity in which the investor living longer would get larger annual payments. In fact, as members of the group died, they were not replaced by new investors, but the shares were divided among fewer and fewer members. The surviving investors literally profited from the death of their fellows: the shares for remaining members increased as more members died and the last investor alive collected the entire pot. When all the investors died, the tontine ended, and the government usually absorbed the remaining capital.
Here is an example: imagine a group of 1000 soon-to-be retirees who join together and pool \$1,000 each to purchase a \$1 million Treasury Bond (perpetual, whose nominal value will never be redeemed), which pays 3% coupons. The bond generate \$30,000 in interest yearly, which is split among the 1,000 participants in the pool, for a 30,000/1,000 = \$30 income per year per member. In the tontine scheme members agree that if and when they die, their guaranteed income of \$30 will be split among those who are still alive. If one decade later only 800 original investors are alive, the \$30,000 coupon is divided only among 800 members, providing a \$37.50 dividend each. Of this, \$30 is the guaranteed income and \$7.50 is the mortality credit, the other people's money. Then, if after two decades, only 100 individuals survive, the annual cash flow to survivor will be \$300 per annum. This procedure theoretically continues as long as there is still a survivor in the pool.

It is obvious that tontines were special products that could offer a solution to longevity risk: in this context, the longer the life span, the greater the benefit received. The risk of outliving resources is significantly reduced and consumption is more sustainable. However, the downside of those members with a bequest motive is that the assets of the dead are not passed to their estate, but are divided among the tontine group.

Today, a growing number of financial advisors, academics, and Fintech firms think that it might be time to take a second look at these financial arrangements. Many authors have examined whether a historical insurance concept such as the tontine has sufficient innovative potential to extend and improve on prevailing private pension solutions in a modern way. A recent stream of literature proposes and analyzes modern versions of tontines. Modern tontines allow people to get the benefits of longevity risk pooling without having to buy a life annuity. As already mentioned, the popularity of annuities has declined over the years because people are afraid of not realising a return on their investment before they die. Considering that modern tontines have lower costs and higher expected returns than annuities, they could be a safer and more cost-effective way of restoring pensions. Moreover, although payments are not as fixed as annuities, they will not decrease.
Modern versions of the tontine could be interesting and valuable ways for people to finance their final years.

There exists two different kinds of modern tontines versions:

- Explicit tontine: In the explicit tontine there is an explicit rule about how assets of dead members are shared. Each year individuals receive an explicit longevity credit into their individual account, usually called mortality credit. The Fair Tontine Annuity by Sabin(2010) [27], the Annuity Overlay Fund by Donnelly et al.(2014) [9] and the Pooled Annuity Fund by Stamos(2008) [30] are examples of explicit modern tontines.

- Implicit tontine: In the implicit tontine the individual receives an income, which is implicitly adjusted according to the mortality experience of the group. The Group Self-Annuitization of Piggott et al.(2005) [26] and the Optimal Retirement Tontine(2016) of Milevsky and Salisbury [20] are examples of implicit tontines.

In the following sections I will analyse in great detail the modern implicit and explicit versions of tontines proposed in the literature. Considering the great opportunity that tontines could offer to investors, as decumulation strategies, I considered it appropriate to devote a entire sections to the full explanation of such products. All of the modern tontines that I will analyze present some common characteristics and some peculiarities. For each of the proposed modern tontines I will examine benefits, payout rates and the wealth dynamic of individuals. In some cases I will also present comparisons with typical insurance products available in today's market, in order to better understand the attractiveness of these policies. But first, let's have a look at their history and origins.

2.1 History of Tontine

The word 'tontine' comes from the name of Lorenzo de Tonti. Lorenzo de Tonti was the governor of Gaeta and a Neapolitan banker who served as exile financial consultant to the French crown in the 1650s and promoted the scheme which now bears his name.
In the early 1650s the French treasury was battered by the Thirty Years' War and the rebellions within France known as the 'Fronde' and needed to raise money. King Louis XIV required money to finance the war and support his military. He used to collect money by raising taxes and issuing bonds, but bonds were too expensive and taxes could not be raised too much. Lorenzo de Tonti therefore proposed the original tontine to Jules Cardinal Mazarin as a means for the French King Louis XIV to raise revenue.[16]
According to Tonti's plan, which is the basis for all tontine schemes developed so far, the French government should issue shares, with 20 million principal, at the price of 300 livre per share. Investors would choose either their life, or the life of a third party (nominee) as the life in interest for their stake in the tontine. Participation would be structured in groups of equal sizes (classes) according to the age of the nominee, 0-7, 8-14 etc. all the way through the age of 63. Each beneficiary should receive an annual payment based on the interest earned from the combined initial capital of the investors in the applicable age cohort. The interest rate should increase with the age of the nominee: for example, an interest rate of 5% is assumed for the 8-14 age group and a rate of 15% for the over 63 age class. If a nominee died, his/her payments will get redistributed among the surviving investors in his class and then the payments of the other surviving nominees in that class would increase. The subscriber represented by the last nominee in each group would get all the interest generated by the capital within that band. On the death of the last investor, the capital would revert to the government.
Tonti did not just propose the financial aspect of the plan, he also made guidelines for administrative matters and pointed out the high benefits that shareholders can extract via purchasing a tontine.
The banker's idea was not without reason, but it was rejected by the French Parliament because it was considered too risky. There were no reliable calculations on life expectancy (these were not available until 1746) and, moreover, it was enough for a person to survive a long time to prolong the payment of the entire capital to the government by several decades. Tonti's original proposal of 1653 was not actually implemented.

Initially rejected by the French, the very first working tontine was established by the city of Kampen in Holland in 1670. [16] The Kampen Tontine paid a fixed coupon which every year was split among the surviving members of the tontine. Upon hearing about its success, many other local Dutch cities offered tontines to their citizens. Following suit, also France issued the first in 1689 in a series of 14 state tontines to raise money to fund military operations.

Also the English government organized a state tontine in 1693.
In January 1693, during the fifth year of the reign of King William III and Queen Mary II, the British Parliament passed the Million Act, designed to raise one million pounds towards carrying on the war against France.[18] The Act specified that prior to 1 May 1693 any British native or foreigner could purchase a tontine share from the Exchequer[1] for £100 and thus gain entry into the first British government-run tontine scheme.
For £100 a contributor (annuitant) could select any nominee of any age – including the contributor himself – on whose life the tontine would be contingent. Dividend payments would be distributed to the investor as long as the nominee was still alive. Notice that the entry investment in the tontine pool far exceeded the annual average industrial wage.[2]
The British tontine was a simpler structure compared to the original tontine scheme envisioned by Lorenzo de Tonti, which involved multiple classes - in order to reduce the transfer of wealth from older to younger participants. In King William's tontine, each share entitled the investor to an annual dividend of 10% for seven years (until June 1700), after which the dividends would be reduced to 7% per share. The declining structure of interest rates was introduced to allow dividends received by the survivors to be smoother.

A tontine was also one of the options proposed by the U.S. Treasury Secretary Alexander Hamilton as a means to reduce the national debt at the beginning of the American Republic in the year 1790. [16]
In order to reduce a crushing national debt, he suggested the U.S. government replace high-interest war debt with new bonds in which coupon payments would be made to a group, as opposed to individuals. The group members would share the interest payments equally among themselves, provided they were alive. But, once a member of the group died, his or her portion would stay in a pool and be shared among the survivors. This process would theoretically continue until the very last survivor would be entitled to the entire interest payment - potentially millions of dollars. Hamilton proposed the tontine in a letter to George Washington, claiming that it would reduce the interest paid on the US debt, and eventually eliminate it entirely. However the US Congress decided not to act on Hamilton's proposal, but the tontine idea itself never died on American soil.

Although tontines were imperfect substitutes for national taxes, they developed a secondary purpose in the UK and the US: project financing. [16] The tontine often took the form of a private subscription, whose profits were used to finance particular projects. These tontines were usually founded as an instrument for municipal improvement including both public and private buildings.

[1] The Exchequer was a component of the governments of England, Wales, Scotland and Northern Ireland which was responsible for the administration and collection of taxes.

[2] The average annual wage of building laborers in England during the latter part of the 17th century was £16. It is therefore quite plausible that the 1693 tontine was an investment mainly for rich people.

From the 18th century tontines were then used to collect private funding for construction projects on both sides of the Atlantic. In Britain some notable projects included hotels, the Freemasons' Lodge in the City of London and Richmond Bridge across the River Thames in west London. Notable private projects in the United States included hotels and public buildings, such as the 'Tontine Coffee House' in Manhattan, which is the first real location for what eventually became the New York Stock Exchange.

In the middle of the 19th century, US insurance companies also began issuing tontine insurance policies to the public and they became very popular. In 1876, seeking a means to stand out in the emerging and highly competitive insurance markets, Henry Hyde, the founder of Equitable Life, introduced the US public to tontine pensions. By the start of the 20th century, historians have documented that half of U.S. households owned a tontine insurance policy, which many used to support themselves through retirement. This was a personal hedge against longevity, with little risk exposure for the insurance company. Within 30 years, the tontine business had quickly outgrown the annuities business with over $5.77bn alone of Tontines sold by the four largest firms.[18] Also France and British insurance companies started issuing tontine as a form of social and retirement insurance. Just as their popularity grew rapidly in America, the fall of tontines was equally precipitous. Shortly after 1900, several spectacular misappropriation scandals[3] in the insurance industry led the influential New York State Insurance Commission to ban tontine insurances in the state, and by 1910 most other states followed suit. Tontines have been illegal in the United States for more than a century, and most insurance executives likely haven't heard of them.

Today, a growing number of financial advisers, academics and fintech companies think it might be time to take a second look at these financial arrangements. Tontines provide the regular income of an annuity - even more income for living members - and thanks to their relatively low costs, they produce higher returns than annuities.
They can also offer a solution to longevity risk and a modern version of them could be a viable way for people to fund their later years.

Today, tontines are regulated in Europe by Directive 2002/83/EC of the European Parliament. The pan-European pension regulation approved by the European Commission in 2019 also contains provisions that specifically allow new generation pension products that comply with the "tontine principle" to be offered in the 27 EU Member States.
In March 2017, the New York Times [24] reported that tontines were achieving new consideration as a way for people to obtain a steady retirement income

[3]Widespread fraud included manipulation of tontine members' registers, overcharging fees/costs to tontine pension funds, and conflicts of interest regarding investments made using members' capital.

and there are now only two States in the U.S. (Louisiana and South Carolina) that specifically prohibit tontines in their statutes.

So far, several new pension architectures have been designed that partially or fully use the tontine risk-sharing structure, such as:

- The Collective Defined Contribution (CDC) pensions offered in the UK after the Pension Freedom.
- Tontine Trust:[35] TontineTrust is a fintech company providing retirement benefits on a digital platform, based on the 'Tontine' principle where subscribers pool their longevity risk unlike annuities. The company is preparing to launch several international pension products for European consumers and to develop a safer, more cost-efficient and transparent marketplace.
- QSuper:[34] an Australian fund which offers lifetime annuities to its clients, including also tontine policies.
- Le Conservateure.[36] Le Conservateur Mutual Associations were created in France in 1844 with the ambition of developing and modernising the tontine system imagined by Lorenzo Tonti during the reign of Louis XIV. In 1976, Le Conservateur created a mutual insurance company that enabled it to extend its offer to life insurance contracts. Today Le Conservateure offers its customers the possibility of purchasing tontine policies.

Authors such as Piggott, Stamos, Milevsky, Donnelly and Sabin introduced modern versions of tontines in their works. I will analyse them in detail in the following sections.

2.2 Group Self Annuitization

The Group Self Annuitization (GSA) plan, introduced by Piggott et al.(2005) [26], is an implicit tontine.
The main feature of this type of product is that annuitants bear their systematic risk, but they share their idiosyncratic risk. A GSA plan allows retirees to join together and form a fund that can provide protection against longevity.
Individuals pay a single premium to the insurance company in order to receive annually, until death, a benefit. This benefit changes and is adjusted according to the mortality experience of the pool and actual investment returns. Hence it is an implicit tontine, in which the income is adjusted annually as mortality and returns evolve.
The authors presented the payment path for different scenarios - from the simplest to the most sophisticated and realistic.
They initially assumed a single cohort with identical mortality characteristics and a pool in which participants contribute equal amounts to the fund and in return, receive equal benefit payments. Then they considered the case where individuals may contribute varying amounts and may therefore receive different annuity payouts. They also extended their analysis to include multiple cohorts joining the pool at arbitrary times.
I will describe in the following subsections all the settings examined by Piggott et al.(2005) [26] with a focus on how payouts are evaluated in a GSA plan.

Single Cohort with Constant Contributions and Constant Annuity Payouts

A GSA plan initially operates like an ordinary lifetime annuity purchased in the private market. Depending on the premium paid by each retiree, the benefit payout is derived - taking into account expected mortality rates and expected investment returns. When the expectations are actually realized over time, the benefit payout determined at the point of entry remains constant. Conversely, when expectations are not fulfilled, the income should be adjusted taking into account actual data.

In this section I consider a pool consisting of l_x annuitants, all aged x, and with identical mortality characteristics. All of them decide to receive an annual benefit set equal to B_0. The initial total fund - equal to the sum of all premiums paid by the retirees - is therefore equal to:

$$F_0 = l_x \cdot B_0 \cdot \ddot{a}_x$$

where $\ddot{a}_x$ is an actuarial notation interpreted as the single premium paid by an x-aged for an immediate unitary lifetime annuity and its value is given by

$$\ddot{a}_x = \sum_{t=0}^{\infty} v^t \left(\frac{l_{x+t}}{l_x}\right)$$

Here it is assumed that the realised rates of return on investments are equal to the expected rates of return. However we consider the case where the actual survival pattern is different from expected, i.e. the number of individuals in the fund surviving is different from expected, and thus the initial payout B_0 must be adjusted to balance the fund.
The number of actual survivors will be denoted by a * so that, for instance, l^*_{x+1} represents the actual number of individuals in the fund surviving at age $x+1$.[4]
In order to determine the final formula for future payouts, it is first necessary to identify the evolution of the fund.
The fund value at time 1 becomes equal to the fund at time 0 minus benefits paid to survivors, all capitalised at the interest rate R:

$$\begin{aligned} F_1 &= (F_0 - l_x\, B_0)(1+R) \\ &= (l_x\, B_0\, \ddot{a}_x - l_x B_0)(1+R) \\ &= l_x\, B_0(\ddot{a}_x - 1)(1+R) \end{aligned}$$

Spreading this across the remaining survivors during their expected future lifetime, the periodic benefit payment received at time 1 becomes:

$$\begin{aligned} B^*_1 &= \frac{1}{l^*_{x+1}} \left(\frac{F_1}{\ddot{a}_{x+1}} \right) \\ &= \frac{1}{l^*_{x+1}} \left(\frac{l_x\, B_0(\ddot{a}_x - 1)(1+R)}{\ddot{a}_{x+1}} \right) \end{aligned}$$

Using the recursive relationship for annuitant factors

$$\ddot{a}_{x+1} = (\ddot{a}_x - 1)(1+R)/p_x \tag{2.1}$$

where p_x is the expected annual surviving rate equal to l_{x+1}/l_x, we have

$$\begin{aligned} B^*_1 &= \frac{1}{l^*_{x+1}} \left(\frac{l_x\, B_0(\ddot{a}_x - 1)(1+R)}{(\ddot{a}_x - 1)(1+R)/p_x} \right) \\ &= B_0 \left(\frac{l_x p_x}{l^*_{x+1}} \right) = B_0 \left(\frac{l_x p_x}{l_x p^*_x} \right) = B_0 \left(\frac{p_x}{p^*_x} \right) \end{aligned}$$

It is clear that the payout at time 1 is an adjusted version of the benefit at time 0. The adjustment factor is based on the ratio of expected to actual survival rates. Of course, if the realized survival rate p^*_x coincides with the expectation p_x, the benefit received at time 1 by the individual is exactly equal to B_0. Hence, when expectations are realized the payout is constant and equal to the benefit paid by a standard annuity. On the other hand, whenever $p^*_x \neq p_x$, the benefit must be adjusted. Indeed, if $p^*_x > p_x$, i.e. survival probability increases, the ratio $\frac{p_x}{p^*_x}$ is lower than 1, resulting in a decrease in benefits. The

[4] l_x is fixed and know since it represents the number of initial annuitants in the pool.

greater the number of survivors in the pool, the smaller the revenue received. Viceversa, if actual survival probability decreases, i.e. $p^*_x < p_x$, the ratio $\frac{p_x}{p^*_x}$ is higher than 1, resulting in an increase in payouts. The smaller the number of survivors in the pool, the higher the mortality credits received and thus the greater the income.
Proceeding intuitively we would determine the benefit payment at any given point in time t.

$$\begin{aligned} B^*_t &= \frac{1}{l^*_{x+t}} \left(\frac{F_t}{\ddot{a}_{x+t}} \right) = \frac{1}{l^*_{x+t}} \left(\frac{l^*_{x+t-1}\, B^*_{t-1} (\ddot{a}_{x+t-1} - 1)(1+R)}{\ddot{a}_{x+t}} \right) \\ &= \frac{1}{l^*_{x+t}} \left(\frac{l^*_{x+t-1}\, B^*_{t-1} (\ddot{a}_{x+t-1} - 1)(1+R)}{(\ddot{a}_{x+t-1} - 1)(1+R)/p_{x+t-1}} \right) \\ &= B^*_{t-1} \left(\frac{p_{x+t-1}}{p^*_{x+t-1}} \right) \end{aligned}$$

Again the benefit paid at time t is an adjusted version of the payout at time $t-1$, where the adjustment factor is based on the ratio between expected and actual survivorship rates. Of course, also in this case, if expected mortality rates are realized, the benefit paid at time t is constant and equal to the payout received at time $t-1$. Hence, if $p^*_{x+t-1} = p_{x+t-1}$, then $B^*_t = B^*_{t-1}$.
Notice that if the expectations are realized for any given point in time t, the payment path of the GSA plan is exactly the same as a standard annuity, since it provides a constant benefit equal to B_0 throughout time.

We now consider also the case when the investment earning pattern is different from the assumed constant rate R.
The actual investment earning rates will be denoted by a * so that, for instance, R^*_1 represents the actual interest rate earned in the first year, in $(0;1)$.
The value of the fund at any time t will be equal to the fund in the previous period minus benefits paid to survivors, all capitalised at the effective interest rate R^*_t.

$$F_t = (F_{t-1} - l^*_{x+t-1} B^*_{t-1})(1+R^*_t) = B^*_{t-1} l^*_{x+t-1} (\ddot{a}_{x+t-1} - 1)(1+R^*_t)$$

Spreading this across the remaining lives during their expected future lifetime, the periodic benefit payment received at time t becomes:[5]

$$\begin{aligned} B^*_t &= \frac{1}{l^*_{x+t}} \left(\frac{F_t}{\ddot{a}_{x+t}} \right) = \frac{1}{l^*_{x+t}} \left(\frac{l^*_{x+t-1}\, B^*_{t-1} (\ddot{a}_{x+t-1} - 1)(1+R^*_t)}{\ddot{a}_{x+t}} \right) \\ &= B^*_{t-1} \frac{l^*_{x+t-1}}{l^*_{x+t}} \frac{(\ddot{a}_{x+t-1} - 1)(1+R^*_t)}{(\ddot{a}_{x+t-1} - 1)(1+R)/p_{x+t-1}} \\ &= B^*_{t-1} \left(\frac{p_{x+t-1}}{p^*_{x+t-1}} \cdot \frac{1+R^*_t}{1+R} \right) \end{aligned} \tag{2.2}$$

[5]Here we assume again that (2.1) holds. The stream of annuity factors $\ddot{a}_{x+t}$ is calculated taking into account expected survival and interest rates, not actual ones.

Hence, we observe that the payment for period t depends on the payment for $t-1$ and two adjustment factors: the first one is related to the difference in expected and realized mortality during the previous period and the second factor is related to the difference in the expected and realized investment earnings rate for the period. Of course if $R = R_t^*$ and $p_{x+t-1}^* = p_{x+t-1}$, the benefit received at time t is the same as the one received at time $t-1$.
The essential feature in the calculations demonstrated above, is that the periodic, here assumed annual, benefit payout rates can be determined from the previous benefit payout rates multiplied by two adjustment factors. The generic formula is given by

$$B_t = B_{t-1} \cdot MEA_t \cdot IRA_t \tag{2.3}$$

where MEA_t is the mortality experience adjustment and IRA_t is the interest rate adjustment for the period from year $t-1$ to t.
This is how a GSA plan operates: it recomputes the benefit payouts periodically using the most recent benefit payouts and multiplying them by adjustment factors. If, for example, the mortality is lighter than expected, i.e. $p_{x+t-1}^* > p_{x+t-1}$, the next period benefit will decrease, i.e. $B_t < B_{t-1}$. The fund would have to be distributed over a larger group of survivors than expected and so the benefit payment would be lower. Similarly, if the investment earnings of the period are worse than expected, i.e. $R_t^* < R$, there will be lower benefit payouts, i.e. $B_t < B_{t-1}$, since the accumulated fund does not grow as predicted. Of course, in this scheme, if expectations are always realized, then the benefit payout is constant throughout time and equal to the one of a standard lifetime annuity, i.e. $B_t = B_0 \; \forall i$.

Another interesting way to explain the benefits of a GSA plan is to recognize mortality credits. In fact, GSA payments, in addition to being adjusted versions of previously received payouts, can also be viewed as a portion of the personal fund plus additional benefits received by dead people.
We define by $F_{i,t}$ the fund owned by each individual i at time t. Notice that:

$$F_t = \sum_{\forall i} F_{i,t} \text{ where } F_{i,t} = F(t) \; \forall i \text{ each member contribute equal amount}$$

We define the following sets:

$$A_t = \{i \in \{1, ..., l_x\} : \text{i-th individual is alive at time } t\}$$

$$D_{t-1} = \{i \in \{1, ..., l_x\} : \text{i-th individual dies between } t-1 \text{ and } t\}$$

Now, the total fund of the plan at time t can be decomposed into the funds of those who are still alive at time t, denoted by the set A_t, and the funds of those who died between $t-1$ and t, denoted by the set D_{t-1}.

$$F_t = \sum_{A_t} F_{i,t} + \sum_{D_{t-1}} F_{i,t}$$

The benefit of GSA plan can thus be viewed as:

$$
\begin{aligned}
B_t &= \frac{F_t}{\ddot{a}_{x+t}} \cdot \frac{1}{l^*_{x+t}} = \frac{\sum_{A_t} F_{i,t} + \sum_{D_{t-1}} F_{i,t}}{\ddot{a}_{x+t} l^*_{x+t}} \\
&= \frac{l^*_{x+t} F(t)}{l^*_{x+t} \ddot{a}_{x+t}} + \frac{\left(l^*_{x+t-1} - l^*_{x+t}\right) F(t)}{l^*_{x+t} \ddot{a}_{x+t}} \\
&= \frac{F(t)}{\ddot{a}_{x+t}} + \frac{\left(l^*_{x+t-1} - l^*_{x+t}\right) F(t)}{l^*_{x+t} \ddot{a}_{x+t}} \qquad (2.4)
\end{aligned}
$$

It is clear that the value of the benefit received at time t is equal to a portion of the personal fund $F(t)$ owned at time t plus a mortality credit. The mortality credit is derived from an equitable redistribution of available funds from those who died during the period.
$\left(l^*_{x+t-1} - l^*_{x+t}\right)$ represents the number of people died between $t-1$ and t: the funds owned by those people are redistributed among those who are still alive at time t, i.e. l^*_{x+t}. This is precisely what is meant by mortality credits: reallocation of funds among survivors.
Through this representation we are able to recognize the tontine mechanism within the GSA plan: funds from dead people are redistributed equally among those who are still alive, in the form of mortality credits.
We can check that if we assume $F(t) = \frac{F_t}{l^*_{x+t-1}}$, this 'new' representation of benefits coincides with the original one.

$$
\begin{aligned}
B^*_t &= \frac{F(t)}{\ddot{a}_{x+t}} + \frac{l^*_{x+t-1} F(t)}{l^*_{x+t} \ddot{a}_{x+t}} - \frac{l^*_{x+t} F(t)}{\ddot{a}_{x+t} l^*_{x+t}} \\
&= \frac{l^*_{x+t-1}\, F(t)}{l^*_{x+t}\, \ddot{a}_{x+t}} \\
&= \frac{l^*_{x+t-1}\, F_t}{l^*_{x+t-1}\, l^*_{x+t}\, \ddot{a}_{x+t}}
\end{aligned}
$$

which corresponds exactly to

$$
\frac{F_t}{\ddot{a}_{x+t}} \cdot \frac{1}{l^*_{x+t}}
$$

The idea is that the total fund value at each time t is equally split among those who are still alive, providing them with an individual fund of $F(t)$.

To summarize, the benefit payout of a GSA plan, when we assume homogeneous investors and constant contribution, can be seen as:

- Adjusted version of previous benefit payout as stated in (2.2);
- Portion of individual fund plus mortality credit as stated in (2.4).

In the following section I will present more complicated and realistic situations in order to extend the formulas presented so far. We will see that, even in the case of variable contributions, the insights presented up to now are valid.

Single Cohort with Varying Contributions and Varying Annuity Payouts

The previous section developed calculation of the benefit payout rates assuming that participants contribute equal amounts to the fund and in return, receive equal annual benefit payouts. We consider now the case where varying amounts of contributions and different annuity incomes are allowed for all participants. Here it is assumed that at time 0 there are l_x individuals, all aged x, joining the group. Each i^{th} annuitant pays the single premium $F_{i,0}$ at time 0, which is contributed to the total fund. The total fund at the beginning of the period would be therefore equal to the sum of all the single contributions made by all retirees.

$$F_0 = \sum_{i=1}^{l_x} F_{i,0}$$

The total benefit payment for the entire group is then given by

$$B_0 = \frac{F_0}{\ddot{a}_x} = \sum_{i=1}^{l_x} B_{i,0}$$

where $B_{i,0}$ represents the initial annual benefit payment for the i^{th} annuitant and is given by:

$$B_{i,0} = \frac{F_{i,0}}{\ddot{a}_x} = \frac{F_0}{\ddot{a}_x}\frac{F_{i,0}}{F_0} = B_0 \left(\frac{F_{i,0}}{F_0}\right)$$

It is clear that each income $B_{i,0}$ is equal to a portion of the total benefit payout B_0: the proportion of benefit due to the i^{th} individual is closely related to the percentage of fund held by the i^{th} annuitant. The higher is the percentage of total fund owned by the i^{th} annuitant, the higher would be the benefit received.

After one period, at time $t = 1$, the entire group's fund value becomes equal to the initial total fund minus the total benefit payment maid to the l_x individuals, all capitalised at the actual interest rate R_1^*.

$$F_1 = (F_0 - B_0)(1 + R_1^*)$$

The next annuity payout for the entire group is therefore given by:

$$B_1^* = \frac{F_1}{\ddot{a}_{x+1}}$$

Let us denote by $F_{i,1}$ the fund owned by the i^{th} individual at time 1 and let us define the following sets:

$$A_1 = \{i \in \{1, ..., l_x\} : \text{i-th individual is alive at time 1}\}$$

$$D_0 = \{i \in \{1, ..., l_x\} : \text{i-th individual dies between 0 and 1}\}$$

Knowing that the single annuity payout for an individual who is still alive at the end of the period is determined as a portion of the total annuity payout

B_1^*, taking into account the proportion of the fund owned by the annuitant, we can compute:

$$B_{i,1}^* = B_1^* \left(\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \right)$$

We can also check that the sum of all the single payouts received by the entire pool, in particular by those who are still alive at time 1, is exactly equal to the total payout of the fund. Indeed:

$$\sum_{i \in A_1} B_{i,1}^* = \sum_{i \in A_1} B_1^* \left(\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \right) = B_1^* \left(\sum_{i \in A_1} \frac{F_{i,1}}{\sum\limits_{i \in A_1} F_{i,1}} \right) = B_1^*$$

According to the insights developed in the previous section, we expect, even in the case of variable contributions, that the benefit received at time 1, is equal to a portion of the individual fund at time 1 plus mortality credits derived from people who died between 0 and 1.
Recall that the total fund value at time 1, can be decomposed into the sum of all the individual retirees' funds, both for those who are still alive , represented by the cohort A_1, and for those who died between $[0, 1)$, denoted by D_0.

$$F_1 = \sum_{A_1} F_{i,1} + \sum_{D_0} F_{i,1}$$

As a matter of fact, we can derive the value of the benefit received at time 1 by the i^{th} annuitant as:

$$B_{i,1}^* = B_1^* \left(\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \right) = \frac{F_1}{\ddot{a}_{x+1}} \left(\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \right) = \frac{\sum\limits_{A_1} F_{i,1} + \sum\limits_{D_0} F_{i,t}}{\ddot{a}_{x+1}} \left(\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \right)$$

$$= \frac{F_{i,1} + \left(\sum\limits_{D_0} F_{i,1} / \sum\limits_{A_1} F_{i,1} \right) F_{i,1}}{\ddot{a}_{x+1}} = \frac{F_{i,1}}{\ddot{a}_{x+1}} + \frac{\frac{F_{i,1}}{\sum\limits_{A_1} F_{i,1}} \sum\limits_{D_0} F_{i,1}}{\ddot{a}_{x+1}}$$

The total benefit received at time 1 by each individual can thus be decomposed into two parts. The first one is derived from the personal fund owned by each participant at time 1. The second one is an additional benefit amount made to the annuitant, derived from a redistribution of the funds available from those who died between 0 and 1. One can think of this as a form of inheritance derived from those who died in the group. We can now extend the formulas inductively to time t, such that the annuity payout rate for an individual who survive at time t can be determined as follow:

$$B_{i,t}^* = \frac{F_{i,t}}{\ddot{a}_{x+t}} + \frac{\frac{F_{i,t}}{\sum\limits_{A_t} F_{i,t}} \sum\limits_{D_{t-1}} F_{i,t}}{\ddot{a}_{x+t}} \tag{2.5}$$

$\sum\limits_{D_{t-1}} F_{i,t}$ represents funds of dead people that must be reallocated among those who are still alive. Also in this case the distribution of mortality credits is

proportional to the percentage of fund owned by each individual. The higher is the percentage of fund owned, the higher is the amount of mortality credit received. We can also see that when a person with a very large fund dies, all other participants in the pool will benefit a lot, since a great amount of money is redistributed. On the other hand, if some poor member dies, someone with a very small fund, all the other participants in the pool will not increase a lot their benefit payouts.
Further, in order to get a better idea of the dynamics of individual funds $F_{i,t}$, let us define by $\hat{F}_{i,t}$ as the numerator of (2.5). We can recognize that

$$\hat{F}_{i,t} = \frac{B^*_{i,t}}{a_{x+t}}$$

The evolution of individual funds depends on $\hat{F}_{i,t}$. Indeed:

$$F_{i,t} = (\hat{F}_{i,t-1} - B^*_{i,t-1})(1 + R^*_t)$$

Each single fund at time t is equal to the fund $\hat{F}_{i,t-1}$ in the previous period, minus benefit paid to the retiree, all capitalized at the actual interest rate R. One interesting feature is that each single fund evolves by taking into account both benefits paid and mortality credits received. Therefore, on one hand, the benefits paid are subtracted, but on the other hand mortality credits are added.
Having in mind how individual funds evolve over time, let us explain the benefits of the GSA plan as an adjusted version of previous payments received, even in the case of varying contributions amount. We attempt to derive, also in this case, benefit payments at time t by recognizing two adjustment factors. As a matter of fact:

$$\begin{aligned} B^*_{i,t} &= \frac{F_t}{\ddot{a}_{x+t}}\left(\frac{F_{i,t}}{\sum\limits_{A_t} F_{i,t}}\right) = \frac{F_t}{\sum\limits_{A_t} F_{i,t}} \frac{(\hat{F}_{i,t-1} - B_{i,t-1})(1+R^*_t)}{(\ddot{a}_{x+t-1} - 1)(1+R)/p_{x+t-1}} \\ &= \frac{F_t}{\sum\limits_{A_t} F_{i,t}} \frac{B^*_{i,t-1}(\ddot{a}_{x+t-1} - 1)(1+R^*_1)}{(\ddot{a}_{x+t-1} - 1)(1+R)/p_{x+t-1}} \\ &= B^*_{i,t-1} \cdot \frac{p_{x+t-1}}{\left(\sum\limits_{A_t} F_{i,t}\right)/F_t} \cdot \left(\frac{(1+R^*_t)}{(1+R)}\right) \end{aligned} \tag{2.6}$$

Again, the next year's payout rate is calculated by adjusting the previous year's benefit payout rate by two factors: one due to mortality experience and another due to interest rates. Of course, if $R^*_t > R$, the benefit received increases, $B^*_{i,t} > B_{i,t-1}$: the higher investments returns, the higher payout rates. On the other hand, the higher the number of people surviving in the pool, i.e. $\sum\limits_{i \in A_t} F_{i,t}$ is high, the smaller benefit received.

We can notice that the mortality adjustment derived in the previous section, when all participant contribute equal amounts, is a special case of this. Let us examine the issue:

$$MEA_t = \frac{p_{x+t-1}}{\left(\sum\limits_{A_t} F_{i,t}/F_t\right)} \overset{[if\ F_{i,t}=F_i\ \forall i]}{=} \frac{p_{x+t-1}}{\left(\frac{F_i l^*_{x+t}}{F_t}\right)} = \frac{p_{x+t-1}}{\left(\frac{F_t l^*_{x+t}}{l^*_{x+t-1} F_t}\right)} = \frac{p_{x+t-1}}{p^*_{x+t-1}}$$

The benefit payout fits the formula (2.3) stated in the previous section:

$$B^*_{j,t} = B^*_{j,t-1} \cdot MEA_t \cdot IRA_t \tag{2.7}$$

The benefit payout is, once again, a twice-adjusted version of the previous year's income. The first adjustment is the mortality experience adjustment given by the ratio of expected survival rate to $(\sum\limits_{A_t} F_{i,t}/F_t)$. The second adjustment is instead the interest rate adjustment calculated as the ratio of realized to expected investment earnings rate for the period.

Multiple Cohorts

The most effective way to exploit the pooling effect in a GSA plan is to introduce multiple cohorts. Such a scheme allows people to join the fund, regardless of their age and irrespective of the point in time. It is a plan able to integrate new entrants with existing members of the pool. The authors listed four criteria that must be met within a GSA plan.

- If all groups experience the expected mortality, payouts should not alter for any group;
- If groups' expected and actual mortality differ, payments should all vary in the same proportion;
- Departures of realized from expected mortality should result in a once-for-all adjustment in all future payments;
- Period by period balance should be preserved.

When we introduce multiple cohorts there are many temporal and age dimensions involved: let's have a look at the notation.
First there is the age at which an individual enters the pool, denoted by $[x]$.
Second we have the current period, indicated with time t.
Lastly we have the length of time that has elapsed since joining the plan, denoted by k. When $k = t$ it means that we are dealing with cohorts who joined the pool at plan inception. When $k \neq t$ it means that we are dealing with cohorts who joined the pool after the plan inception and k represents the time spent in the fund by individuals.

The fund value at time t, of the i^{th} annuitant, belonging to the cohort who entered at age $[x]$, k periods ago is denoted by ${}^{k}_{[x]}F_{i,t}$ and computed as:

$$ {}^{k}_{[x]}F_{i,t} = ({}^{k-1}_{[x]}\hat{F}_{i,t-1} - {}^{k-1}_{[x]}B^*_{i,t-1})(1+R^*_t) $$

As before, the fund realised at time t corresponds to the fund $\hat{F}$ at time $t-1$ minus the benefits paid, capitalised at the current interest rate.
The total fund of the plan F_t is then given by the sum of all single funds, aggregated over all ages, over all durations of funds and over all individuals who are still alive at time t.[6]

$$ F_t = \sum_k \sum_x \sum_{A_t} {}^{k}_{[x]}\hat{F}_{i,t} $$

The payout rate at time t, for the i^{th} annuitant, belonging to the cohort who entered at age $[x]$, k periods ago is denoted by ${}^{k}_{[x]}B^*_{i,t}$ and computed as:

$$ {}^{k}_{[x]}B^*_{i,t} = \frac{{}^{k}_{[x]}\hat{F}_{i,t}}{\ddot{a}_{[x]+k}} $$

By combining the formulas stated above we can rewrite the total fund value at time t as:

$$ F_t = \sum_{k\geq 1} \sum_x \sum_{A_t} {}^{k}_{[x]}B^*_{i,t}\ddot{a}_{[x]+k} $$

This equation balances the fund and the actuarial present value of future payments. Following the logic of the GSA plan, with ${}^{k}_{[x]}B^*_{i,t} = {}^{k}_{[x]}B^*_{i,t-1} \cdot MEA_t \cdot IRA_t$, we get:

$$ \begin{aligned} F_t &= MEA_t \cdot \left(\frac{1+R^*_t}{1+R}\right) \sum_{k\geq 1} \sum_x \sum_{A_t} {}^{k}_{[x]}B^*_{i,t-1}(\ddot{a}_{[x]+k-1} - 1)\frac{(1+R)}{p_{[x]+k-1}} \\ &= MEA_t \cdot \sum_{k\geq 1} \sum_x \sum_{A_t} ({}^{k}_{[x]}B^*_{i,t-1}\ddot{a}_{[x]+k-1} - {}^{k}_{[x]}B^*_{i,t-1})\frac{1+R^*_t}{p_{[x]+k-1}} \end{aligned} $$

Notice that

$$ ({}^{k}_{[x]}B^*_{i,t-1}\ddot{a}_{[x]+k-1} - {}^{k}_{[x]}B^*_{i,t-1})(1+R^*_t) = ({}^{k}_{[x]}\hat{F}_{i,t-1} - {}^{k}_{[x]}B^*_{i,t-1})(1+R^*_t) = {}^{k}_{[x]}F_{i,t} $$

Thus the total fund value becomes:

$$ F_t = MEA_t \cdot \sum_{k\geq 1} \sum_x \sum_{A_t} \frac{{}^{k}_{[x]}F_{i,t}}{p_{[x]+k-1}} $$

The mortality adjustment factor results thus to be equal to:

$$ MEA_t = \frac{F_t}{\sum\limits_{k\geq 1} \sum\limits_x \sum\limits_{A_t} \frac{{}^{k}_{[x]}F_{i,t}}{p_{[x]+k-1}}} \tag{2.8} $$

[6]The summation is done considering only those who entered the pool before time t, not for those who enter at t. We consider only those exposed to risk in the previous period, but not new entrants.

It can be proved, after a little bit of calculation and transformation, that the mortality adjustment factor under a multiple cohort scheme, is a weighted harmonic mean of the mortality adjustment factor under a single cohort scheme. We can also recognize that the mortality adjustment factor of the single cohort case is a special case of this adjustment. When we are dealing with homogeneous cohorts there is only one age x and only one time of entry $t-k$. Equation (2.2) is therefore simplified and turns out to be in line with equation (2.6).

Once the mortality adjustment factor has been calculated, once and for all future payments, as required by the four initial criteria, we can state the general formula for calculating the payment rate at time t for the i^{th} annuitant, belonging to the cohort who entered at age $[x]$, k periods ago. Again, the benefit payout fits the formula (2.3) and (2.7) stated in the previous sections:

$$ {}^{k}_{[x]}B^{*}_{i,t} = {}^{k}_{[x]}B^{*}_{i,t-1} \cdot \left(\frac{F_t}{\sum\limits_{k\geq 1} \sum\limits_{x} \sum\limits_{A_t} \frac{{}^{k}_{[x]}F_{i,t}}{p_{[x]+k-1}}} \right) \cdot \left(\frac{1+R^{*}_{t}}{1+R} \right) \tag{2.9} $$

The periodic benefit payment in a group self annuitization plan is always determined based on the previous payment, adjusted for any deviations in mortality and interest rates from expectations. Once again, if the earning rate of return is higher than expected, i.e. $R^{*}_{t} > R$, the payout rate increases.
Such a tontine is able to determine a payout level a priori, on the basis of mortality and earnings expectations. Then, in the event that expectations are not met, the GSA plan adjusts benefits. This income adjustments are made periodically whenever deviations from expectations occur.

Examples

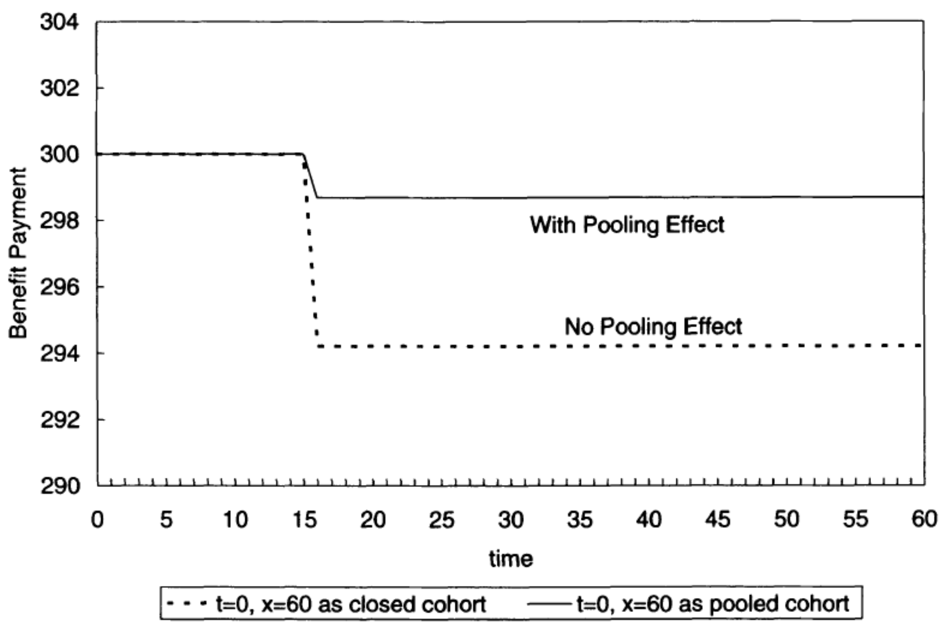
In this section we will show the difference between a GSA plan with a single cohort and a GSA plan with multiple cohorts. We consider an example in which expected mortality rates are taken from the United States RP-2000 Male Healthy Annuitant, and expected interest rates are fixed to 4%. We assume a 50% deviation in mortality rates occurred in only one period, at time $t = 15$, and no deviations in earning rates.

For the single cohort case, we consider an individual belonging to the age-60 cohort joining the plan at inception, and an initial benefit payout established at \$300 for all per period, i.e. we are dealing with an homogeneous cohort in which all participants contribute equal amounts. A deviation in mortality rate at period 15 requires a benefit adjustment from the initial value at time 16. In particular, if we assume a 50% drop in mortality, the benefit payment is reduced by about 2 percentage points to \$294. As expected, whenever the actual survival rates increase, thus mortality rates drop, the level of the payouts decreases. The idea is that when the actual number of participant in the GSA is higher than expected, the benefits paid to survivors must be adjusted so

that they decrease. The fund is not able to guarantee high benefits to a larger pool of people than expected.
In the case of multiple cohorts, other cohorts of different ages are allowed to enter the pool at a later date. Deviation in mortality, for the purpose of simplifying the illustration, is assumed only for the 60-age cohort: new entrants do not encounter any deviation from expected mortality. Also in this case, a deviation in mortality rate at period 15 requires a benefit adjustment from the initial value at time 16. Now, if we assume again a 50% drop in mortality, the benefit payment is reduced only by about 0.5 percentage points to $298,7.

In figure 2.1 is shown the effect of pooling longevity over several age cohorts when a deviation in mortality occurs in a single period. When we pool all cohorts together, the effect of drop in mortality in benefit payouts is less dramatic. Whenever the fund is open to multiple cohorts, allowing individuals to share idiosyncratic risks, any variation in benefits may be reduced. By sharing the risk, people are able to reduce the adjustments to their payout resulting from mortality and interest rate experience.

Figure 2.1: Benefit payments GSA plan



Source: Piggott et. al (2005)[26]

I will come back to the group self annuitization plan in chapter 3, dedicated to simulations. Thus, I will simulate the evolution of benefits, mortality credits and mortality adjustments and I will analyze the dynamic of funds. For the sake of simplicity, I will perform simulation only for the homogeneous cohort case, but I will consider both the constant contribution and varying contribution case.

2.3 Optimal Retirement Tontine

Another implicit tontine structure has been proposed by Milevsky and Salisbury (2016). [20]
By examining the original tontines proposed in history, the authors decided to suggest a different type of scheme in which payment rates were not constant over time. As seen in the case of historical tontines, if a constant payment rate is assumed, the dividends received by the survivors turn out to be increasing over time: as the number of survivors decreases more than exponentially over time. Such a payout path, where the last survivor receives hundreds of multiples of his or her initial investment, is at odds with the economic desires of pensioners. Indeed, many studies have shown that older retirees prefer stable or decreasing consumption across years, rather than increasing cash flows as proposed by historical tontines. For this reason, Milevsky and Salisbury proposed a different tontine, designed to better meet the needs of pensioners. By maximizing the expected utility of consumption over a lifetime, they were able to find the optimal tontine structure, in other words, to calculate the optimal payment rate $d(t)$ that is expected not to remain constant over time. As in most literature articles, the authors first examined a simpler case and then moved into a more complicated scenario. They initially considered the case of a single homogeneous cohort, in which all pool participants have the same age and mortality characteristics. In this setting they defined the natural tontine as the function for which the payout declines in exact proportion to the survival probabilities, proving that this is optimal under logarithmic utility function. Thereafter, they extended their analysis to multiple inhomogeneous cohorts. Here they were looking for optimal participation rates, optimal pricing shares and optimal payout paths that would make the scheme equitable. They defined the proportional tontine and provided conditions for the existence of such a product.
In the following sections I will review in great details all the arguments presented in this study.

Single Homogeneous Cohort

In this section I will analyze the tontine scheme proposed by Milevsky and Salisbury when a single homogeneous cohort is assumed.
The proposed tontine is expected to cost a dollar and continuously pay a payout function of $d(t)$, instead of annually or monthly. The authors also assumed a constant[7] risk free interest rate r and an objective survival function ${}_tp_x$ which applies to all individuals aged x.
Given that the basic comparator for a tontine is an annuity in which policyholders initially pay one dollar to the insurer and receive in return a lifetime income stream of $c(t)$, I will now examine the expected utility of a rational

[7] All contributed funds are invested at time 0 in a static bond portfolio, whose interest rate is equal to r.

annuitant in order to compare it with the plan proposed by Milevsky and Salisbury. Letting $u(c)$ denote the instantaneous utility of consumption, a rational annuitant, having no bequest motive, will choose a life annuity payout function $c(t)$ that maximizes the expected discounted lifetime utility function:

$$\max_{c(t)} \mathbb{E}\left[\int_0^{\infty} e^{-rt} u(c(t))dt\right] = \max_{c(t)} \left[\int_0^{\infty} e^{-rt} \, {}_tp_x u(c(t))dt\right]$$

This maximization problem is moreover subject to a constraint. The constraint on these annuities is that they are fairly priced, meaning that the initial payments invested at the risk free rate are able to fund the benefit payouts in perpetuity:

$$\int_0^{\infty} e^{-rt} \, {}_tp_x c(t)dt = 1$$

By the Euler-Lagrangian theorem, it is possible to solve the constrained maximization problem, and come up with the following result:

$$\exists \lambda \ \mid \ e^{-rt} \, {}_tp_x u'(c(t)) = \lambda e^{-rt} \, {}_tp_x \quad \forall t$$

Simplifying the equation on both sides we get:

$$u'(c(t)) = \lambda$$

Provided that the utility function is strictly concave, the optimal annuity payout function $c(t)$ must be constant. The stable level of income can be easily determined through the budget constraint as:

$$c(t) = c_0 = \left[\int_0^{\infty} e^{-rt} \, {}_tp_x dt\right]^{-1} \tag{2.10}$$

Let us now examine the tontine structure proposed by the authors. As mentioned above the main purpose of the authors is to determine the optimal level of payout $d(t)$ that can maximize the expected utility. The main point of the paper is that there is no reason for the tontine payout function to be a fixed percentage of the initial dollar invested, as it was historically.
Suppose there are initially n subscribers to the tontine plan at time 0, each depositing a dollar with the tontine sponsor. Let $N(t)$ be the random number of live subscribers at time t. If we consider one of these subscribers, and assume that this individual is alive, we can define $N(t) - 1$ as the number of other live subscribers. The variable $N(t) - 1$ is assumed to be binomially distributed with probability of success equal to the survival rate ${}_tp_x$.
As with the lifetime annuity, here again the individual will choose the tontine optimal payout function $d(t)$ that maximizes the individual's discounted utility:

$$\max_{d(t)} \mathbb{E}\left[\int_0^{\infty} e^{-rt} u\left(\frac{nd(t)}{N(t)}\right) dt\right]$$

Let us note that $nd(t)$ represents the total fund payment at time t, which must be divided among those who are still alive, i.e. $N(t)$. As the expected value function is solved, the maximization problem becomes as follows:

$$\max_{d(t)} \int_0^\infty e^{-rt} \,{}_tp_x \mathbb{E}\left[u\left(\frac{nd(t)}{N(t)}\right)dt\right] =$$

$$\max_{d(t)} \int_0^\infty e^{-rt} \,{}_tp_x \sum_{k=0}^{n-1} \binom{n-1}{k} \,{}_tp_x^k (1-\,{}_tp_x)^{n-1-k} u\left(\frac{nd(t)}{k+1}\right) dt$$

Also this maximization problem is subject to a constraint. The constraint on the tontine payout function is that the sponsor of the tontine can not sustain a loss, meaning that the initial deposit of n should be sufficient to sustain withdrawals in perpetuity. The sponsor is not subject to longevity risk; it is the pool that bears the risk entirely. The budget constraint is therefore:

$$\int_0^\infty e^{-rt} d(t) dt = 1$$

By the Euler-Lagrangian theorem, it is possible to solve the constrained maximization problem, and come up with the following result:

$$\exists\lambda \;\mid\; e^{-rt} \,{}_tp_x \sum_{k=0}^{n-1} \binom{n-1}{k} \,{}_tp_x^k (1-\,{}_tp_x)^{n-1-k} \frac{n}{k+1} u'\left(\frac{nd(t)}{k+1}\right) = \lambda e^{-rt} \quad \forall t$$

It is possible but challenging to solve this equation when a generic utility function is considered. But if a Constant Relative Risk Aversion (CRRA) utility is assumed the solution is greatly simplified. Indeed with a CRRA utility function, the optimal tontine withdrawal rate is a function of the longevity risk aversion coefficient γ, of the number of initial subscribers n and of the survival probability ${}_tp_x$:

$$d(t) = D_{n,\gamma}(\,{}_tp_x) = D_{n,\gamma}(1)\beta_{n,\gamma}(\,{}_tp_x)^{1/\gamma} \tag{2.11}$$

where:

$$\beta_{n,\gamma}(\,{}_tp_x) = \,{}_tp_x \sum_{k=0}^{n-1} \binom{n-1}{k} \,{}_tp_x^k (1-\,{}_tp_x)^{n-1-k} \left(\frac{n}{k+1}\right)^{1-\gamma}$$

$$D_{n,\gamma}(1) = \left[\int_0^\infty e^{-rt} \beta_{n,\gamma}(\,{}_tp_x)^{1/\gamma} dt\right]^{-1}$$

This optimal payout function results now to be decreasing with t. Given that the actual periodic amount received by each individual is given by $nd(t)/N(t)$, and considering that the number of survivors $N(t)$ decreases over time, the optimal tontine pays survivors a cash value that should remain relatively constant over the retirement years.

We consider now a special case: we fix the longevity risk aversion coefficient $\gamma = 1$, assuming then a logarithmic utility function.

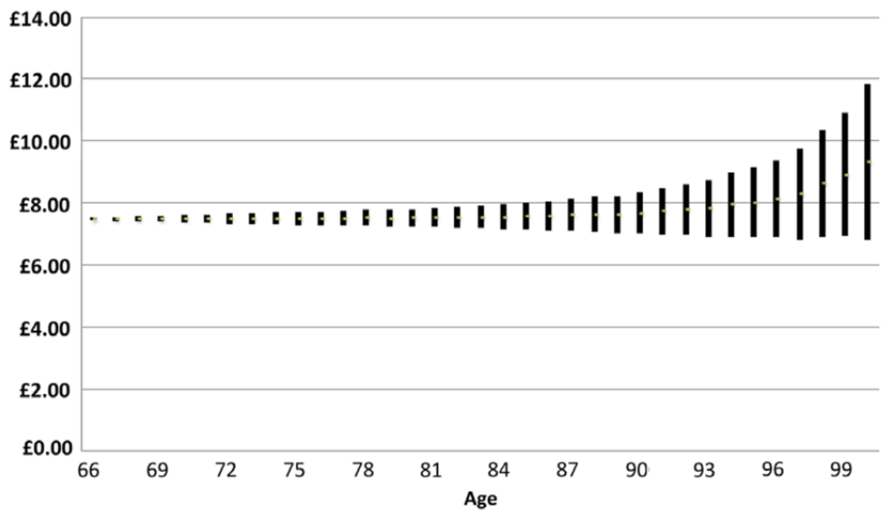
When this utility function is taken into account the optimal withdrawal function is greatly simplified. Under this assumption indeed $d(t)$ is simply a rate that decreases in exact proportion to the survival probabilities. The authors called this special type of tontine *Natural Tontine* and defined its payout with $D_N({}_tp_x)$.

$$D_N({}_tp_x) = {}_tp_x \cdot c_0 = {}_tp_x \cdot \left[\int_0^\infty e^{-rt}\, {}_tp_x dt\right]^{-1} \tag{2.12}$$

Milevsky and Salisbury also proved an important result: the *Natural Tontine* is quite optimal for all given risk aversion coefficients γ when n is sufficiently large. They showed that when an individual with $\gamma \neq 1$ is faced with a tontine structure which is only optimal for someone with $\gamma = 1$, the welfare loss they experience is minuscule. For this reason the authors advocate the *Natural Tontine* as the basis for the 21^{st}-century tontines: rather than expecting insurance companies to offer a range of products for all the longevity risk aversion coefficients γ, they proposed this scheme capable of being quite optimal for all longevity risk aversion coefficients and precisely optimal when considering a logarithmic utility function.

In the following graph we report the range of possible outcomes from the optimal natural tontine. Figure 2.2 is computed by solving for the value of c_0 and constructing $D_N({}_tp_x)$ for $n = 400$, $r = 4\%$ and $\gamma = 1$. Once the payout function $D_N({}_tp_x)$ is known for all t, the optimal cash value effectively paid to survivals, i.e $[n \cdot D_N({}_tp_x)/N(t)]$, is computed. The number of survivals $N(t)$ at the 10^{th} and 90^{th} percentile of the binomial distribution is used to bracket the range of benefits from age 65 to age 100. Clearly the expected income per survivor is relatively constant over the retirement years. This structure results to be optimal when $\gamma = 1$, but also nearly optimal for all the other levels of longevity risk aversion.

Figure 2.2: Range of natural tontine payout



Note: Interest rate r=4%, $n = 400$, Gompertz law of Mortality assumed with parameters m= 88.721 and b=10.

Source: Milevsky and Salisbury (2015)[19]

Combining Heterogeneous Cohorts

Milevsky and Salisbury extended also their analysis by combining heterogeneous cohorts into one pool.
The idea of the authors was as follows: allowing cohorts of different ages (and mortality) to mix in the same pool, by assigning different participation rates or shares, based on their age at the time of purchase. If anyone, regardless of the age, is allowed to participate equally in the same annuity pool, there would be an immediate transfer of wealth from the members who are expected to die sooner (older) to those who are expected to live longest (younger). In fact, the tontine proposed by the authors tries not to discriminate against anyone: depending on the age of the subscriber at inception, the price of shares is determined. Individuals decide on their investment, then depending on their age, the price of shares is determined and thus also the number of shares received. For example, a 55 year-old allocating \$10,000 to the tontine should pay \$200 per share, and therefore he should receive 50 shares; while a 75 year-old allocating \$8,000 to the tontine should pay only \$40 per share and then receive as many as 200 shares. Heterogeneous individuals, paying different amount per share, would all be mixed together into the same pool. However each tontine share would provide equal income rights.
Milevsky and Salisbury focused the analysis of this tontine on the concept of *equitable*, which is different from the concept of *fair*. In a *fair* tontine the expected present value of benefit paid to retirees should always be equal to the amount contributed to or invested in the tontine. The authors have shown that such a scheme, in which there is a chance for everyone to die before the maximum age, and thus leave inheritance to the pool, can never be made fair unless it incorporates some form of payment to estates. They have found that there is always a positive probability of leaving money in the pool when someone dies, making it impossible to create a fair tontine. So they decided to focus on another type of concept, *equitability*: in a *equitable* tontine the expected present value of benefit paid to retirees should always be the same for all participants in the scheme regardless of the age. The authors once again, decided to focus their attention on a scheme that does not discriminate against any cohort. Hence, they aimed to come up with the optimal tontine scheme, to determine the proper share prices to charge participants, so that the plan is *equitable* and does not discriminate against any age of the group. The tontine proposed is moreover a closed pool that does not allow anyone to enter (or obviously exit) after the initial set-up. The mathematical questions they answered are: does a collection of share prices that makes the tontine equitable exist? Are such prices unique? I will present in detail the results achieved by the authors.

Let's introduce some of the notations used:

- n is the number of subscribers at inception;
- K is the number of homogeneous cohorts into which individuals can be

grouped (those with the same age and contribution level).

For $i = 1, ..., K$ we denote by:

- x_i the starting age of individuals in the i^{th} cohort;
- w_i the dollars invested by subscribers in the i^{th} cohort;
- n_i the number of individuals in the i^{th} cohort, such that $n = \sum n_i$;
- $1/\pi_i$ the i^{th} subscriber's price for share;
- u_i the number of shares purchased by each individual in the i^{th} cohort. It is straightforward to realize that:
$$w_i = u_i \cdot \frac{1}{\pi_i} \rightarrow u_i = w_i \pi_i$$
- $N_i(t)$ the number of survivors at time t in the i^{th} cohort. As for the single cohort analysis, also here the variable $N_i(t) - 1 \sim Bin(n_i - 1, \, {}_tp_{x_i})$
- w the total initial investment, given by the sum of all the single contributions: $w = \sum n_i w_i$;
- $d(t)$ the rate of return for initial dollar invested. Hence $wd(t)$ is the total time t payout of the tontine. This total payout must be divided among those who are still alive, and each surviving investor will receive payments at a rate proportional to the amount of shares owned as can be clearly seen in the following formula:
$$wd(t) \cdot \frac{u_i}{\sum u_j N_j(t)} = wd(t) \cdot \frac{\pi_i w_i}{\sum \pi_j w_j N_j(t)}$$

The goal of Milevsky and Salisbury was to find a set of equitable shares prices, thus to define a collection $\{\pi_1, ..., \pi_K\}$. The expected present value of future benefit for an individual in the i^{th} cohort can be written as:

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\left(wd(t) \cdot \frac{\pi_i w_i}{\sum \pi_j w_j N_j(t)}\right) dt\right] = \int_0^\infty e^{-rt} \, {}_tp_x wd(t) \cdot \mathbb{E}\left[\frac{\pi_i w_i}{\sum \pi_j w_j N_j(t)}\right] =$$
$$w_i \cdot \int_0^\infty e^{-rt} \, {}_tp_x wd(t) \cdot \mathbb{E}\left[\frac{\pi_i}{\sum \pi_j w_j N_j(t)}\right] = w_i \cdot F_i(\pi_i)$$

We can see that the present vaue of future benefit is equal to a percentage $F_i(\pi_i)$ of the initial investment w_i. In order to make a tontine equitable we require the percentage earned on the initial investment by each cohort to be equal for all. Ideally we require

$$F_i(\pi_i) = F_j(\pi_j) \quad \forall i \neq j$$

The authors were able to derive the necessary and sufficient conditions for the existence of a set of equitable share prices when a tontine structure $d(t)$ has already been fixed. Here the main theorem presented:

Theorem. *Fix d(t) as well as n_i, x_i and w_i $\forall\ i = 1, ..., K$.*

1. *If there exists an equitable choice of $\pi = (\pi_1, ..., \pi_K)$ such that $0<\pi<\infty$ for each i, then this choice is unique up to an arbitrary multiplicative constant.*

2. *A necessary and sufficient condition for such a π to exist is the following*

$$\int_0^\infty e^{-rt} d(t) (\prod_{i \notin A} {}_tq_{x_i}^{n_i})(1 - \prod_{i \in A} {}_tq_{x_i}^{n_i}) dt < \alpha_A (1 - \epsilon)$$

where

$$A \in \{1, ..., K\}$$

$$\alpha_A = \frac{1}{w} \sum_{k \in A} n_k w_k$$

$$\epsilon = \int_0^\infty e^{-rt} d(t) \mathbb{P} \left(\sum_j N_j(t) \right) dt$$

This theorem is very relevant, but it is capable of providing the necessary and sufficient conditions for the existence of a set of equitable prices, only when the tontine scheme has already been defined. However the equitable prices that meet the necessary and sufficient condition may not be optimal for all cohorts. The natural question is whether it is possible to design a tontine to be optimal for multiple age cohorts. The authors tried therefore to find a set of equitable shares prices $\{\pi_1, ..., \pi_K\}$ and optimal rates of return $d(t)$ that can maximize the discounted expected utility of individuals. The maximization problem consists in finding the optimal $d(t)$ and π simultaneously, such that the tontine plan is optimal for all cohorts.

$$\max_{d(t), \{\pi_1, ..., \pi_K\}} \int_0^\infty e^{-rt} {}_tp_{x_i} \mathbb{E} \left[U \left(wd(t) \cdot \frac{\pi_i w_i}{\sum \pi_j w_j N_j(t)} \right) \right] dt$$

This maximization problem is subject to two constraints.
The budget constraint:

$$\int_0^\infty e^{-rt} d(t) dt = 1$$

And the equitable constraint. Recall that the authors wish to establish a tontine scheme that does not discriminate against anyone. The percentage earned on the initial investment must therefore be the same for all cohorts and thus for all subscribers.

$$\int_0^\infty e^{-rt} {}_tp_{x_i} \cdot wd(t) \mathbb{E} \left[\frac{\pi_i}{\sum \pi_j w_j N_j(t)} \right] dt = \int_0^\infty e^{-rt} {}_tp_{x_m} \cdot wd(t) \mathbb{E} \left[\frac{\pi_m}{\sum \pi_j w_j N_j(t)} \right] dt \quad \forall i \neq m$$

This optimization problem turns out not to be possible to solve, except when $n \to \infty$. Hence, the authors were able to define the tontine that asymptotically optimizes the utility of each cohort simultaneously, calling it the *Proportional tontine*. The *Proportional tontine* has the following characteristics:

- $$\pi_i = \left[\int_0^\infty e^{-rt} \, {}_tp_{x_i} dt \right]^{-1} \tag{2.13}$$

- $$d(t) = \sum_{j \in K} \frac{n_j w_j}{w} \cdot {}_tp_{x_j} \cdot \pi_j \tag{2.14}$$

We can recognize that in the homogeneous cohort case, when $K = 1$, the *Proportional tontine* agrees with the *Natural tontine* defined in the single cohort analysis. Again the tontine scheme $d(t)$ is proportional to the survival rate ${}_tp_{x_j}$. Moreover we can also notice that the payout rate per surviving individual from the i^{th} cohort at time t remains constant over time:

$$d(t)w \cdot \frac{w_i \pi_i}{\sum\limits_{j \in K} w_i \pi_i N_j(t)} = \sum_{j \in K} \frac{n_j w_j}{w} \, {}_tp_{x_j} \pi_i \cdot w \cdot \frac{w_i \pi_i}{\sum\limits_{j \in K} w_j \pi_j n_j \, {}_tp_{x_j}} = \pi_i w_i$$

If the i-th cohort invests a dollar at the inception of the tontine (i.e. $w_i = 1$) , the payment rate will remain constant over time (i.e. π_i) and it will correspond to the payment rate of a standard fixed lifetime annuity. Recall that the stable level income provided by an annuity for initial dollar invested is:

$$c_0 = \left[\int_0^\infty e^{-rt} \, {}_tp_x dt \right]^{-1}$$

This particular design of tontine matches (in the limits, when $n \to \infty$) the payment structure and cost of a standard annuity for each subscriber. The *proportional tontine* can thus be compared to such policies and proposed as a plausible alternative when the number of subscribers is really huge.

2.4 Fair Tontine Annuity

The Fair Tontine Annuity(FTA), introduced by Sabin(2010)[27] is an explicit tontine, in which members receive an explicit longevity credit into their personal account when an individual in the pool dies.
The FTA is an arrangement that provides lifetime payments whose expected present value matches that of a fair annuity. The main purpose of the author is to demonstrate that his FTA is favourable compared to the unfair annuities actually offered by insurance companies. We know that in general the annuities proposed by insurers are not fair: companies sell annuities to make money, so they typically charge premiums higher than they effectively need to. Actual annuities than can be purchased from an insurer, typically cost therefore more than a fair annuity. On the other hand, the cost of an FTA, because of the manner in which it was designed, can be compared to that of a fair annuity. Hence, it is evident that an FTA is an attractive alternative to conventional annuities: it is more cost effective and it can be offered by many providers, not just insurance companies.

Moreover, the FTA of Sabin is based on a fair tontine. The author indeed attempted to solve the problem of unfair tontines, by relying on a different type of scheme in which members in the pool can be of any age and gender, and new members can join at any time. The fair tontine allows each member to contribute any desired amount, but the distribution of asset to surviving members is not made in equal amount (unlike traditional tontines). Instead of providing equal portion to each individual in the pool, making the tontine unfair and advantaging someone over others (the younger over the older), in a Fair Tontine the distribution to surviving members is made in unequal proportion, according to a plan, called Fair Transfer Plan, that provides each member with a fair bet. The distribution of dividends is proportioned in such a way that it appropriately compensates each member according to his or her probability of dying and his or her amount of risk. Payouts in a fair tontine, just as in a traditional tontine, occur at random times, whenever a group member dies. Even the amounts of payments are random, as they depend on the balance account of the person who died and the proportions to which each member is entitled. The FTA is formed by adding few enhancements, which I will better analyze later, to the fair tontine, so that it can mimic a fair annuity.

I will present in the following sections the analysis made by Sabin. First, I will present the Fair Transfer Plan proposed by the author and the necessary and sufficient conditions for the existence of such a plan. After that I will analyze a fair tontine, in which payments to survivors are made in accordance with the Fair Transfer Plan. Finally, I will present the Fair Tontine Annuity, a rearrangement of a fair tontine in which payments are made to match those of a fair annuity.

Fair Transfer Plan

In a tontine, when a member dies, his account balance is distributed to surviving members of the pool. Traditionally, the distribution was made in equal portions to each survivor, or possibly in proportion to surviving members' balances. In general, this results in an unfair situation - for example, it favors younger members who are likely to live longer. In the fair tontine proposed by Sabin, surviving members do not get equal portions of a dying member's balance. Instead, the distribution is made in unequal portions, carefully chosen to make it a fair bet for all members. The distribution is governed by a Fair Transfer Plan (FTP) that takes into account each member's probability of dying and each member's account balance. In this section I describe in detail the procedures for deriving such a plane and the conditions of its existence.
Let's introduce some of the notations used:

- n is the number of members in the pool;
- ϵ_1, ..., ϵ_n are the random time of death of members $i = 1, ..., n$;
- τ is the time of the next member death: $\tau = \min(\epsilon_1, ..., \epsilon_n)$;

- p_j is the probability that the dying member is j.
  Assuming that t is the observed value of τ

$$p_j = \mathbb{P}\left[\tau = \epsilon_j | \tau = t\right] = \frac{\mu_j}{\sum_{i=1}^{n} \mu_i}$$

  where μ_i is the members' i force of mortality;

- J is an integer random variable such that $\mathbb{P}\left[J = j\right] = p_j$;

- s_j is the dollar value of the j's account balance at time t;

- $\alpha_{i,j}$ is the portion of j's balance transferred to member i. Member i receives $s_j \alpha_{i,j}$. Indeed if $i \neq j$, i is a surviving member and receives a non-negative dividend, while if $i = j$, $\alpha_{i,j} = -1$, since the dying member j forfeits all his balance.
  Notice that $\sum_{i=1}^{n} \alpha_{i,j} = 0$, meaning that the amount forfeited by j corresponds to the sum of the amount received by all the surviving members;

- R_i is the amount received by member i

$$R_i = \sum_{j=1}^{n} \alpha_{i,j} s_j \mathbb{1}_{\{J=j\}}$$

  which is equal to $-s_i$ if the dying member is i and a non negative value if some other member died.

- ER_i is the expected amount received by member i

$$ER_i = \sum_{j=1}^{n} \alpha_{i,j} s_j p_j$$

Notice that the sum of all members' expected amount is

$$\sum_{i=1}^{n} ER_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{i,j} s_j p_j = \sum_{j=1}^{n} s_j p_j \sum_{i=1}^{n} \alpha_{i,j} = 0$$

Hence, if the expected amount of some member is positive, there must be some other member with a negative expected amount. Such a situation would generate an unfair scheme, in which some individual has the advantage over someone else. To make it fair, Sabin imposed the requirement that each member's expected amount is zero, i.e. $ER_i = 0 \; \forall i$.

The Fair Transfer Plan is therefore defined by $\alpha_{i,j}$. These quantities represents the exact amount that must be transferred to each balance account i, when a member j dies. It is a plan that satisfies the following conditions in order to ensure fair bets to the entire pool.

$$\alpha_{j,j} = -1 \; for \; j = 1, 2, ..., m \tag{2.15}$$

$$0 \leq \alpha_{i,j} \leq 1 \; for \; j = 1, 2, ..., m \; when \; i \neq j \tag{2.16}$$

$$\sum_{i=1}^{n} \alpha_{i,j} = 0 \; for \; j = 1, 2, ..., m \tag{2.17}$$

$$ER_i = \sum_{j=1}^{n} \alpha_{i,j} s_j p_j = 0 \; for \; i = 1, 2, ..., m \tag{2.18}$$

The solution to such a problem requires finding n^2 unknowns, namely $\alpha_{i,j}$, subject to $3n$ equality constraints and $n(n-1)$ inequality constraints. The necessary condition that ensures the existence of such a plane is stated in the following theorem.

Theorem. *Let t be the time of death of a member and let n be the number of members in the pool at time t. Let p_j be the probability that member j is the one who died at time t. Let s_j be the balance held by member j at time t. Then a Fair Transfer Plan exists if and only if*

$$p_i s_i \leq \sum_{j=1}^{n} p_j s_j \; \forall i = 1, ..., n$$

The theorem is met by not allowing any one member to contribute an excessive amount. The member i's risk of loss, i.e. $p_i s_i$, can not be higher than the risk of loss of the entire pool. Otherwise, whenever the necessary condition is not satisfied, member i is not fairly compensated for his risk.

Let me show you an example in order to better understand the essence of a Fair Transfer Plan. Consider a pool composed of 4 individuals with the following characteristics.

i	Age	Gender	s_i	p_i
1	80	male	2	0.55464
2	71	female	6	0.15983
3	70	female	3	0.14447
4	65	male	2	0.14107

Here a Fair Transfer Plan for the pool described above. This FTP satisfies all the conditions required for its existence and fairness.

-1	0.75754	0.54370	0.52162
0.61302	-1	0.39621	0.38012
0.23766	0.14814	-1	0.09826
0.14932	0.09432	0.06010	-1

The actual amount received by surviving members depends on which member has died. Consider for instance that member 3 died: his balance of $s_3 = 3$ is distributed to the pool according to the third column of the FTP. Member 1 receives $0.54370 \cdot 3 = 1.6331$ dollars; member 2 receives $0.39621 \cdot 3 = 1.18863$ dollars; member 3, the one who died, forfeits his entire balance and member 4 receives $0.06010 \cdot 3 = 0.18030$ dollars.

Fair Tontine

The Fair Transfer Plan is used to build a fair tontine. In a fair tontine, members of any age and gender join the pool by contributing a desired amount. The initial contribution of each member may be invested, allowing the fund to grow or shrink based on investment performance. Investments may be managed collectively for the entire pool, or individually. Dividends and other cash earnings from investments are retained in the pool and credited to the member's balance. Then, each time a member dies, his balance is distributed to surviving members according to a Fair Transfer Plan. If the dying member is j, then each surviving member i receives $\alpha_{i,j}s_j$ dollars, where $\alpha_{i,j}$ is derived from a FTP. If the pool is large, members die often, causing very frequent payments to survivors. New members can join at any time, by making a contribution of a desired amount; thus, the fair tontine can be considered to operate in perpetuity. No member may withdraw from his account balance, ever. Once a contribution is made, it remains in the pool, along with any investment gain; at the member's death, the balance is distributed to other members as survivor benefits. The situation is identical to a conventional annuity: once the premium is paid, there is no refund of it, ever.
An important and surprising property of a fair tontine is that the expected payout of surviving members does not depend on the the characteristics of the pool. The number of members of the pool, ages, genders and contributions of other people does not matter. The expected payout of a survivor depends only on the member's own balance and and own probability of dying. Let's see the formula derivation.

Consider the situation where a member i survives some time interval, say $(t_1; t_2)$. During the interval, some random collection of zero or more members die. Each time a death occurs, survivor benefits are calculated by first deriving a FTP, then using the FTP to distribute the dying member's balance. The total benefit that member i receives for surviving the interval - meaning the sum of the survivor benefits paid to member i from FTPs during $(t_1; t_2)$ - is a random quantity. It depends on which members die, when they die, and what their balances are when they die. This is a complicated function. However, the expected value of member i's total benefit has a simple expression, stated in the following theorem.

Theorem. *In a fair tontine, let s(t) for $t \in (t_1; t_2)$ be the balance, at time t, of a member who joins prior to t_1. Let $R(t_1; t_2)$ be the sum of payments received by the member as survivor benefits during $(t_1; t_2)$. Let $A(t_2)$ be the event that the member is alive at t_2. Then*

$$\mathbb{E}\left[R(t_1; t_2)|A(t_2)\right] = \int_{t_1}^{t_2} s(t)\mu(t)dt$$

where $\mu(t)$ is the member's force of mortality function.
Furthermore, if $s(t)$ is constant on $(t_1; t_2)$, i.e. the member contribution is not

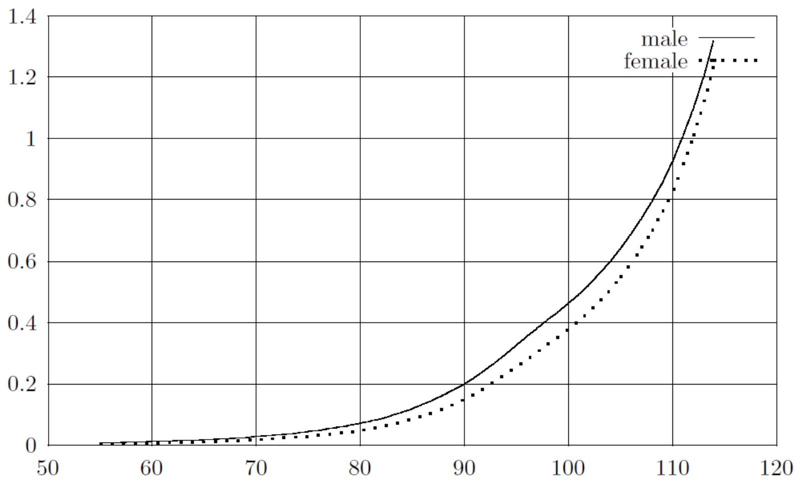
invested, then

$$\mathbb{E}\left[R(t_1;t_2)|A(t_2)\right] = s(t)\ ln\left(\frac{1-F(t_1)}{1-F(t_2)}\right) = s(t)\frac{q_t}{1-q_t}$$

where $F(t)$ is the member's distribution function for time of death and q_t is the probability of dying during the interval.

A member's expected benefit for surviving is given by a simple formula that depends only on the member's own balance and own probability distribution for time of death. The demographics of the pool does not matter; all that matters is that FTPs are used to compute survivor benefits. A young person in a pool dominated by elderly members has the same expected benefit as that young person would have in a pool made up of individuals matching her age. The actual payout that a surviving member receives during a time interval is, of course, a random variable. The probability distribution function of that random variable most definitely depends on the parameters of the other members. However, regardless of those parameters, its mean value is as just described.
Figure 2.3 shows the normalized expected payout, i.e. considering $s(t) = 1$, for a surviving member in a one-year time interval.[8] We can see that the benefit payout increases with age: an older member's probability of dying is higher than that of a younger one. The fairness of a fair tontine requires that the older member is compensated for the higher risk, providing him with a larger expected benefit. Similarly, the benefit for men is higher than for women, because men are more likely to die.

Figure 2.3: Expected normalized benefit for surviving member in a Fair Tontine



Source: Sabin(2010) [27]

[8]In general the member's expected benefit is also proportional to her account balance s.

Fair Tontine Annuity

In the fair tontine just described, members make a one-time payment (the contribution), of a desired amount, and in return receive a stream of benefits that lasts for their lifetime. In some way, the fair tontine has the same features as an annuity. Where an annuity and a fair tontine differs is in the timing and values of payments: the annuity makes payments on a fixed schedule, say monthly, with each payment having a constant expected amount. While the fair tontine makes payments at random times (whenever a member of the pool dies), with each benefit being a random amount whose mean value increases with the age of individuals.

Sabin attempted to replicate the structure of a fair annuity by proposing a new type of insurance policy: the Fair Tontine Annuity. In this section I present the Fair Tontine Annuity which is constructed by adding two enhancements to the fair tontine.

- Monthly accrual of payouts. To solve the problem of random timing of payments of fair tontines and to mimic the fixed timing of annuity payouts, we can modify the tontine as follows. Each time a member dies, payments from the dying member are not immediately distributed, instead they are accrued within surviving member's accounts. At the end of the month, accrued payments in each surviving member's account are paid out in a lump sum. If a member dies during the month, she forfeits both the original contribution plus any accumulated payments for that month. Thus, a member receives a payout only if she is alive at the end of the month—just as with an annuity. With monthly accrual as just described, the member receives monthly payments as in an annuity.
- Self payback. To solve the problem of random amount of payments of fair tontine and to mimic the constant amount of annuity payouts, we can modify the tontine as follows. At the end of each month, each surviving member receives a benefit derived from the balances of dead people, as in traditional fair tontine, plus also a portion of their own balance. This "self-payback" is made according to a predefined plan such that the expected value of the total monthly payment - meaning the expected value of the sum of the survivor benefit plus the self payback - behaves like the monthly payment of a fair annuity. We know that survivor benefits, in accordance with FTP, are increasing with age; therefore, we expect "self paybacks" to be decreasing with age so that we can ensure constant total monthly payments over time.

I analyze in the following subsections the derivation of the expected payouts for both a fair annuity and a fair tontine annuity, so that we can better understand their similarity.

Fair annuity - monthly payments

The fair annuity that Sabin wishes to replicate has the following features: the initial premium paid at time t_0 by the retiree is allocated to an investment portfolio. Benefit amounts d_n are paid monthly - at time $t_n = t_0 + n \cdot \frac{1}{12}$ for $n = 1, 2...$ - and fluctuate according to the performance of the fund. d_0 is the initial payout amount specified in the annuity contract and a is the assumed interest rate. i_n, for $n = 1, 2...$, are instead the actual interest rates earned by the underlying investment in the time interval $(t_{n-1}; t_n)$. If the actual interest rate i_n always matches a - i.e. the assumed interest rate of the contract - then each payment d_n is the constant value d_0; otherwise d_n fluctuates according to:

$$d_n = d_0 \prod_{k=1}^{n} \left(\frac{1 + i_k}{1 + a} \right)^{1/12}$$

The initial premium $v(t_0)$ paid at time t_0 by the annuitant is calculated assuming that each i_n= a and it is defined as:

$$v(t_0) = d_0 \cdot L_a(t_0)$$

where $L_a(t_0)$ is the annuity factor, i.e. the fair premium paid at time t_0 for normalized monthly payments $d_0 = 1$ dollar.
Similarly, at time t_n, if the annuitant is still alive, the balance $v(t_n)$ is the premium the annuitant would pay to purchase an annuity with initial payment amount of d_n.

$$v(t_n) = d_n \cdot L_a(t_n) = d_{n-1} \left(\frac{1 + i_n}{1 + a} \right)^{\frac{1}{12}} \cdot L_a(t_n) \tag{2.19}$$

$$= \frac{v(t_{n-1})}{L_a(t_{n-1})} \left(\frac{1 + i_n}{1 + a} \right)^{\frac{1}{12}} \cdot L_a(t_n) \tag{2.20}$$

Let u_n be the change in balance that occurs at time t_n:

$$u_n = v(t_{n-1})(1 + i_n)^{1/12} - v(t_n) \tag{2.21}$$

u_n represents a partial distribution of the account balance to the annuitant. It is a "self-payback", a withdrawal by the individual of the invested premium. By combining (2.14) with (2.15) we obtain:

$$u_n = v(t_{n-1})(1 + i_n)^{1/12} \left[1 - \frac{L_a(t_n)}{L_a(t_{n-1})(1 + a)^{1/12}} \right] \tag{2.22}$$

A part of the premium received at time t_n is therefore given by u_n, i.e. the "self-payback". The remaining portion of payment is defined by b_n:

$$b_n = d_n - u_n = ... = v(t_{n-1})(i + i_n)^{1/12} \frac{q_n}{1 - q_n}$$

where q_n is the probability of dying in the time interval $(t_{n-1}; t_n)$. The quantity b_n can be seen as the the annuitant's "survivor benefit" for $(t_{n-1}; t_n)$. It is paid by the insurer, funded by the forfeited balances of other annuitants who have died. In summary, a fair annuity can be viewed as a private account held by the annuitant. The initial account balance $v(t_0)$ is the premium paid by the annuitant. At the end of month n, if the annuitant is alive, a portion of the current balance is paid back to the annuitant, reducing the balance to $v(t_n)$. This self payback u_n is supplemented by a survivor benefit b_n, paid by the insurer. The sum $d_n = u_n + b_n$ is the monthly payment received by each individual.

FTA

The Fair Tontine Annuity has the same characteristics as the fair tontine, but, as mentioned above, exhibits two adjustments so that it can perfectly mimic the fair annuity just described. At time t_0 the annuitant pays the initial contribution $v(t_0)$, which is allocated into an investment portfolio. Each member's portfolio can be individually managed. We define by $\phi_i(t_{n-1}, t_n)$ the growth function of the member i's investment, meaning that if the investment has value $v_i(t_{n-1})$ at time t_{n-1}, then its value at time t_n is $v_i(t_{n-1})\phi_i(t_{n-1}, t_n)$. Each member receives payments at time t_n, monthly. These income payouts are, as in the case of a fair annuity, made up of two amounts: a self payback and a survivor benefit.
The self payback amount u_i behaves in line with that of the fair annuity, according to the formulas (2.14),(2.15).

$$u_i(t_n) = v_i(t_{n-1})\phi_i(t_{n-1}, t_n) - v_i(t_n)$$

$$v_i(t_n) = \frac{v_i(t_{n-1})}{L_i(t_{n-1})} \cdot L_i(t_n) \frac{\phi_i(t_{n-1}, t_n)}{(1 + a_i)^{1/12}}$$

The survivor benefit received at time t_n is instead defined by $B_i(t_n)$. Consider a period $(t_{n-1}; t_n)$. Let N be the number of people in the pool at time t_{n-1}. Let k be the number of members died during the period. We define by $\tau_1, ..., \tau_k$ the times of death occurring in the time interval, with $t_{n-1} = \tau_0 < \tau_1 < ... < \tau_k < t_n$. The survivor benefit $B_i(t_n)$ consists of the accrued proceeds from the Fair Transfer Plan. Initially $B_i(\tau_0) = 0$, then at each member death time τ_k, it is updated as follows:

$$B_i(\tau_m) = B_i(\tau_{m-1})\phi_i(t_{m-1}, t_m) + \alpha_{i,j}(\tau_m)s_j(\tau_m) \quad for \;\; m = 1, ..., k$$

The balance at each death time is therefore equal to the accrued accumulated fund plus the benefit received from the balance of member j who died. Hence,

- $\alpha_{i,j}(\tau_m) \quad for \;\; m = 1, ..., k$ is the Fair Transfer Plan defined at time τ_m;
- $s_j(\tau_m)$ is the balance of member j at time τ_m which evolves itself according to

$$s_j(\tau_m) = v_j(t_{n-1})\phi_j(t_{n-1}, \tau_m) + B_j(\tau_{m-1})\phi_j(\tau_{m-1}, \tau_m)$$

At time t_n if the individual is still alive he is paid out the accrued survivor benefit

$$B_i(t_n) = B_i(\tau_k)\phi_j(\tau_k, t_n)$$

The total payment that living member receives at time t_n is

$$D_i(t_n) = u_i(t_n) + B_i(t_n)$$

Sabin stated in a theorem that the expected value of this payment turns out to be exactly equal to the one of a fair annuity.

Theorem. *In an FTA, let $v(t_0)$ be a member's initial contribution at a payment time t_0. Let $L_a(t_0)$ be the annuity factor of a fair annuity purchased by the member at t_0 with assumed interest rate a. Let $\phi(t'; t'')$ be the growth rate of the member's investment. Let $D(t_n)$ be the payment the member receives at payment times $t_n = t_0 + \cdot \frac{n}{12}, n = 1, 2, \ldots$. Let A_n be the event that the member is alive at time t_n. If an FTP exists at each member death in $(t_{n-1}; t_n)$, then*

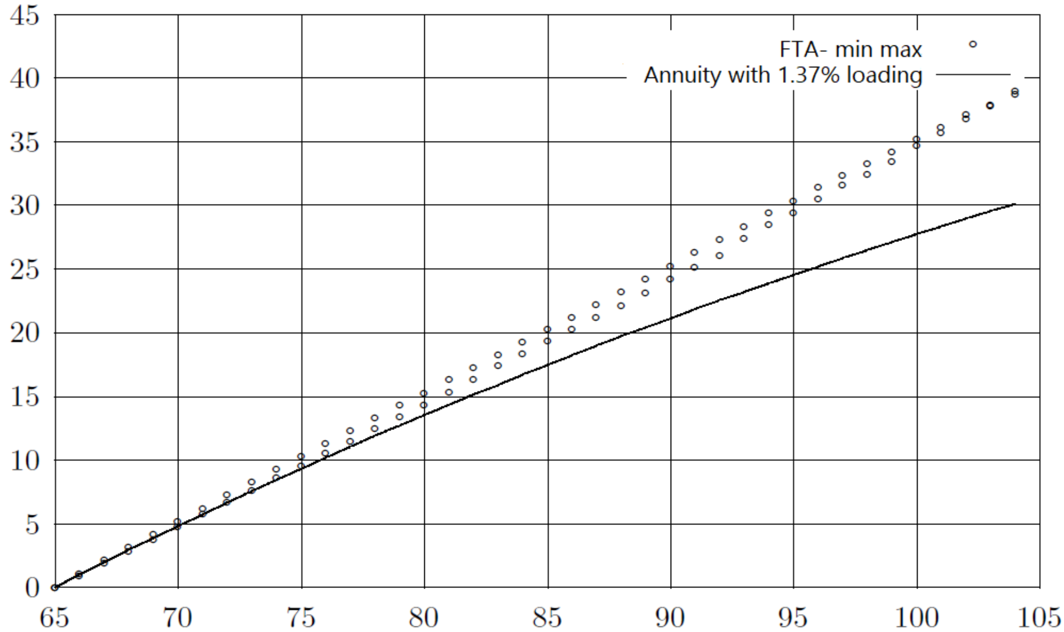
$$\mathbb{E}\left[D(t_n)|A(t_n)\right] = \frac{v(t_0)}{L_a(t_0)} \cdot \frac{\phi(t_0; t_n)}{(1+a)^{\frac{n}{12}}}$$

which behaves in line with the definition of d_n in a fair annuity.

Again, according to the theorem, the parameters of other members of the pool - their ages, genders, balances, assumed interest rates, etc. - do not affect an individual member's expected monthly payment. All that matters is that FTPs exist so that they can be used to compute survivor benefits.

Finally, the author, in order to emphasize the importance of the new type of insurance policy suggested, made a comparison between a fair tontine annuity and a conventional annuity sold on the market with a 1.37% loading charge. His main goal is to underline that an FTA outperforms in terms of cost effectiveness the actual annuities that can be purchased from a typical insurance. In the following graph he shows the maximum and minimum cumulative payout over all members in a FTA: he made a simulation and plots the largest and smallest cumulative payout that any such member receives at a given age. For comparison, Figure 2.4 also includes the normalized cumulative payout of a fair annuity that imposes a 1.37% mortality and expense charge.
What is perhaps surprising is that, in the figure, the FTA outperforms the annuity for every member who lives beyond, roughly, age 75. The variance in FTA payouts is more than offset by the cost charged by insurers, making it preferable. This phenomenon is emphasized even more when the number of people in the tontine pool is larger: the variance of payouts in a FTA is greatly reduced when the pool is large, making this type of product even more preferred. The FTA exhibited a better payout than a typical insurer-provided annuity not just on the average, but for virtually every member who lived more than just a few years.

Figure 2.4: Maximum and Minimum Cumulative payout over all FTA members vs Cumulative payout of an Annuity with 1.37% loading charge



Source: Sabin(2010) [27]

The FTA would seem to be an attractive alternative to insurer-provided annuities, at least to those who can tolerate variation in monthly payments in exchange for a better long-term result. The FTA is more cost-effective (no insurer profit), allows for more flexible portfolio management (such as a self-directed brokerage account), and is mostly void of creditworthiness concerns (the assets are owned by the members). Moreover the FTA could be offered by low-cost providers such as mutual fund houses, retail brokers, etc. and not only by insurance companies.

2.5 Annuity Overlay Fund

The Annuity Overlay Fund of Donnelly et al. (2014)[9] is another explicit tontine scheme. As in the Fair Tontine Annuity of Sabin(2010)[27], described in the previous section, members receive an explicit longevity credit into their personal account when an individual in the pool dies. On the other hand, unlike Sabin's plan, the AOF proposed by Donnelly is actuarially fair for any heterogeneous group, it does not require some necessary conditions for the existence, for example, of a fair transfer plan, which ensures fair bets.

The main goal of the authors is to enable cost transparency in life annuity products: in general, consumers have no idea whether annuity prices are fair, whether insurance companies are making excessive profits, or whether available information is adequate. Sources of cost are not as clear defined, there is a lack of information for retirees, and this means that individuals can not make informed decision when managing their assets. Thus, Donnelly aimed to create a very transparent product, in which costs are attributed to each source independently and charged to each individual when they occur. The motivation for the Annuity Overlay Fund is to arrive at a transparent market, in which people really understand what they are paying for and can determine if the costs charged are reasonable. Hence Donnelly envisioned a scheme in which each year participants receive an investment statement that details their current individual wealth, how much they have gained from investments, how much they have earned from mortality credits (as always in tontines the wealth of the dead people is divided among participants), management fees, administrative costs, and all other sources of expenses. In addition, each participant could receive annual information on future projected mortality credits in order to get a better idea of the income to which they will be entitled.

Another important feature of the tontine proposed by Donnnelly is the following: individuals can leave the pool before death without incurring in financial penalties. This is a very surprising property, never suggested in any other tontine scheme, that allows individuals to remove the Annuity Overlay Fund from all or some of their assets at any time. Similarly, participants can decide to add the overlay to more of their assets at any time. This is made possible because a portion of mortality credit is also given to the estates of people who have just died, making the Annuity Overlay Fund actuarially fair at every instant in time. The amount of flexibility given to retirees is very high: they can also control their own investments, they can decide how much to invest and where to allocate their wealth across asset classes.

In the following subsections I will analyze the Annuity Overlay Fund in detail. First I will present a toy example, to give a clear idea of the essence of the plan. Then, I will present the actual real plan proposed by Donnelly, with a full analysis of mortality credits, payments structure and evolution of wealth. Finally I will show some important results achieved by the authors through a numerical simulation.

Basic Example

I first illustrate the Annuity Overlay Fund with a toy example that provides the basic idea of the plan. Here, the investment risk is not considered: no financial returns on wealth are assumed. Each participant has a fixed initial wealth which is not invested in the market. The only things that really matter in this example are the risk pooling mechanism and how mortality credits are shared.

Let's imagine that two people, Alice and Bob, decided to enter into an Annuity Overlay Fund and pool their mortality experience together for one month. The characteristics of the two individuals are summarized in the following table:

Name	Wealth	q_i
Alice	\$1 000 000	0.2%
Bob	\$50 000	0.1%

where q_i represents the probability of dying in the next month of each member. Whenever a member dies, his wealth is put into a notional mortality account. At the end of the month, the money in the notional mortality account is shared among all the participants, including those who just died during the month - in this case money is paid to estate. The payment that each participant receives from the notional mortality account, called mortality credit, is proportional to his individual mortality rate and to his own wealth. If Alice dies during the month, then her wealth is put into the notional mortality account, and the same rule applies for Bob. Thus, at the end of the month, the money in the notional mortality account is shared among Alice and Bob, in proportion to their wealth and probability of dying.
Suppose Bob is the only one to die during the month. His wealth of \$50 000 is put into the notional mortality account and then, at the end of the month, the money in the fund are shared out as follows:

- Alice gets:

$$\$50\,000 \cdot \frac{\$1\,000\,000 \cdot 0.2\%}{\$1\,000\,000 \cdot 0.2\% + \$50\,000 \cdot 0.1\%} = \$48\,780$$

- Bob gets:

$$\$50\,000 \cdot \frac{\$50\,000 \cdot 0.1\%}{\$1\,000\,000 \cdot 0.2\% + \$50\,000 \cdot 0.1\%} = \$1\,220$$

The actuarial gain for Alice is equal to \$48 780, given by the amount of money that is added to her fund, providing her with a total wealth of \$1 048 780 at the end of the month.
By the other side, Bob forfeits his wealth of \$50 000, but his estate gain \$1 220, providing him with a total actuarial gain equal to $-\$48\,780$. What is lost by Bob is earned by Alice. No money is created by pooling mortality risk: the wealth of dead people is simply re-distributed among all the participants. The expected actuarial gain for each participant is equal to zero over all scenarios. For this reason at the end of the month neither Alice nor Bob have any further actuarial obligation to each other; they can then withdraw their money and choose whether or not to pool their mortality for another month. The sum paid to Bob's estate can indeed be thought of as a balancing item to make the Annuity Overlay Fund work for any group of heterogeneous participants.
The same approach can be used to pool mortality risk among a larger group of individuals. Let's see how it operates.

AOF

Here I will present how the Annuity Overlay Fund operates theoretically, on a instantaneous basis. Unlike the toy example, here we will assume returns on investment: we will model therefore also the dynamics of assets available in the financial market.
Let's introduce some of the notations used:

- $M \in \mathbb{N}$ represents the number of homogeneous groups of individuals who participate in the AOF. Individuals in the same group are homogeneous in the sense that they have the same mortality characteristics, risk preferences, age and initial wealth;

- L_0^m represents the number of individuals alive aged x_m in the m^{th} group at time 0;

- $N_t^{m,i}$ models the survival of the i^{th} individual in the m^{th} group at time t. It is a Poisson process which has value 0 if the individual is alive at time t, and 1 if the individual is dead;

- λ_t^m is the force of mortality at time t of the Poisson process $N_t^{m,i}$;

- N_t^m models the survival of the m^{th} group at time t. It denotes the number of death that have occurred up to and including time t in the m^{th} group.

$$N_t^m = \sum_{i=1}^{L_0^m} N_t^{m,i}$$

  It is again a Poisson process with rate $\lambda_t^m L_{t^-}$ at time t, where L_{t^-} represents the number of people alive up to t^-;

- B_t is the price of the risk free asset available in the financial market:

$$dB_t = rB_t dt$$

  where r is the constant risk free rate of return.

- S_t is the price of the risky asset available in the financial market. It evolves according with a geometric brownian motion.

$$dS_t = S_t(\mu dt + \sigma dZ_t)$$

$$S_0 > 0 \quad Constant$$

  where Z_t is a standard brownian motion, μ is the drift component and σ the volatility.

- W_t^m is the wealth at time t of each participant in the m^{th} group. If an individual in the m^{th} group dies during the short interval of time $(t^-; t)$, then her wealth $W_{t^-}^m$ is put in the notional mortality account;

- U_t represents the amount of money which has passed through the notional mortality account up to time t. The amount of money which is put into the notional mortality account during the short interval of time $(t^-; t)$ is:

$$dU_t = \sum_{m=1}^{M} W_{t^-}^m dN_t^m$$

The amount dU_t is then shared out at time t among all the participants who were alive at time t^-. The amount allocated to each participant is proportional to their individual wealth and force of mortality. Thus each individual in the k^{th} group who was alive at time t^- receives a payment at time t of:

$$dU_t \cdot \frac{W_{t^-}^k \lambda_t^k}{\sum\limits_{m=1}^{M} W_{t^-}^m \lambda_t^m L_{t^-}^m}$$

These payments, called mortality credits, are made independently of whether or not participant are alive at time t.
Let us remind that the main objective of the authors is to ensure cost transparency of annuity products and to increase the level of information provided to investors. Therefore, it is very important to clearly differentiate the financial risk from the mortality risk: it is necessary to separate and define in detail the differences between the investment gains derived from the financial market and the actuarial gains derived from sharing the mortality risk.
We denote by $G_t^{k,i}$ the total actuarial gains up to time t of an individual i in the k^{th} group. The change in the actuarial gains at time t is given instead by

$$dG_t^{k,i} = \begin{cases} dU_t \cdot \frac{W_{t^-}^k \lambda_t^k}{\sum\limits_{m=1}^{M} W_{t^-}^m \lambda_t^m L_{t^-}^m} - W_{t^-}^k & \text{if individual } i \text{ dies between} (t^-;t) \\ dU_t \cdot \frac{W_{t^-}^k \lambda_t^k}{\sum\limits_{m=1}^{M} W_{t^-}^m \lambda_t^m L_{t^-}^m} & \text{if individual } i \text{ is alive at time } t \\ 0 & \text{if individual } i \text{ dies before } t^- \end{cases}$$

The authors stated in a theorem a very important insight regarding instantaneous changes in actuarial gains.

Theorem. *The expected instantaneous actuarial gains for a participant in the Annuity Overlay Fund are zero at all times, i.e for individual i in the k^{th} group*

$$\mathbb{E}\left[dG_t^{k,i} | \mathbb{F}_{t^-}\right] = 0 \ \forall i = 1, ..., L_0^k \ \text{and} \ \forall k = 1, ..., M$$

The theorem emphasizes that the Annuity Overlay Fund is actuarially fair at any given point in time: this important property ensures that individuals can exit from the pool whenever they want and without incurring in financial penalties.
However, even though the expected actuarial gains are zero, the incentive to join the AOF is that the actuarial gains for a participant who survives are always nonnegative. Indeed, conditional upon survival, the expected instantaneous actuarial gains for a participant in the AOF are nonnegative at all times. As long as participants survive, they do not lose financially from participating in the fund; they do not lose any of their money from pooling mortality risk until they die.
If we denote by π_t^k the portion of wealth invested in risky assets at time t, the total wealth dynamic of an individual i in the k^{th} group is described by:

$$dW_t^k = W_{t^-}^k \left[\left(r + \pi_t^k (\mu - r) \right) dt + \sigma \pi_t^k dZ_t \right] + dG_t^{k,i} \tag{2.23}$$

Subject to:

$$W_0^k = w_0^k > 0$$

Equation (2.23) clearly defines the differences in gains: the first two terms are due to investment in the financial market, they represents the income derived from both the risky and risk free asset. The third term $dG_t^{k,i}$, on the other hand, represents the instantaneous actuarial gains from participating in the fund. Again, we can see the clear separation between the two sources of revenue in the AOF.

In order to re-emphasize the decomposition of the two sources of risk, we also report expectation and variance of the instantaneous return on wealth, conditional upon survival:

$$\mathbb{E}\left[\frac{dW_t^k}{W_{t^-}^k}|N_t^{k,i}=0\right] = \left(r + \pi_t^k(\mu - r) + \lambda_t^k\left(1 - \frac{W_{t^-}^k\lambda_t^k}{\sum\limits_{m=1}^{M} W_{t^-}^m\lambda_t^m L_{t^-}^m}\right)\right)dt$$

$$Var\left[\frac{dW_t^k}{W_{t^-}^k}|N_t^{k,i}=0\right] = \left((\sigma\pi_t^k)^2 + (\lambda_t^k)^2\left(\frac{\sum\limits_{m=1}^{M}(W_{t^-}^m)^2\lambda_t^m L_{t^-}^m - (W_{t^-}^k)^2\lambda_t^k}{(\sum\limits_{m=1}^{M} W_{t^-}^m\lambda_t^m L_{t^-}^m)^2}\right)\right)dt$$

Within both of these quantities, we can clearly recognize a component due to the investment of the i^{th} individual in the financial market and a component due to the sharing of mortality risk. In this way, the investor is able to identify in detail the sources of risk, actually understand where the gains are coming from and decide whether the plan implemented is effectively suited to his needs.

Infinite AOF

Here I present an idealized version of the Annuity Overlay Fund, called Infinite Annuity Overlay Fund, in which there are infinitely many participants in each homogeneous group. The actuarial fairness previously presented and analyzed continues to hold regardless of the number of participants and the heterogeneity between groups.

In this idealized version of AOF, we assume that at time $t > 0$, each group has exactly the same number of members, so that $L_{t^-} = L_{t^-}^1 = L_{t^-}^2 = ... = L_{t^-}^M$. The instantaneous actuarial gains provided to individual i in the k^{th} group , who is assumed to be alive at time t, can then be rewritten as:

$$dG_t^{k,i} = dU_t \cdot \frac{W_{t^-}^k\lambda_t^k}{L_{t^-}\sum\limits_{m=1}^{M} W_{t^-}^m\lambda_t^m}$$

Now, if we let the number of participant tend to infinity, i.e. $L_{t^-} \to \infty$, the expected instantaneous actuarial gain for an individual who is still alive at time t is

$$\mathbb{E}\left[dG_t^{k,i}|N_t^{k,i}=0\right] = \lambda_t^k W_{t^-}^k dt$$

while the variance is

$$Var\left[dG_t^{k,i}|N_t^{k,i}=0\right]=0$$

In an infinite annuity overlay fund, deaths occur continuously, which releases a continuous flow of money into the notional mortality account. As this is shared among infinitely-many participants, their individual wealth increases at a continuous rate equal to their own force of mortality, with zero volatility. In this perfect pool, the volatility of return on wealth arises solely from investment in the financial market.
Having this in mind, we can rewrite the wealth dynamic as:

$$dW_t^k = W_{t^-}^k\left[\left(r+\pi_t^k(\mu-r)+\lambda_t^k\right)dt+\sigma\pi_t^k dZ_t\right] \tag{2.24}$$

The additional term $\lambda_t^k W_{t^-}^k dt$ emphasizes the benefit of joining the infinite annuity overlay fund. The pooling of mortality risk increases the wealth of single individuals in the fund, providing them with higher expected returns.

Final conclusions

To point out the relevance of the product proposed by Donnelly et al., the authors attempted to make a comparison between an Annuity Overlay Fund and a mortality linked fund.
A mortality linked fund is a widely available insurance product very similar to the AOF: benefits received by annuitants are due to both investment gains and to actuarial gains. In this case, actuarial gains are not random amount derived from other dead members' accounts, but are deterministic amount provided directly by the insurance company. The deterministic mortality-linked interest rate that the insurer pays on a member's wealth is equal to the member's force of mortality but with a reduction to allow for costs. The costs are what the insurer of the mortality–linked fund charges to the individual to remove the latter's mortality risk. Unlike the AOF, where the longevity risk is shared among the members of the fund, here it is the insurance company that bears the risk: hence it asks to be rewarded by charging loadings. If we denote by a_t^k the costs charged to individual k at time t, the wealth dynamic of a mortality-linked fund owner is:

$$dW_t^k = W_{t^-}^k\left[\left(r+\pi_t^k(\mu-r)\right)dt+\sigma\pi_t^k dZ_t\right]+(1-a_t^k)\lambda_t^k W_{t^-}^k dt \tag{2.25}$$

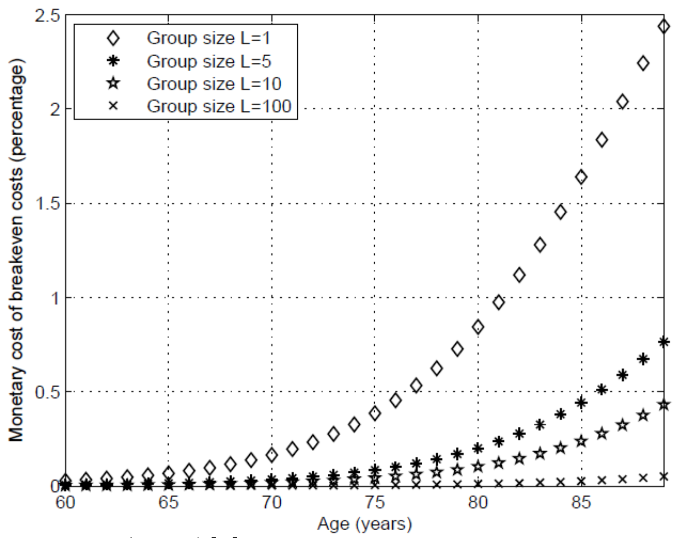
Subject to:

$$W_0^k = w_0^k > 0$$

which results to be very similar to (2.23).

In the simulation performed by the authors, their main objective was to demonstrate that their product, the Annuity Overlay Fund, actually outperforms the mortality linked funds available in the market. To accomplish this, they first calculate the instantaneous break-even costs: the a_t^k costs borne by a mortality-linked fund such that, given the same instantaneous volatility of wealth return, the expected instantaneous returns on wealth from the AOF and the mortality-linked fund are equal. The idea is that if the actual costs charged by insurance companies are higher than the instantaneous break-even costs calculated by the authors, then an individual can obtain an higher expected return from the AOF for the same amount of volatility on wealth. In figure 2.5 the authors showed their results by expressing the break-even costs as monetary rate per unit of wealth: $1 - e^{\lambda_k a_k}$

Figure 2.5: Break-even costs for different pool sizes



Source: Donnelly et al.(2014)[9]
Note: Break-even costs are expressed in monetary terms.

The authors performed the simulation for different pool sizes: the first observation is that there is an inverse relationship between the total number of participants and break-even costs. The higher the amount of people who decide to share their mortality risk, the lower break-even costs: hence the more an AOF outperforms a mortality linked fund. If break-even costs are low, it means that costs charged by insurance companies must be low, to ensure the same return of an AOF. If we look at the graph, we clearly understand that costs charged by insurers of a mortality linked fund, must be lower than 0.5% per year of wealth - across all groups - in order to be as attractive as the Annuity Overlay Fund.
In general, even when the total number of participant is small, break-even costs are small; but the effect is amplified when the group size increases.
Considering that insurance companies actually charge costs higher than 0.5% of wealth to their annuitant, in order to cover all expenses, the Annuity Overlay Fund proposed by Donnelly outperforms almost surely the mortality-linked fund available in the market.
We discover that also this new type of tontine scheme, proposed by Donnelly et al.(2014), turns out to be most cost-efficient for customers than other insurance products actually sold by companies. Moreover, the proposed AOF seems very peculiar, both for its innovative attractive features and for its high focus on transparency.

2.6 Pooled Annuity Fund

The last modern tontine scheme that I would like to present is the Pooled Annuity Fund proposed by Stamos(2008)[30]. The PAF is an explicit tontine in which longevity credits are determined for an homogeneous pool. In fact, members of the PAF are assumed to have the same age, to follow the same investment strategy and to have the same force of mortality. The assumption of homogeneity seems to be a bit restrictive, but in this way the author was able to determine analytical solutions. Stamos aimed to solve a utility maximization problem and to determine both the optimal amount of wealth invested in risky assets and the optimal consumption rate of the fund. The author modeled the population through a Poisson process, the risky assets through a geometric Brownian motion, and in this way he was able to present the wealth dynamics of both single individuals and the total Pooled Annuity Fund.

The Pooled Annuity Fund constitutes an alternative way to protect against longevity risk compared to purchasing a standard lifetime annuity. We know that insurance companies, whenever they sold a lifetime annuity, collect premiums of annuitants and invest them into capital markets. A part of the reward received by each retiree is therefore derived from the investment returns achieved by the insurer; instead, another part of the benefit is called mortality credit. The mortality credit is paid out in a deterministic manner by the insurer, and is made of redistributed wealth of deceases former members. The mortality risk in standard annuity is thus directly transferred to the insurance company and the annuitant bears no risk.
A similar mechanism applies for the PAF: part of the investor's gain is due to investment returns while another part of the gain is due to mortality credits, redistribution of wealth of dead members. Unlike standard products available on the market, in the PAF, members have access to equity markets, can decide on the optimal asset allocation strategy, can diversify their portfolios, and take advantage of risk sharing. Moreover, in the PAF, mortality credits are not deterministic, they depend on population development: therefore, in this case, participants have to bear some remaining systematic mortality risk.

Stamos, running a simulation, showed that in most cases the Pooled Annuity Fund is preferred to fixed payout annuities: investors value having access to the stock market and the freedom to choose more than having perfect longevity insurance. In general people prefer to decide on the optimal asset allocation strategy rather than completely eliminate their risks. Only very risk averse investors are more likely to pay risk premium to access private lifetime annuities: they prefer to completely lay off mortality risk. However, even for a small pool size, Stamos showed that PAFs are very effectively against longevity risk. Obviously, if we increase the effect of pooling, if the number of members in the PAF increases, significant utility gains are generated. The more pronounced is the sharing effect, the more significant will be the increases in wealth of fund

members.
Hence, Stamos presented in his paper a modern tontine capable of outperforming standard annuities. I will examine in the following sections the PAF in detail: I will present its mechanism, the population model, financial markets models and wealth dynamics.

Population Model

In order to determine the dates at which wealth of perished members are reallocated among survivors, we have to employ an appropriate population model. Stamos sets up a model in which the time of death of each pool member is determined by the first jump of a Poisson Process with time-dependent intensity. Let L_0 be the initial number of pool members.
Each pool member's time of death τ_i, $i \in \{1, ..., L_0\}$, is determined by the first jump of a Poisson Process $N_{t,i}$ with intensity parameter $\lambda_{t,i}$.

$$\tau_i = \{\min_t : N_{t,i} = 1\}$$

The Poisson Process $N_{t,i}$ takes value 1 if member i dies before t, and value 0 if member i is still alive at time t.
Recall that the Pooled Annuity Fund proposed by Stamos assumes homogeneous investors: the hazard rates $\lambda_{t,i}$ are then equal for all pool members.
For the sake of simplicity the paper assumes that individual hazard rates evolve deterministically according to the Gompertz-Makeham mortality law:

$$\lambda_{t,i} = \lambda_t = \frac{1}{b} e^{(t-m)/b}$$

where m denotes the modal time of death and b is a dispersion parameter.
The Gompertz-Makeham law of mortality constitutes the standard mortality law in population models since it is parsimonious to handle, as only two parameters are to be estimated. Further, it captures empirical mortality rates remarkably well.
In order to determine the total evolution of the pool's population, let us define N_t as

$$N_t = \sum_{i=1}^{L_0} N_{t,i}$$

which represents the number of dead members up to time t. N_t is the sum of Poisson processes, therefore it is itself a Poisson Process with intensity parameter equal to the sum of the single intensity parameters of the processes $N_{t,i}$.
The number of living member instead, follows from the above assumptions and is given by

$$L_t = L_0 - \sum_{i=1}^{L_0} N_{t,i}$$

It is really important to have in mind the population evolution of the fund: in each time interval of length dt, such as $(t-dt; t)$, a random amount of members

$dN_t = N_{t-dt} - N_t$ dies. The wealth owned by dN_t dead members has to be distributed among survivor members L_t. Hence, it is only by knowing accurately and precisely the evolution of the fund's population that it is possible to determine the wealth dynamic of investors, and therefore to maximize their utility.

Financial markets

Another important assumption that has to be made in order to define a Pooled Annuity Fund regards financial markets. In fact, the wealth dynamic of single investors depends also on returns from investments. Members of the PAF can invest their wealth into one risky asset and into one riskless asset. They can decide the riskiness of their portfolio and choose their optimal asset allocation strategy.
The price dynamic of the riskless asset is given by the standard formula

$$dB_t = B_t r dt$$

where r is the locally and globally riskless interest rate.
The price dynamic of the risky asset is instead given by a standard geometric Brownian motion:

$$dS_t = S_t[\mu dt + \sigma dZ_t]$$

where dZ_t is the increment of a one-dimension standard Brownian motion, μ is the drift component and σ is the volatility of the process.

Wealth Dynamic

At time $t = 0$ each investor pools his wealth $W_{i,0}$ inside the Pooled Annuity Fund. Recall that the assumption of the PAF requires homogeneous investors, meaning that each member of the fund invests the same amount of money within it.

$$W_{i,0} = W_0 \;\; \forall i \in \{1, ..., L_0\}$$

The total initial value of the Pooled Annuity Fund is therefore given by

$$W_{PAF,0} = L_0 \cdot W_0$$

The value of the PAF evolves then according to:

$$dW_{PAF,t} = W_{PAF,t}[r + \pi_t(\mu - r) - c_t]dt + \pi_t W_{PAF,t} \sigma dZ_t$$

The fund is invested in financial markets; thus it earns instantaneous returns from both risky and risk-free asset, provided that π_t represents the portion of wealth allocated in the risky asset at each time t. However a proportion c_t of the total fund is withdrawn at each time t in order to reward each member of the pool. In fact, each investor receives benefit at time t according to the evolution of the optimal withdrawal rates. c_t represents the percentage of fund

collected by each investor at time t and give us information about people's consumption behavior.
Providing that each investor has the same initial endowment W_0, the fraction of wealth owned by each investor at time t is given by

$$W_{i,t} = \frac{W_{PAF,t}}{L_t}$$

The total Pooled Annuity Fund is split in equal shares among those who are still alive at time t. However, we can also define the dynamic of single investors' wealth in this way:

$$dW_{i,t} = W_{i,t} - W_{i,t^-} = W_{i,t^-}[r + \pi_t(\mu - r) - c_t]dt + \pi_t W_{i,t^-}\sigma dZ_t + \frac{1}{L_t}dN_t W_{i,t^-} \tag{2.26}$$

with

$$\tau_i > t \ \ and \ \ L_t > 1$$

Anytime a pool member dies, his remaining wealth W_{i,t^-} is reallocated among survivors L_t. The wealth dynamic of each individual is given by gains earned from financial markets plus gains earned from mortality credits minus the amount of fund withdrawn.
Equation (2.26) give us a clear representation of the Pooled Annuity Fund: each investor receives benefits from financial gains but also additional returns due to the sharing mechanism of the tontine. Further, a portion c_t of the fund is withdrawn.

- $W_{i,t^-}[r + \pi_t(\mu - r)]dt + \pi_t W_{i,t^-}\sigma dZ_t$ represents gains from both the risky and riskless asset.
- $\frac{1}{L_t}dN_t W_{i,t^-}$ represents mortality credits: the wealth of all the members dead in dt is split among survivors.
- $c_t W_{i,t^-}dt$ is the level of consumption.

The smaller is the size of the PAF, the higher would be mortality credits, since the wealth of dead people is reallocated among a smaller number of individuals. However, on the other hand, the smaller is the size of the PAF, the smaller is the probability that one of the other investors dies. Of course, if $L_t = 1$, the last survivors earns no mortality credits anymore. This intuition is better made clear by the calculation of the instantaneous expected mortality credit (for unit of wealth):

$$\mathbb{E}\left[\frac{1}{L_t}dN_t\right] = \lambda_t dt$$

We can notice that expected instantaneous mortality credits are independent from the pool size L_t. As already explained, the higher is the pool size, the lower would be the amount of mortality credits, but the higher would be the probability of earn additional gains. These two effects cancel out, resulting in the pool size not impacting the expected instantaneous mortality credits. The only relevant state variable is the age: the older the pool and thus the higher λ_t, the more released fund would be shared among fewer survivor making mortality credits increase.

The variable that is instead affected by the size of the pool is the variance of mortality credits.

$$Var\left[\frac{1}{L_t}dN_t\right] = \frac{\lambda_t}{L_t}dt$$

The higher the number of investors, the more predictable becomes mortality credits. The higher the pool size, the more deaths occur, so the more frequently wealth redistributions among survivors take place. On the other hand, if the number of participant in the PAF is really small, the variance of mortality credits increase: it becomes more difficult to predict times of death.

Imagine having access to a perfect insurance pool, in which the number of participants tends to infinity, such that

$$\lim_{L_t\to\infty} Var\left[\frac{1}{L_t}dN_t\right] = 0$$

In this context, the mortality risk is completely eliminated. We are dealing with a perfect pool which yields deterministic income $\lambda_t dt$ from earned mortality credits. There is no uncertainty about mortality credits. The wealth dynamic of investors is therefore restated as follows:

$$dW_{i,t} = W_{i,t^-}[r + \pi_t(\mu - r) - c_t]dt + \pi_t W_{i,t^-}\sigma dZ_t + \lambda_t W_{i,t^-}dt$$

which behaves in line with the infinite Annuity Overlay Fund presented by Donnelly et al (2014)[9].

Optimization Problem

The main goal of Stamos was to determine the optimal amount of wealth invested in risky assets at each time t and the optimal amount of consumption c_t. In order to determine these quantities, he solved a utility maximization problem. He assumed that all investors have homogeneous preferences described by a CRRA utility function:

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \quad \gamma \neq \ , \gamma > 0$$

where $C_t = c_t W_t$ is the level of consumption at time t and γ denotes the level of relative risk aversion. The problem consists in maximizing the expected discounted utility of consumption:

$$\max_{[\pi_t, c_t]} \mathbb{E}\left[\int_0^\infty e^{-\delta t} u(C_t)dt\right]$$

Subject to

$$dW_{i,t} = W_{i,t^-}[r + \pi_t(\mu - r) - c_t]dt + \pi_t W_{i,t^-}\sigma dZ_t + \frac{1}{L_t} dN_t W_{i,t^-}$$

After solving the first order condition for π_t, Stamos obtained the following optimal portfolio policy

$$\pi_t = \frac{\mu - r}{\gamma\sigma^2} \quad \forall t \tag{2.27}$$

It is clear that the optimal asset allocation strategy results in a constant mix rule: the portion of wealth invested in risky asset is constant over time and it is independent of wealth dynamic and pool size.
In order to derive the optimal consumption strategy, a more complicated approach is required.
We first define the value function $V(w, l, t)$, as the value assumed by the expected discounted utility function at time t, when $W_t = w$, $L_t = l$ and when the optimal state variables - optimal π_t and c_t - are plugged in.
This value function is assumed to have the following form:

$$V(w, l, t) = f(l, t)\frac{w^{1-\gamma}}{1-\gamma}$$

$$\text{with } \lim_{t\to\infty} f(l, t) = 0 \text{ and } f(0, t) = 0.$$

Then, the optimal consumption policy is obtained by

$$c_t = f(l, t)^{-\frac{1}{\gamma}} \tag{2.28}$$

Theorem. *$f(l, t)$ that solves the first-order condition for c_t and is able to derive the optimal consumption path, must obey the following set of ordinary differential equations:*

$$\frac{f_t(l, t)}{f(l, t)} + \gamma f(l, t)^{-1/\gamma} + (A - \lambda_t l) + \lambda_t(l - 1)\left(\frac{l}{l-1}\right)^{1-\gamma}\frac{f(l-1, t)}{f(l, t)} = 0 \tag{2.29}$$

where $f_t(l, t)$ denotes the partial derivative with respect to t and A being a constant

$$A = (1 - \gamma)\left[r + \frac{1}{2\gamma}\left(\frac{\mu - r}{\sigma}\right)^2\right] - \delta$$

Analytical solutions to the set of ODEs can only be derived for the two extreme cases $l = 1$ and $l = \infty$.

When $l = 1$, only one investor lives - either because no other investors are there or because all other investors have already perished. It can be verified that

$$f(1, t) = \left[\int_t^\infty e^{\frac{1}{\gamma}\int_t^s (A - \lambda_u)du} ds\right]^\gamma \tag{2.30}$$

Thus the optimal consumption policy becomes

$$c(1,t) = \left[\int_t^\infty e^{\frac{1}{\gamma}\int_t^s (A-\lambda_u)du} ds\right]^{-1} \tag{2.31}$$

It can be seen that the optimal consumption rate increases over time: the present value of consumption becomes more valuable as the individual gets older, hence as hazard rates increase. The higher is the probability of death, the more people wish to withdraw higher amounts.

When instead $l = \infty$ we derive the function $f(l,t)$ by maximizing the expected discounted utility of consumption and assuming the wealth dynamic of a perfect pool. It can be verified that

$$f(\infty,t) = \left[\int_t^\infty e^{\frac{1}{\gamma}\int_t^s (A-\gamma\lambda_u)du} ds\right]^{\gamma} \tag{2.32}$$

Thus the optimal consumption policy becomes

$$c(\infty,t) = \left[\int_t^\infty e^{\frac{1}{\gamma}\int_t^s (A-\gamma\lambda_u)du} ds\right]^{-1} \tag{2.33}$$

Also in this case, if we assume $\gamma > 1$, the optimal consumption rate increases over time. However, the higher the number of pool participant is, the higher are withdrawal rates, since more pool members benefit from mortality credits. Thus, if we compare $c(\infty,t)$ and $c(1,t)$ we can notice that $c(\infty,t) > c(1,t)$ almost surely.
Recall that the higher the number of participants, the lower the variance of mortality credits. When $l = \infty$, the mortality risk is completely eliminated, leading members to obtain deterministic returns and thus increase their benefits. People do not need to preserve their capital, since they already know that they will receive extra income from deceasing pool members.
On the other hand, when $l = 1$, people cannot afford do withdraw high amounts, since they will not receive extra money from other deceasing members. The last survivor does not receive mortality credits, so compared with a very large pool, the consumption amount withdrawn is much lower.

As already mentioned, analytical solutions are not available when $1 < l < \infty$. The set of ODEs (2.29) has to be solved numerically by using finite difference methods. The function $f(l,t)$ is indeed approximated by a step function which takes values $\hat{f}_{l,i}$ at each time $t = i\Delta t$ and the partial derivative $f_t(l,t)$ is appoximated by a forward differential quotient. The new set of ODEs becomes thus:

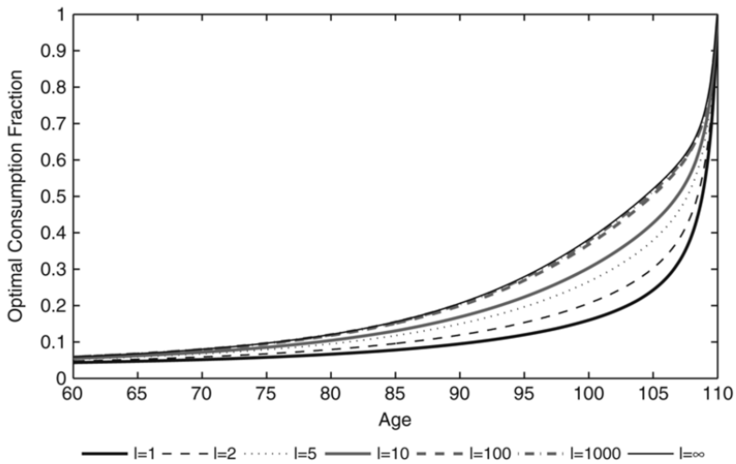
$$\frac{\hat{f}_{l,i+1}\hat{f}_{l,i}}{\hat{f}_{l,i}\Delta t} + \gamma\hat{f}_{l,i}^{-1/\gamma} + (A-\lambda_t l) + \lambda_t(l-1)\left(\frac{l}{l-1}\right)^{1-\gamma}\frac{\hat{f}_{l-1,i}}{\hat{f}_{l,i}} = 0$$

This has to be solved recursively. Then $\hat{f}_{l,i}$ is used to compute the approximate solution of the optimal consumption strategy

$$\hat{c}_i = \hat{f}_{l,i}^{-1/\gamma}$$

Figure 2.6 shows the impact of pooling on the optimal withdrawal decision. It can be clearly seen that the higher the number of members in the PAF is, the higher are consumption rates, because the more pool members can benefit from the mortality credit. The idea is that the higher are mortality credits, the more people can afford to withdraw higher amounts during their entire lifespan.

Figure 2.6: Optimal consumption rates of PAF



Source: Stamos(2008) [30]

Note: The graph presents the optimal descrete withdrawal function $1 - e^{-c_t}$ for different ages and population sizes. The expected instantaneous stock return is $\mu = 0.06$ and the expected instantaneous volatility is $\sigma = 0.18$. The real interest rate is set to $r = 0.02$. Under this parameterization the optimal portion invested in risky asset is $\pi = 24.69\%$.

We can notice that even when the pool size is equal to 100 or more, the optimal consumption rate is only slightly below the one of the case $l = \infty$. This indicates that individuals receive a stream of mortality credits that is very similar to the perfectly diversified pool and can afford to withdraw amounts very similar to them.

Chapter 3

Simulations of GSA plan

In this chapter I decided to focus my attention on the group self-annuitization plan: the GSA is one of the first modern tontine schemes proposed by Piggott et al. (2005) [26] and represents one of the main innovative and interesting plans suggested so far. Many other experts, such as Donnelly and Hanewald, have analyzed the GSA plan, comparing it with other products available in the insurance market and evaluating its attractiveness in relation to various contracts. The GSA is a very significant implicit tontine scheme, in which the benefits paid to retirees at each time t, are periodically adjusted taking into account investment and mortality experience. Recall that in a group self-annuitization plan, the retiree pays a single premium at inception, and receives annually payments B_t until death. The benefit received at each time t evolves according with two general rules. It can be seen as:

1. $$B_t = B_{t-1} \cdot MEA_t \cdot IRA_t$$

   Where MEA_t is the mortality experience adjustment and takes different values in the constant and varying contribution case; IRA_t is the interest rate adjustment ad it is equal to $(1 + R_t^*)/(1 + R_t)$.
   The periodic benefit of a GSA plan is determined based on the previous payment, adjusted for any deviations in mortality and interest rates from expectations. Whenever expectation of mortality rates are not met, the mortality experience adjustment MEA_t, corrects the benefit. On the other hand , in the event that actual earning rates deviate from expectations, the interest rate adjustment IRA_t, modifies the payout.

2. $$B_t = \frac{F_{i,t}}{\ddot{a}_{x+t}} + MC_t$$

   Where MC_t is the mortality credits achieved at time t.
   The benefit payout of a member in the GSA plan can also be decomposed into two parts: the first one corresponds to the annuity he would receive if at time t he paid a single premium equal to $F_{i,t}$; the second one is the mortality credit. The mortality credit represents the amount of fund of

people who died between $t-1$ and t, reallocated among members who are still alive at time t. With this representation we are able recognize the tontine structure of the GSA plan. People really pool their longevity risk and share their accumulated fund.

Considering the revolutionary features offered by a GSA plan and the great appeal of this product, I decided to devote this chapter to the illustration of practical and realistic examples. Thus, I present here some numerical representations of a GSA plan. I analyze in detail all of the insights introduced in section 2.1, which is entirely devoted to the description and explanation of this product.

I first present the assumptions made to conduct the simulations, paying particular attention to the derivation of actual survival rates.
Then, I focus my attention on the constant contribution case when an homogeneous cohort is assumed. I consider ten different types of pools, each with size equal to 100 000. The first pool consists of individuals whose initial B_0 is fixed at 100. The second pool consists of members with initial B_0 equal to 200. In the third pool, $B_0 = 300$, all the way up to an initial benefit of 1 000. For all these cases, I simulate the actual survival pattern, the evolution of benefits, mortality credits, mortality adjustments and the dynamics of funds over time. For simplicity, I the ignore interest rate risk, by assuming $R_t^* = R_t$.
Finally, I considered the varying contribution case for an homogeneous cohort. Here, I analyze a pool consisting of 100 000 members: 10 000 of them with an initial benefit B_0 fixed at 100, other 10 000 of them with initial B_0 equal to 200, other 10 000 with $B_0 = 300$, all the way up to an initial benefit of 1 000. Also in this case, I simulate the survival pattern, the evolution of benefits, mortality credits, mortality adjustments and the dynamics of funds over time for each homogeneous group and for the entire pool. Again, I ignore deviations from earning rates by assuming $R_t^* = R_t$.
Instead, I decide not to consider the case of inhomogenous cohorts, in which members of different ages are allowed to join the pool, regardless of the point in time.
I also decide to made some comparison of the GSA plan with standard annuity contracts, in order to better understand the appeal of this product.

3.1 General Assumptions

In the simulations performed in this chapter I make some important assumptions. First, I set the starting age of the cohort to 65. The individuals who join the group self-annuitization plan are all 65 years old at inception, i.e. $x = 65$ at time $t = 0$. The final age at which every members of the pool is dead is fixed to 110. The maximum time t considered is therefore 45, after that time the pool is totally exhausted. The pool size at inception l_x is fixed to 100 000: in both the constant and the varying contribution cases I set the number of members at inception to 100 000. The initial benefit payment B_0 varies across examples. For both the constant and varying contribution cases, I will explain in more detail the value assumed by B_0.

As anticipated, one of the main assumption is related to deviations in interest rates: the expected investment earning rate R is assumed to be constant over time, i.e. $R_1 = R_2 = ... = R_{45} = R$, and the actual investment rate is assumed to coincide with the expected one, i.e. $R^* = R$. As Piggott et al.(2005) [26], I do not attempt to measure the impact of deviations from returns in this illustration. Certainly, interest rate deviations have a large impact on the calculation of benefit payments, perhaps even greater than the impact resulting from mortality deviations, but in this context I preferred to focus on longevity variation effects. The level of interest rate assumed for all the simulations performed is 1%.

Another important assumption made concerns survival rates.
Expected survival rates are extrapolated from mortality tables available in the Human Mortality Database (HMD) [6]. I decide to consider an Italian mortality table based on the 1910 cohort in which the gender screened is male. This mortality table reports survival rates p_x for $x = 0, ..., 110$.
The survival probabilities p_x for $x = 65, ..., 110$ are therefore taken as the expected survival rates in my simulations.

Other important quantities that I calculate to derive the development of a GSA plan are the annuity factors. Recall that the quantities $\ddot{a}_x$ are critical and must be calculated taking into account expected survival rates and expected interest rates. In the simulation, I first derive the value of $\ddot{a}_{65}$ as:

$$\ddot{a}_{65} = \sum_{t=0}^{45} \frac{1}{(1+R)^t} \, {}_tp_{65}$$

Then, the value of the others annuity factors is derived recursively according to the formula:

$$\ddot{a}_{65+t} = \frac{(\ddot{a}_{65+t-1} - 1) \cdot (1+R)}{p_{65+t-1}}$$

As anticipated, in this illustration I attempt to measure the impact of deviations from mortality rates on the calculation of benefit payments. To do so, I first require to simulate actual survival rates p_x^*.
I decide to follow an approach very similar to the one used by Milevsky and Salisbury(2015) [19]: I therefore select a binomial model to simulate actual survival pattern. The evolution of number of survivors within the pool l_{x+t}^* is indeed simulated according to this formula:

$$l_{x+t}^* \sim Bin(l_{x+t-1}^*, p_{x+t-1})$$

Where p_{x+t-1} is taken from the mortality table and l^*_{x+t-1} comes from the simulation at the previous stage. Once the actual number of survivors in the pool has been simulated for each time t, it is possible to derive the actual survival probabilities:

$$p^*_{x+t-1} = \frac{l^*_{x+t}}{l^*_{x+t-1}}$$

For each simulation performed I therefore compute the actual number of survivors l^*_{x+t} at each time $t = 0, ..., 45$. Starting from the initial pool size $l_x = 100\,000$, I simulated l^*_{x+1} from a binomial distribution with parameters l_x and p_x. Then, exploiting the knowledge of the actual number of survivors l^*_{x+1} I derive the number of survivors at time 2. I repeat this process until the entire pattern of survivors is simulated. For each simulation computed, I then repeat all the described steps: each simulation involves a different survival scenario.

3.2 Constant Contribution case

In this section I will present some numerical results obtained simulating a constant contribution GSA plan. However, before presenting some outcome, it is better first to explain the detailed procedures that allowed me to obtain such illustrations. Recall that the variables available are:

- $x = 65$ initial age of members in the pool;
- $l_x = 100\,000$ initial pool size;
- p_x for $x = 66, ..., 110$ expected survival probability taken from the mortality table;
- $R = 1\%$ expected and actual interest rate;
- $\ddot{a}_{x+t}$ for $t = 0, ..., 45$ annuity factors calculated according to (2.1)
- B_0 initial benefit payout;
- l^*_{x+t} for $t = 0, ..., 45$ actual number of survivors at time t, simulated from a binomial distribution;
- p^*_{x+t} for $t = 0, ..., 45$ actual survival rate derived from l^*_{x+t}.

In order to obtain the evolution of benefits over time we can apply the following recursive formula:

$$B_t = B_{t-1} \cdot \frac{p_{x+t-1}}{p^*_{x+t-1}}$$

where we recognize that the ratio of expected to actual survival rates corresponds to the mortality experience adjustment at time t.

$$MEA_t = \frac{p_{x+t-1}}{p^*_{x+t-1}}$$

Another important quantity that I simulate is the mortality credit. The mortality credit has the following form:

$$MC_t = \frac{(l^*_{x+t-1} - l^*_{x+t})}{l^*_{x+t}\,\ddot{a}_{x+t}} \cdot F(t)$$

where $F(t)$ represents the individual fund owned by each participant of the GSA plan at time t. We saw that:

$$F(t) = \frac{F_t}{l^*_{x+t-1}}$$

where F_t corresponds to the total fund of the GSA plan at time t. It evolves according with the following recursive formula:[1]

$$F_t = (F_{t-1} - B_{t-1} l^*_{x+t-1})(1 + R)$$

First Simulation - $B_0 = 100$

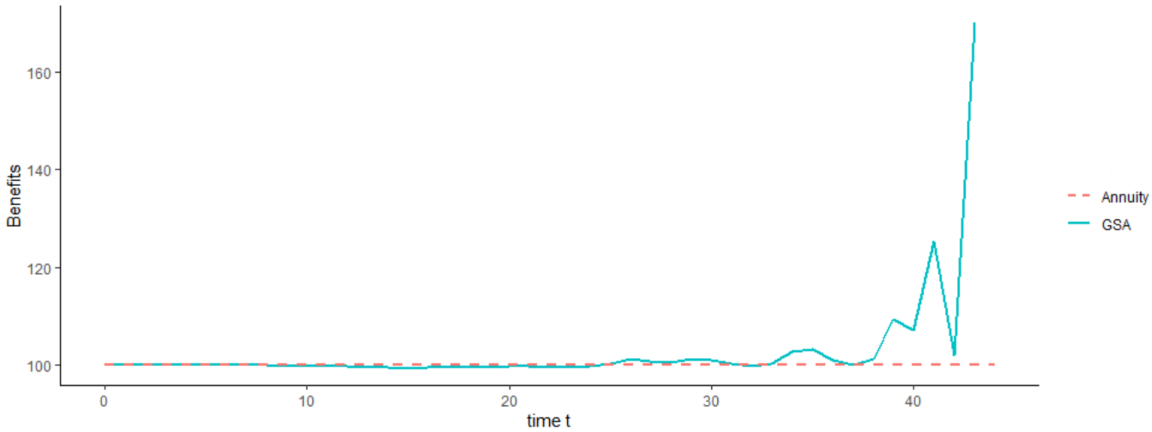
I introduce in this section one simulation realized when the initial benefit is fixed to 100. In the following table I report the evolution of the actual number of survivors in the pool, simulated through a binomial distribution.

Age x	Time t	l^*_x	Age x	Time t	l^*_x
65	0	100000	88	23	26744
66	1	97848	89	24	23002
67	2	95574	90	25	19187
68	3	93282	91	26	15719
69	4	90833	92	27	12778
70	5	88283	93	28	10114
71	6	85614	94	29	7727
72	7	82857	95	30	5810
73	8	79950	96	31	4243
74	9	77212	97	32	3020
75	10	74189	98	33	2061
76	11	70943	99	34	1338
77	12	67877	100	35	844
78	13	64687	101	36	526
79	14	61472	102	37	311
80	15	58125	103	38	171
81	16	54323	104	39	87
82	17	50516	105	40	46
83	18	46722	106	41	19
84	19	42705	107	42	10
85	20	38532	108	43	3
86	21	34559	109	44	0
87	22	30681	110	45	0

[1] Recall that the initial total fund at time $t = 0$ is $F_0 = B_0 l_x \ddot{a}_x$

We can see that the number of participants in the pool obviously decreases over time: in this simulation the pool is already exhausted at time 44, since the last survivors of the pool live up to 108 years. In order to understand actual deviations from expectations, let's have a look at the evolution of benefits. I report in the following graph the results obtained:

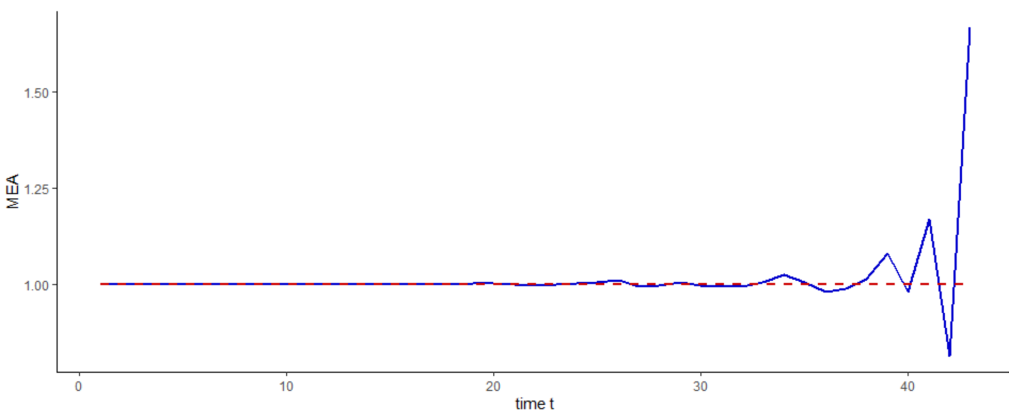
Figure 3.1: GSA benefit payouts - First Simulation whit $B_0 = 100$.



The blue line represents benefits received from a GSA plan under this simulation, while the red line represents benefits paid by an ordinary lifetime annuity with $B_0 = 100$ constant. In this case, we can see that it is convenient for individuals to enter in a group self-annuitization plan, instead of buying an annuity. Hence, it is evident that benefits paid by the GSA plan are almost always higher than 100.

The path of GSA payouts can be explained by two factors. Let us consider Figure 3.2 which reports the evolution of mortality experience adjustments MEA_t. First of all we can notice that deviations in survival rates occurs, and we can see a significant upward jump in mortality around age 41 and 43. Recall that whenever the number of members in the pool is lower than expected, the amount of benefit paid to survivors increases. We can clearly understand this phenomenon by looking at the graph:

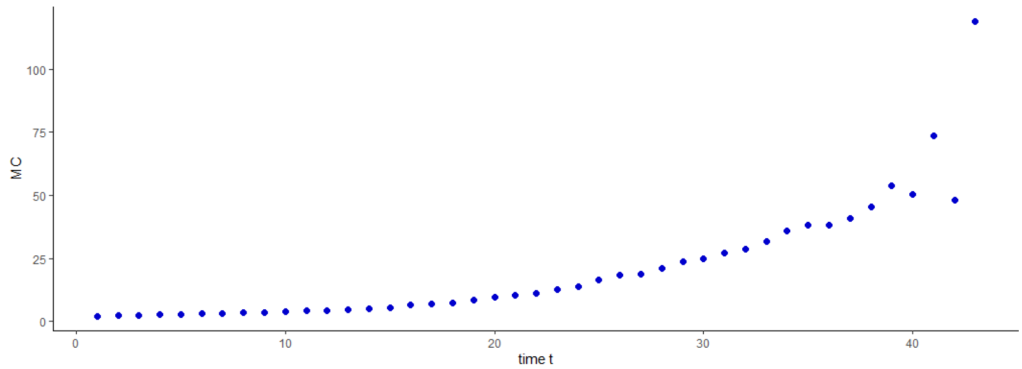
Figure 3.2: GSA mortality experience adjustment - First simulation with $B_0 = 100$.



We can see that MEA_t are almost always higher than 1. Actual survival probability are therefore lower than expected: precisely for this reason benefit payouts are higher than 100. We can also notice that any upward/downward jump in MEA_t is totally reflected in little upward/downward jump in benefit payouts B_t.

Another way to explain the payment path of the GSA plan is to examine the evolution of mortality credits. In general we expect mortality credits to increase over time, since they represents fund of dead people redistributed among individuals who are still alive. As time passes there are many more people dying, so many more funds that must be reallocated among a decreasing number of members. If we look at the evolution of survivors in the table above, we can also notice that at time $t = 40$ there are 46 survivors; in the period after, at time $t = 41$, there are instead only 19 individuals still alive. This means that the funds owned by the 27 members dead are all allocated to the pocket of the individuals still alive. Precisely for this reason, the figure below reports a jump in mortality credits at time $t = 41$. The same thought can be made for the jump that occurs at time $t = 43$. Hence, we can notice at time $t = 42$ there are 10 survivors; in the period after, at time $t = 43$, there are instead only 3 individuals still alive. Again the funds owned by the 7 members dead are all allocated to the pocket of the remaining 3 individuals still alive.

Figure 3.3: GSA mortality credits - First Simulation with $B_0 = 100$.



The steep growth of mortality credits in later years is therefore also reflected in the increasing behavior of benefit payments. As already mentioned, in this simulation, benefits paid by the GSA plan seems to be higher with respects to those paid by an ordinary annuity. In order to better compare the two contracts I calculate the expected present discounted value of both products, also called money's worth by Mitchell [21]. The EPDV is derived by summing up all the payments expected to be paid by the insurance policies, discounted back taking into account nominal interest rates. Considering that in this simulation the last benefit is paid at time $t = 43$ we can compute:

$$EPDV(GSA) = \sum_{t=0}^{43} \frac{B_t \, {}_tp_x}{(1+R)^t}$$

$$EPDV(Annuity) = B_0 \sum_{t=0}^{43} \frac{{}_tp_x}{(1+R)^t}$$

The value obtained in this simulation are reported in the following table:

	GSA	Annuity
EPDV	3040.48	3000.72

It is again evident that, in this precise simulation, individuals would prefer the GSA plan to an annuity. The expected present discounted value of the GSA plan is better than the annuity by 1.34%.
However, let us consider in the next section a new simulation of survival pattern. Thus, it is very critical to investigate whether systematically and consistently the GSA plan outperforms the annuity product, or whether it was just a casual event for this scenario.

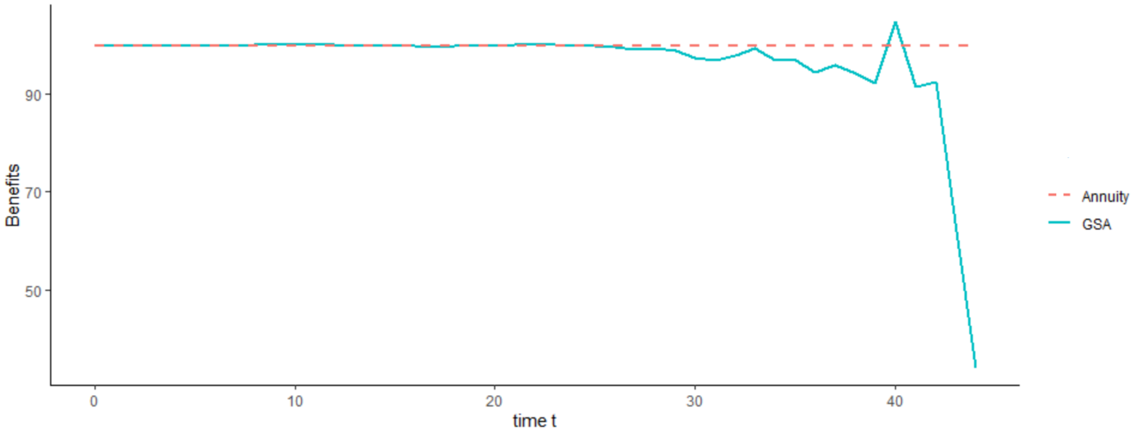
Second Simulation - $B_0 = 100$

I consider again a GSA plan in which the initial benefit payout B_0 is fixed to 100 and the actual number of members in the pool is simulated from a binomial distribution. The values of the 'new' l_x^* simulated are summarized in the following table:

Age x	Time t	l_x^*	Age x	Time t	l_x^*
65	0	100000	88	23	26544
66	1	97923	89	24	22977
67	2	95698	90	25	19250
68	3	93427	91	26	15989
69	4	90959	92	27	12972
70	5	88390	93	28	10225
71	6	85660	94	29	7912
72	7	82757	95	30	6014
73	8	79843	96	31	4384
74	9	76983	97	32	3077
75	10	73833	98	33	2076
76	11	70654	99	34	1417
77	12	67596	100	35	898
78	13	64405	101	36	563
79	14	61165	102	37	324
80	15	57801	103	38	184
81	16	54090	104	39	103
82	17	50382	105	40	47
83	18	46454	106	41	26
84	19	42529	107	42	11
85	20	38410	108	43	8
86	21	34387	109	44	5
87	22	30448	110	45	0

We can immediately recognize that the survival pattern generated is totally different with respect to the one presented in the previous section. Here the last survivor lives up to 109 years and the pool exhausts at time 45. It is already evident that this survival model is more optimistic than the previous one: the number of survivors is almost always higher than before. I report in the following graph the evolution of benefits obtained in order to get a better idea of the GSA plan generated:

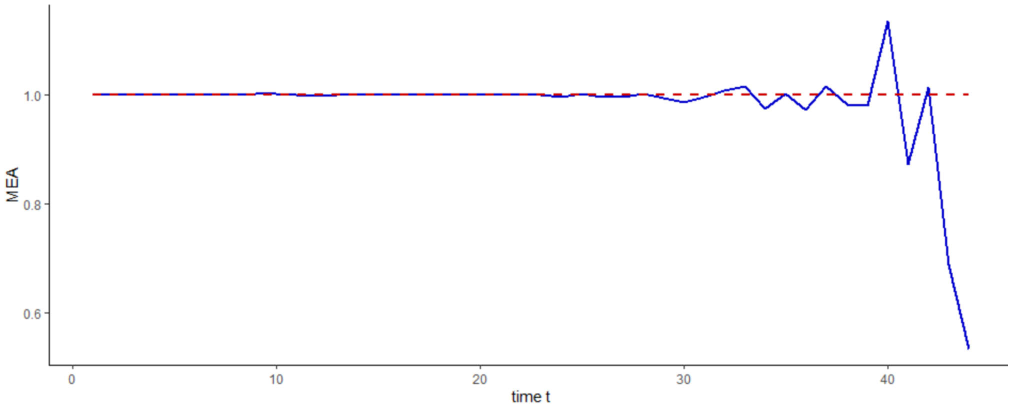
Figure 3.4: GSA benefit payouts - Second simulation with $B_0 = 100$.



Also in this case the blue line represents benefits paid by the GSA plan for this simulation, while the red line corresponds to benefits paid by an ordinary lifetime annuity with $B_0 = 100$. Now the figure is completely reversed: payouts received from a GSA plan are almost always lower than 100. In this case, an individual would prefer to buy an annuity instead of entering into a group self-annuitization plan.

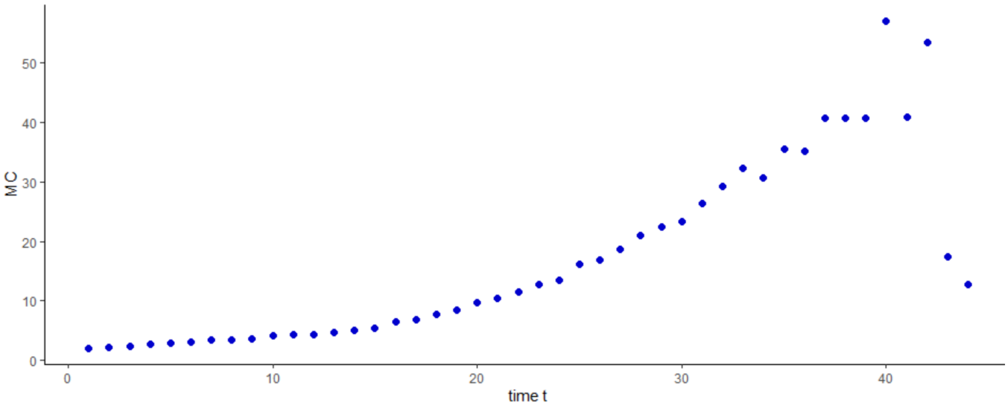
If we look at the mortality experience adjustments reported in figure 3.5, is indeed evident that actual survival rates are almost always higher than expected, i.e. $MEA_t < 1$. Recall that whenever the amount of survivors in the pool is higher than expected, the amount of benefits paid to survivors decreases. In fact, the total fund must be distributed among a larger number of individuals than expected.

Figure 3.5: GSA mortality adjustments - Second Simulation with $B_0 = 100$.



Due to the high number of survivors in the pool, the amount of mortality credits is not so large. If we look at the figure below, which illustrates the evolution of mortality credits over time, we can see that their dynamics do not show a strong increase, especially in final years. The number of survivors is so high that the funds distributed are not very significant. In fact, if we compare the mortality credits of this simulation, with the previous one, we can see that the scale of values is totally different. Here the maximum value of MC_t taken is 60, while before it was 130.

Figure 3.6: GSA mortality credits - Second Simulation $B_0 = 100$.



The evolution of mortality credits, which is not as increasing as before, is therefore reflected in the evolution of benefits paid. As already mentioned, in this simulation benefits paid by the GSA plan seems to be lower with respects to those paid by an ordinary lifetime annuity. In order to better compare the two products, I calculate again the expected present discounted value of the two. Now, considering that the last payments occurs at time $t = 44$, I compute:

$$EPDV(GSA) = \sum_{t=0}^{44} \frac{B_t \, {}_tp_x}{(1+R)^t}$$

$$EPDV(Annuity) = B_0 \sum_{t=0}^{44} \frac{{}_tp_x}{(1+R)^t}$$

The value obtained in this simulation are reported in the following table:

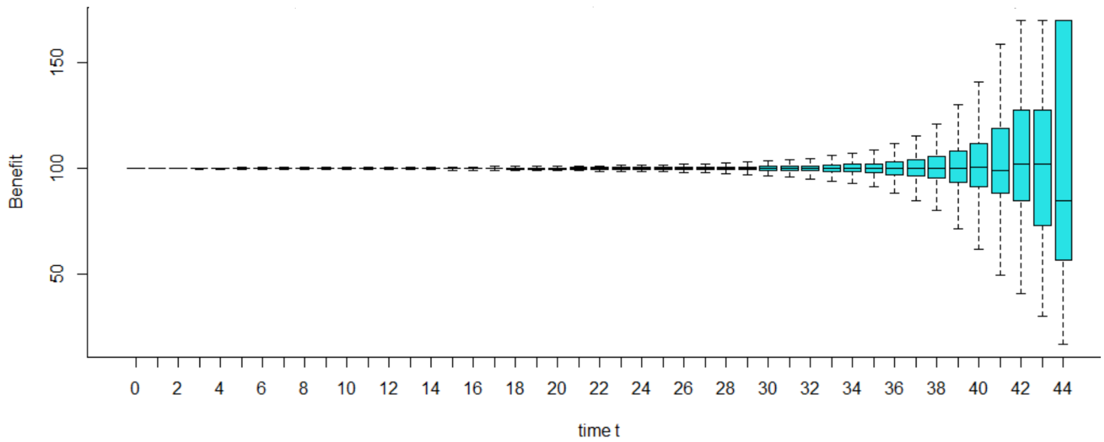
	GSA	Annuity
EPDV	2975.99	3022.23

It is again evident that, in this precise simulation, individuals would prefer the annuity to the GSA plan. Thus, the expected present discounted value of the GSA plan is worse than the annuity by 1.52%.
We obtained two completely different simulations: in one, individuals would prefer the GSA plan, while in the other, annuity would be better. The only way to actually learn if one product is better than the other statistically and significantly is to perform a large number of simulations in order to capture the overall evolution.

General trend - B_0 =100

In order to investigate the real-life attractiveness of a GSA plan relative to an annuity, I run 100 000 simulations.
I repeat the same steps described in the previous sections 100 000 times and I report in the following chart the general trend of GSA benefits when B_0 is fixed to 100.

Figure 3.7: GSA benefit payouts - General trend with $B_0 = 100$.



The graph presented above is a box-plot. Box-plots are very effective methods for fully understanding the distribution of variables and give a good indication of how values are spread out. They display five numbers:[2]

- The median: 50^{th} percentile of the dataset;
- The 25^{th} percentile;
- The 75^{th} percentile;
- The maximum value;
- The minimum value.

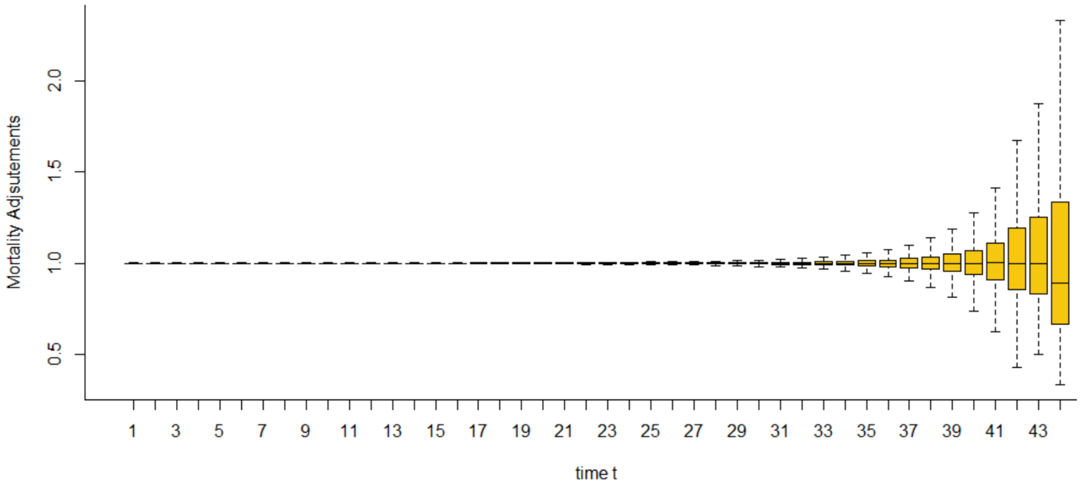
It is evident that the benefits of a GSA plan are fairly evenly distributed around 100. At early ages, variability in payments is minimal; at older ages, however, it is clear that variance in payments increases. Thus, it is not possible to state definitively that a GSA plan is preferable to an annuity with $B_0 = 100$. Up to $t = 33$, i.e. $x = 98$, the two products are very similar, whereas when $t > 34$ there are cases where one is really preferable to the other and other cases where the pattern is totally reversed.

The evolution of benefits is totally linked to the value of mortality experience adjustments. In fact, we know that whenever actual survival probabilities are

[2] Usually also outliers are visualized in boxplots, but I decided not to show them in this diagram since they were irrelevant for the main focus of the search.

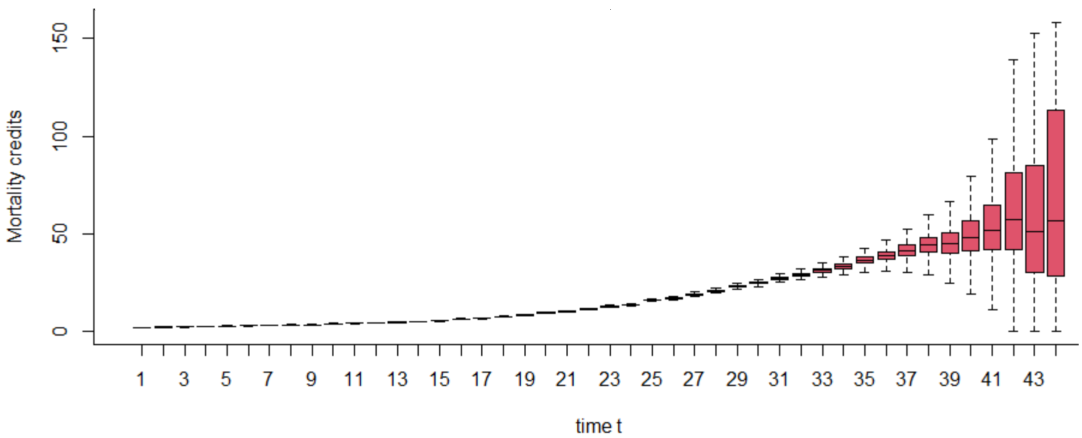
higher than expected, benefits decrease and whenever survival probabilities are lower than expected, benefits increase.
Looking at the box plot below, which shows the distribution of MEA_t, it is clear that the values are uniformly distributed around 1. They are neither statistically greater nor less than 1: again, it is not possible to say a priori and with certainty that a GSA plan is preferable to an annuity. Everything depends on the survivors pattern.

Figure 3.8: GSA mortality adjustments - General trend with $B_0 = 100$.



For the sake of completeness, I also report the trend of mortality credits in the next graph. As expected, it is clear that mortality credits increase over time and that their variability rises substantially in later years. Regardless of whether or not a GSA plan is preferable to a life annuity, the evolution of MC_t over time is increasing. The idea is that more and more people die over time, so an increasing amount of funds must be redistributed among those who are still alive.

Figure 3.9: GSA mortality credits - General trend with $B_0 = 100$.

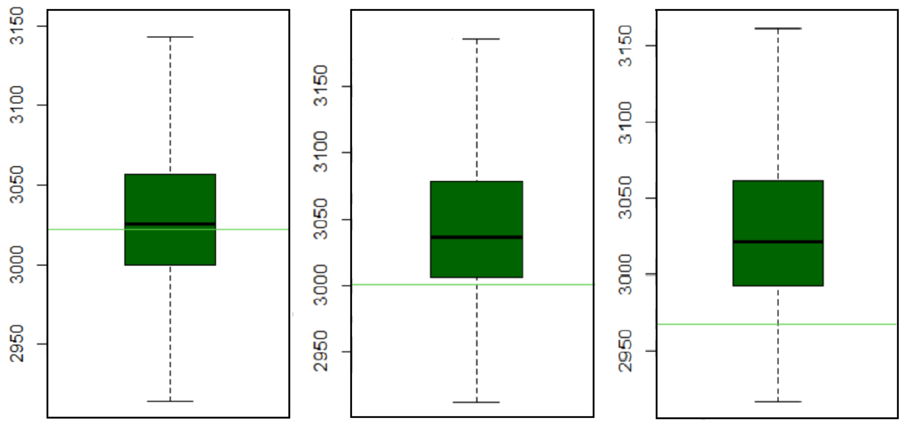


Finally, to quantitatively compare a constant contribution GSA plan and an annuity with constant benefit equal to 100, I also simulate the expected discounted present value (EPDV), i.e. the money's worth described by Mitchell (1999) [21]. I consider 3 different cases:

- Simulations in which all individuals are dead at age 109, as the second simulation presented. In these scenarios, payments continue until time $t = 44$, so the EPDV must be compared with that of the annuity in which payments continue until time $t = 44$, i.e. with the value of 3022.23.
- Simulations in which the pool is already depleted when individuals are 108 years old, as the first simulation presented. Here, payments continue until time $t = 43$. The EPDV benchmark level of the annuity is now 3000.72.
- Simulations in which all individuals die at age 107. In this rare simulations payments continue until time $t = 42$ and the baseline EPDV of the annuity is 2968.12.

I report in the following graph the box-plots concerning EPDV in the three different cases explained. The thin green line represents the EPDV of annuities.

Figure 3.10: EPDV - General trend with $B_0 = 100$.



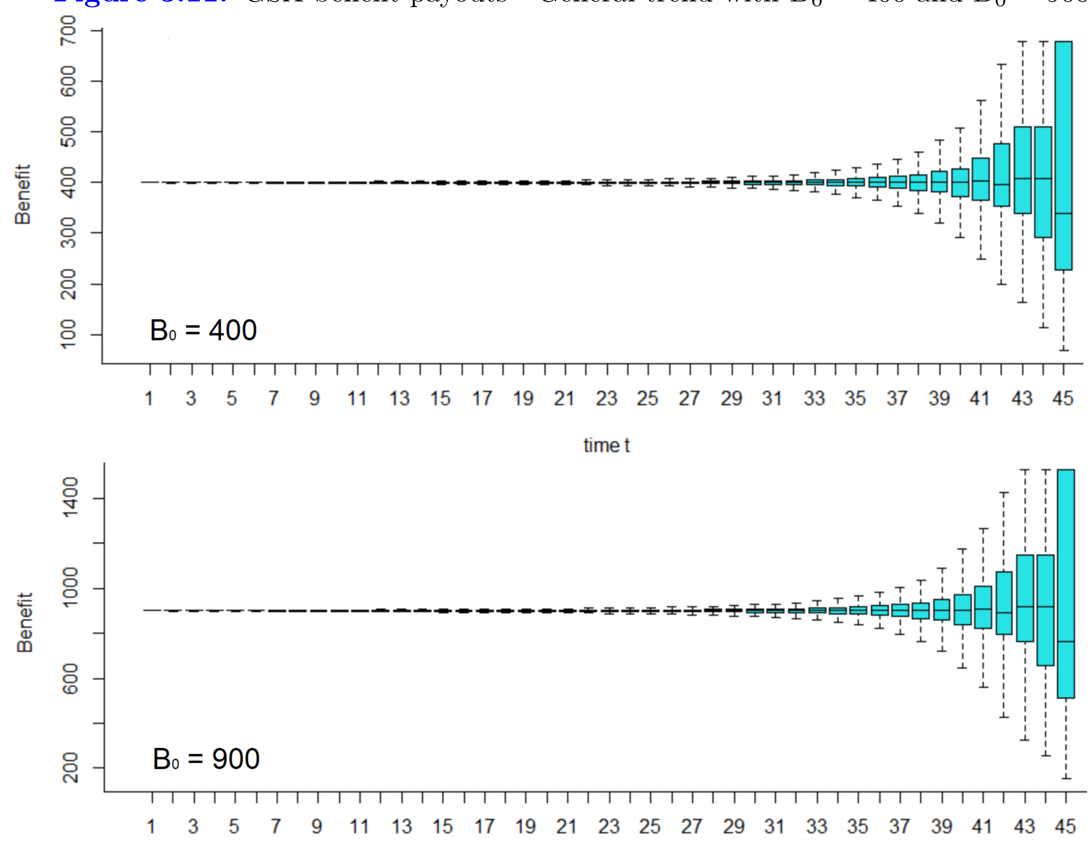
Note: Starting from the left, the first graph shows the simulations when the payments continue until time $t = 44$, the second when t=43, and the third when t=42.

It is clear that in cases where the survival pattern is standard, as expected, and at least one member of the pool survives to 109 years, the GSA plan offers a money worth in line with that of the annuity. These scenarios represent 81.71% of the simulations performed. On the other hand, when mortality increases, so people die earlier, the GSA plan is preferable to an annuity. The payments to which individuals are subject are greater, and therefore the money worth of these products is larger. As expected, only when survival rates drop, the GSA plan outperforms the annuity: hence, the amount of benefit paid to survivors increase with respect to the initial B_0 fixed. However, these examples represent a small percentage of the simulations performed: in fact, in 18% of the scenarios, payments continue until time $t = 43$, and only in 0.29% of the cases payments continue until time $t = 42$.

Overall trend - General B_0

I repeated all the simulations presented so far assuming different levels of initial payout. In fact, I set B_0 at 200, 300, 400, 500, 600, 700, 800, 900 and 1 000. I found that even for levels of B_0 different from 100, the general behavior of a GSA plan is always the same. It is not possible to state that in general annuities are better than the tontine proposed by Piggott, but neither the opposite. For simplicity, I report in the following graphs the benefits trend of a GSA plan when $B_0 = 400$ and $B_0 = 900$. The overall pattern is consistent: there are cases where benefit payouts are greater in the GSA plan, and other cases where the annuity pays higher benefits. Obviously, the scale of values changes: for $B_0 = 400$, the values are distributed evenly around 400, while for $B_0 = 900$, the payouts are distributed around 900.

Figure 3.11: GSA benefit payouts - General trend with $B_0 = 400$ and $B_0 = 900$



I also simulate EPDVs for all other levels of initial B_0 and the same reasoning as before applies. When mortality is standard, the GSA plan is really in line with that of an annuity; however, when many more people die than expected, the GSA plan turns out to be better.

This is a very interesting result. The annuities considered so far are fair, and do not take into account the loading levels normally charged by insurance companies. However, we know that in general the annuities proposed by insurers are not fair: companies sell annuities to make money. In order to cope with all the commitments, expenses and risks to which they are subject, they have to charge higher premiums, or provide lower benefits. It is therefore evident that if we would consider the actual annuities available on the market, the GSA plan would be better of annuities even in the case of normal survival patterns. This kind of tontine would turn out to outperform typical annuities in nearly all the simulations.

3.3 Varying Contribution case

In this section, I will present some results obtained simulating a varying contribution GSA plan. But before, I will explain the procedure that allowed me to obtain such illustrations.
The variables assumed are:

- $x = 65$ initial age of members in the pool;
- p_x for $x = 66, ..., 110$ expected survival probability taken from the mortality table;
- $R = 1\%$ expected and actual interest rate;
- $\ddot{a}_{x+t}$ for $t = 0, ..., 45$ annuity factors calculated according to (2.1)
- $l_x = 100\,000$ initial total pool size;
- 10 homogeneous groups of investors;
- $l_{i,x}$ initial pool size of the i^{th} group. We assume $l_{i,x} = 10\,000\,\forall i = 1, ..., 10$;
- $B_{i,0}$ initial benefit payouts of single homogeneous groups.
  We assume $B_{i,0} = i \cdot 100$ for $i = 1, ..., 10$;
- B_0 initial total benefit. $B_0 = 10\,000 \sum_{i=1}^{10} B_{i,0}$;
- $l^*_{i,x+t}$ for $t = 0, ..., 45$ and $i = 1, ..., 10$ actual number of survivors at time t in the i^{th} group, simulated from a binomial distribution. The simulation procedure is exactly the same as in Section 3.2: for each homogeneous group I simulate the survival pattern independently.

In order to obtain the evolution of benefits over time for each homogeneous group of investors we can apply the following recursive formula:

$$B_{i,t} = B_{i,t-1} \frac{p_{x+t-1}}{\sum_{A_t} F_{j,t}/F_t} \text{ for } i = 1, ..., 10$$

where we recognize that

$$MEA_t = \frac{p_{x+t-1}}{\sum_{A_t} F_{j,t}/F_t}$$

The mortality experience adjustment is the same for all the groups in the GSA plan. Departures of realized from expected mortality rates results indeed in a once and for all adjustment: whenever actual and expected mortalities differ, all payments vary in the same proportion. However, In order to calculate MEA_t we must know the value of the total fund F_t and the values of single funds $F_{i,t}$ for all the individuals that are still alive at time t.

The total fund F_t evolves according to this formula:[3]

$$F_{t+1} = (F_t - B_t)(1 + R)$$

where B_t represents to the total benefit paid at time t. It corresponds to the sum of single benefits paid to all the members alive at time t.

$$B_t = \sum_{i=1}^{10} \sum_{A_t} B_{i,t} = \sum_{i=1}^{10} l^*_{i,x+t} B_{i,t}$$

The single individual funds $F_{i,t}$ instead have the following form:[4]

$$F_{i,t+1} = (\hat{F}_{i,t} - B_{i,t})(1 + R)$$

where $\hat{F}_{i,t}$ is simply derived from:

$$\hat{F}_{i,t} = B_{i,t} \ddot{a}_{x+t}$$

Another important quantity that I simulate for each homogeneous group is the mortality credit. The mortality credits MC_t are derived from funds of people who die between $t-1$ and t, taking into account the total percentage of funds held by each group.

$$MC_{i,t} = \frac{\sum_{D_{t-1}} F_{j,t} \cdot \frac{F_{i,t}}{\sum_{A_t} F_{j,t}}}{\ddot{a}_{x+t}} = \frac{\sum_{j=1}^{10} (l^*_{j,x+t-1} - l^*_{j,x+t}) F_{i,t} \cdot \frac{F_{i,t}}{\sum_{j=1}^{10} l^*_{j,x+t} F_{j,t}}}{\ddot{a}_{x+t}}$$

All simulations are done recursively: exploiting the knowledge of the variables at time $t-1$ the variables at time t are derived. At each time t I first derive the total value of the fund F_t and all the value of the individual funds $F_{i,t}$. Then I derive the mortality experience adjustment MEA_t that allows me to calculate the benefit payments for each homogeneous group, i.e. $B_{i,t}$. Once that all individual benefits are known, I derive the benefit for the entire group, i.e. B_t. Finally, I compute the mortality credits.

First Simulation

In this section I will present one simulation of a varying contribution GSA plan. The survival pattern for each homogeneous group of investor is derived from a binomial distribution. Here, the simulated values of $l^*_{i,x}$ for $i = 1, ..., 10$ are summarized:

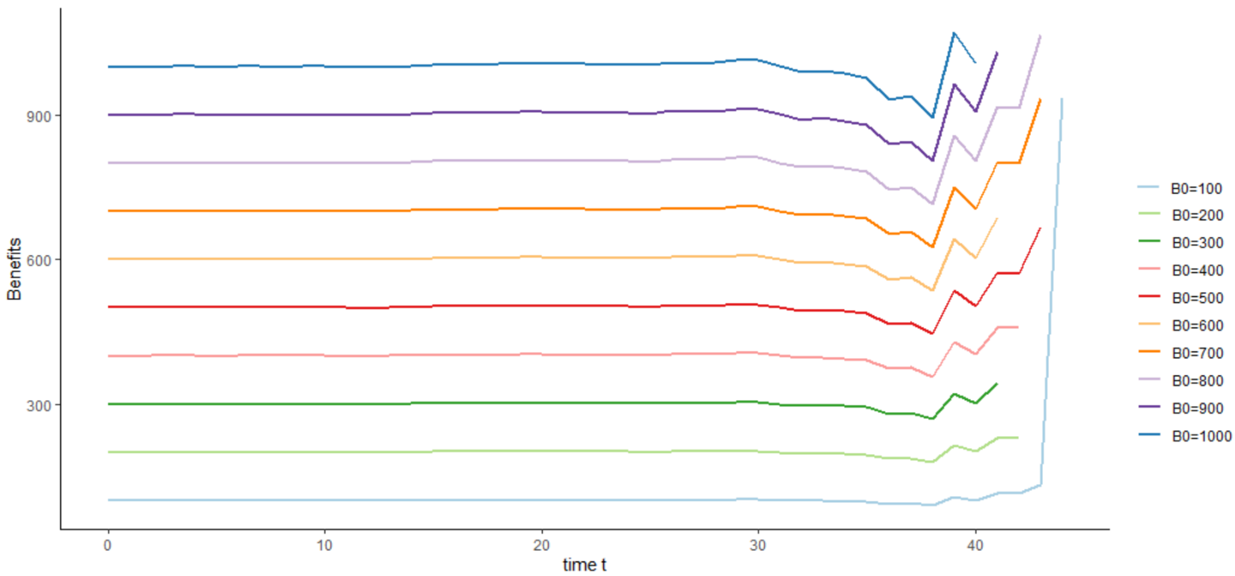
[3] The initial total fund $F_0 = 10\,000 \sum_{i=1}^{10} B_{i,0} \ddot{a}_x$

[4] The initial individual fund $F_{i,0} = 10\,000 B_{i,0} \ddot{a}_x$

t	$l^*_{1,x}$	$l^*_{2,x}$	$l^*_{3,x}$	$l^*_{4,x}$	$l^*_{5,x}$	$l^*_{6,x}$	$l^*_{7,x}$	$l^*_{8,x}$	$l^*_{9,x}$	$l^*_{10,x}$
0	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000
1	9785	9771	9777	9795	9791	9778	9815	9774	9793	9783
2	9565	9543	9534	9566	9540	9572	9590	9518	9565	9550
3	9337	9302	9289	9333	9289	9298	9344	9317	9327	9327
4	9127	9075	9037	9097	9030	9069	9087	9079	9077	9087
5	8879	8814	8803	8853	8765	8828	8845	8835	8829	8834
6	8606	8558	8507	8566	8491	8567	8580	8532	8553	8573
7	8334	8281	8208	8281	8216	8285	8295	8230	8294	8300
8	8045	7952	7918	7969	7940	8005	8014	7931	7996	8033
9	7747	7673	7623	7687	7687	7735	7736	7671	7695	7714
10	7444	7358	7337	7366	7394	7429	7445	7377	7377	7406
11	7099	7030	7028	7018	7048	7140	7172	7062	7070	7081
12	6815	6740	6730	6724	6700	6846	6851	6736	6760	6775
13	6468	6416	6410	6386	6364	6506	6533	6380	6424	6465
14	6142	6111	6066	6052	6032	6186	6173	6045	6072	6097
15	5818	5760	5720	5703	5706	5831	5811	5710	5740	5726
16	5437	5385	5348	5318	5334	5449	5433	5350	5350	5368
17	5038	5001	4973	4948	4979	5049	5059	5004	4967	4992
18	4675	4608	4589	4543	4596	4675	4655	4616	4600	4598
19	4307	4226	4235	4150	4202	4233	4253	4241	4203	4202
20	3916	3812	3808	3737	3786	3819	3837	3837	3797	3794
21	3517	3401	3391	3313	3413	3418	3441	3494	3432	3414
22	3123	3011	3015	2939	3034	3016	3011	3096	3024	3042
23	2721	2621	2658	2554	2652	2613	2614	2721	2612	2684
24	2363	2264	2289	2197	2301	2215	2288	2360	2241	2303
25	1988	1905	1921	1862	1931	1865	1930	1967	1895	1932
26	1668	1550	1583	1529	1585	1531	1592	1627	1533	1601
27	1312	1255	1285	1224	1274	1251	1283	1317	1239	1314
28	1067	1000	1041	965	996	985	1029	1021	986	1031
29	825	773	818	742	763	739	795	775	758	775
30	623	587	635	554	555	559	587	579	571	588
31	446	411	466	399	422	391	443	428	413	441
32	310	296	328	292	280	267	331	307	289	331
33	215	187	229	169	190	191	217	216	203	231
34	145	120	159	109	121	132	147	128	147	158
35	99	78	98	68	81	91	96	87	89	96
36	60	49	62	49	56	56	65	50	55	63
37	38	31	34	27	31	28	43	30	34	34
38	22	15	19	16	14	14	25	17	22	23
39	18	9	11	10	10	8	13	6	8	7
40	10	5	4	3	6	5	7	3	8	2
41	3	2	1	3	4	2	5	2	1	0
42	1	1	0	1	1	0	3	2	0	0
43	1	0	0	0	1	0	1	1	0	0
44	1	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0

We can notice that up to time $t = 40$ there is still at least one living individual in all the homogeneous groups. However, at time $t = 41$ the 10^{th} group, the one with an initial $B_0 = 1\,000$, is already exhausted. On the other hand, we can observe that the 1^{st} group is the one that lives at most: at time $t = 44$ there is only one individual still alive in the entire pool and he belongs to the first group. I report in the following graph the evolution of benefits $B_{i,t}$ for each homogeneous group, in order to get a better idea of the GSA plan generated:

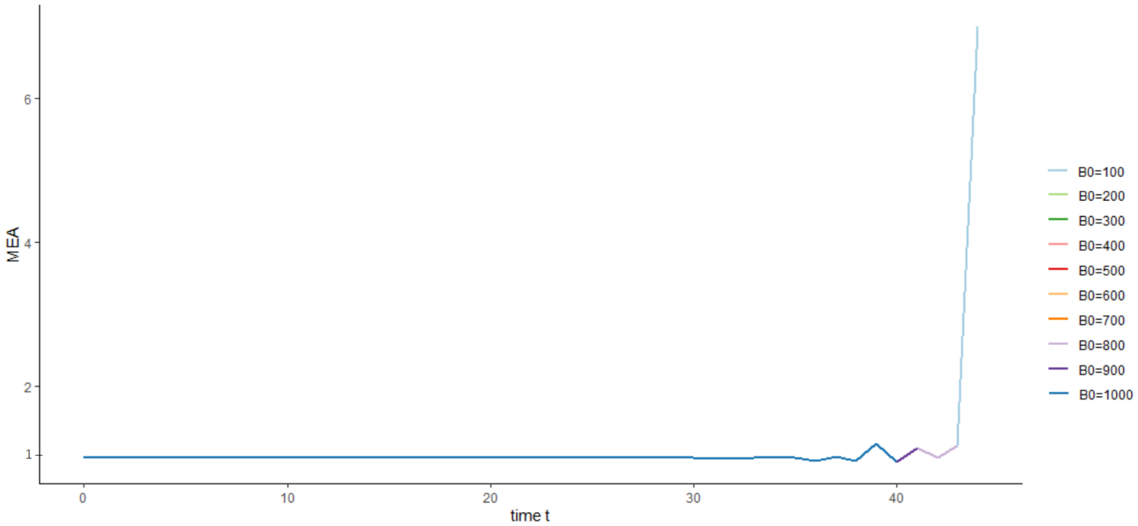
Figure 3.12: Varying contribution GSA benefit payouts - First Simulation



Each line represents the benefit pattern of single groups. The color legend is listed to the right of the chart.
We can notice that up to time $t = 40$, when there is at least one living individual in all groups, the trend is the same for everyone. All payments vary in the same proportion: when there is an increase of, say, 5% in $B_{2,t}$, all other payments also increase by the same percentage. After time $t = 40$, we can observe that the benefit trend of the 10^{th} group stops, since no one receives payments anymore , i.e. they are all dead. The same is true for all other groups, when there are no more living members the payments end. For instance, it is evident that at time $t = 41$ the benefit pattern of 3^{rd}, 6^{th} and 9^{th} groups, stops.
One of the things that immediately becomes apparent in the graph is the evolution of the benefits of the first group: we can see a significant increase in benefits for this group in the last year, a significant jump upwards. This is due to the fact that the only individual still alive at the time $t = 44$ is in the first group, so the funds of all the people who died during the previous year are distributed to him; considering that all the other groups had much larger funds than his, because they had a larger initial contribution, the portfolio of the lucky individual still alive at the time $t = 44$ grows significantly.
In general, the evolution of benefits, for only those groups in which there is someone still alive, remains consistent: payments vary in the same proportion

across groups. This is evident when we look at the evolution of mortality adjustments. For varying contribution GSA plans, the MEA_t are unique for the entire pool; there is no diversification of adjustments among groups and this is made even clearer if we look at the following figure:

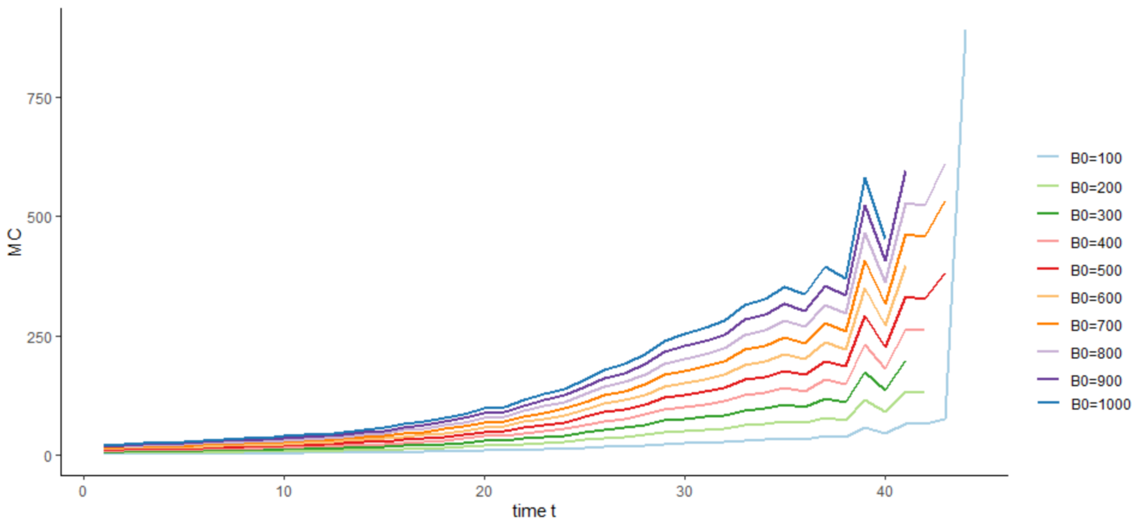
Figure 3.13: Varying contribution GSA mortality adjustments - First Simulation



We can see that the mortality experience adjustment is unique for all groups. Until time $t = 40$ the representative group still alive is the tenth, so we see painted the line of its corresponding color. At time $t = 41$ the representative group becomes the ninth so the color changes, then at time $t = 42$ turns into the eighth and finally at time $t = 44$ it becomes the first.
It is evident that MEA_t are more or less stable around the value of 1 up to $t = 43$: there are small upward jumps that are totally reflected in the behavior of the payment pattern. The strange thing happens at time $t = 44$ in correspondence of the first group, due to the reason already previously described.

Now, let's have a look at the evolution of mortality credits:

Figure 3.14: Varying contribution GSA mortality credits - First Simulation



Once again, it is clear that the overall trend across groups is consistent: mortality credits vary in the same proportion across groups.
The path of mortality credits is in general increasing, due to the growing number of people dying over time. However, the higher are funds owned by members in the pool, the higher are mortality credits received.
Further, we can clearly identify a jump in the mortality credits of the first group, which is completely related to the reason explained earlier: at time $t = 44$ the funds of people who died in group 5, 7 and 8 are completely allocated to the unique individual left alive in group 1.

From the simulation performed, it seems that entering into a varying contribution GSA plan is preferable to purchasing a life annuity for all investors considered. In fact, the benefit pattern seems to be always higher than the B_0 set at the beginning. However, in order to better understand the two products and to better compare them, I report in the following table the expected present discounted values of the two contracts:

B_0	GSA	Annuity	% Difference
100	3 222.46	3 022.23	6.62
200	5 955.80	5 936.25	0.33
300	8 836.83	8 819.72	0.19
400	11 911.60	11 872.5	0.33
500	15 107.07	15 003.61	0.69
600	17 673.67	17 639.45	0.19
700	21 149.90	21 005.05	0.69
800	24 171.31	24 005.77	0.69
900	26 510.50	26 459.17	0.19
1 000	29 087.69	29 078.05	0.03

It can be seen that in all cases presented, the value of the EPDV is greater under a GSA plan than under a life annuity. In other words, in this simulation, regardless of the initial contribution level, all individuals turn out to be better off with a GSA plan rather than with an annuity. There are cases where the difference between the two values is quite significant, as in the case of the first group, and others where instead the two values are very similar. However, in order to have a general picture of the situation, exactly as done in the case of constant contribution, I will perform in the following section another simulation.

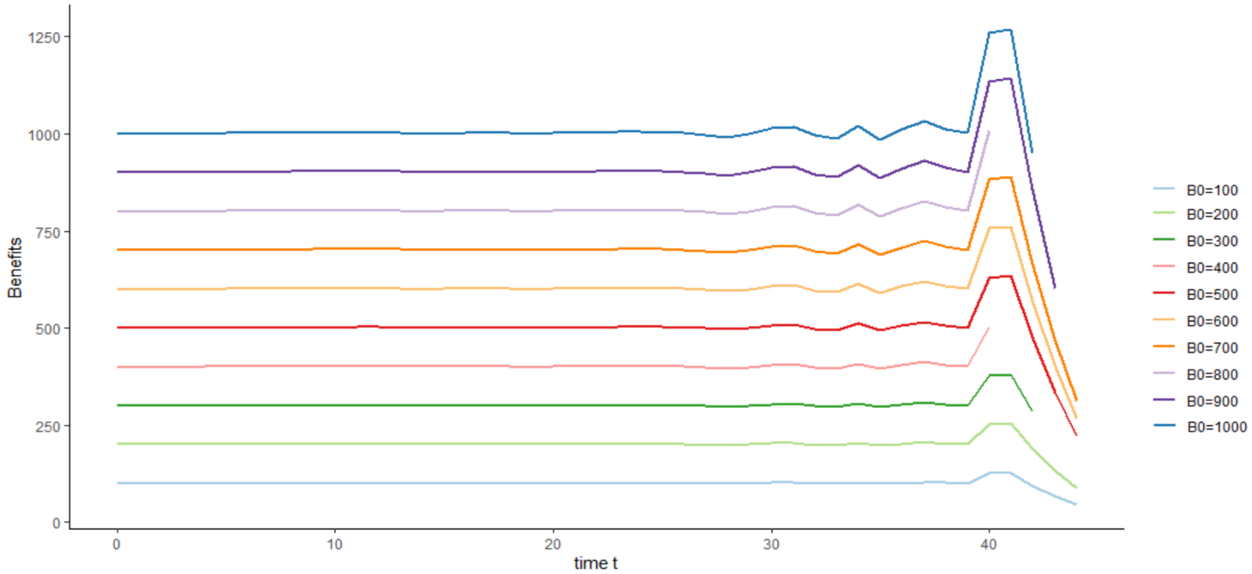
Second Simulation

In this section I will perform a new simulation of a varying contribution group self-annuitization plan. I simulate again the survival pattern $l^*_{i,x}$ for each homogeneous group of investors and I report them in the table below. I recall that all the simulations are done independently: the survival pattern of each group does not depend on the others.

t	$l^*_{1,x}$	$l^*_{2,x}$	$l^*_{3,x}$	$l^*_{4,x}$	$l^*_{5,x}$	$l^*_{6,x}$	$l^*_{7,x}$	$l^*_{8,x}$	$l^*_{9,x}$	$l^*_{10,x}$
0	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000
1	9787	9775	9783	9782	9795	9810	9778	9792	9786	9766
2	9537	9536	9532	9560	9556	9593	9568	9584	9541	9545
3	9303	9304	9294	9338	9322	9342	9337	9358	9314	9314
4	9062	9046	9038	9108	9040	9074	9091	9112	9089	9076
5	8808	8793	8767	8859	8785	8823	8822	8831	8832	8795
6	8547	8535	8494	8586	8511	8560	8523	8584	8567	8517
7	8283	8256	8226	8328	8214	8251	8228	8319	8291	8221
8	8007	7947	7963	8028	7940	7945	7913	8038	7979	7942
9	7697	7690	7681	7739	7650	7652	7640	7740	7671	7680
10	7407	7390	7387	7450	7345	7343	7345	7456	7340	7375
11	7068	7054	7022	7155	7020	7025	7039	7129	7019	7040
12	6771	6749	6719	6873	6694	6740	6697	6831	6700	6707
13	6460	6406	6381	6516	6381	6436	6395	6519	6380	6391
14	6143	6057	6058	6197	6073	6094	6080	6197	6067	6111
15	5810	5745	5711	5864	5728	5752	5744	5872	5754	5729
16	5465	5374	5347	5470	5355	5391	5364	5506	5349	5352
17	5094	5004	4977	5114	4955	5016	5004	5133	4982	4982
18	4713	4609	4607	4734	4617	4589	4581	4752	4582	4610
19	4301	4205	4226	4308	4237	4236	4209	4360	4201	4238
20	3905	3753	3786	3856	3793	3836	3793	3941	3805	3853
21	3517	3384	3377	3479	3404	3448	3401	3550	3387	3449
22	3119	2947	2972	3069	3036	3064	3052	3139	2967	3039
23	2708	2557	2594	2672	2644	2663	2657	2737	2585	2637
24	2346	2215	2244	2283	2267	2312	2306	2346	2237	2252
25	1947	1857	1901	1935	1930	1934	1935	1962	1869	1903
26	1614	1553	1608	1606	1582	1584	1621	1613	1554	1585
27	1285	1258	1304	1293	1278	1301	1291	1308	1266	1299
28	1009	987	1013	1055	1013	1014	1036	1041	1020	1025
29	784	753	784	775	764	772	809	809	775	769
30	575	561	586	588	583	564	597	588	572	566
31	407	418	430	420	407	409	439	428	411	413
32	284	298	340	294	275	296	326	294	296	311
33	200	197	225	185	190	205	235	200	213	214
34	130	135	155	123	127	134	164	116	135	130
35	86	85	102	76	85	85	99	78	86	97
36	61	58	60	47	51	47	54	42	55	60
37	38	33	34	22	33	25	32	28	32	31
38	25	17	18	13	17	15	22	12	19	18
39	12	9	8	7	9	11	11	5	8	14
40	5	4	6	1	5	4	4	1	5	5
41	3	2	4	0	3	1	2	0	1	4
42	2	1	2	0	2	1	2	0	1	1
43	1	1	0	0	2	1	2	0	1	0
44	1	1	0	0	1	1	1	0	0	0
45	0	0	0	0	0	0	0	0	0	0

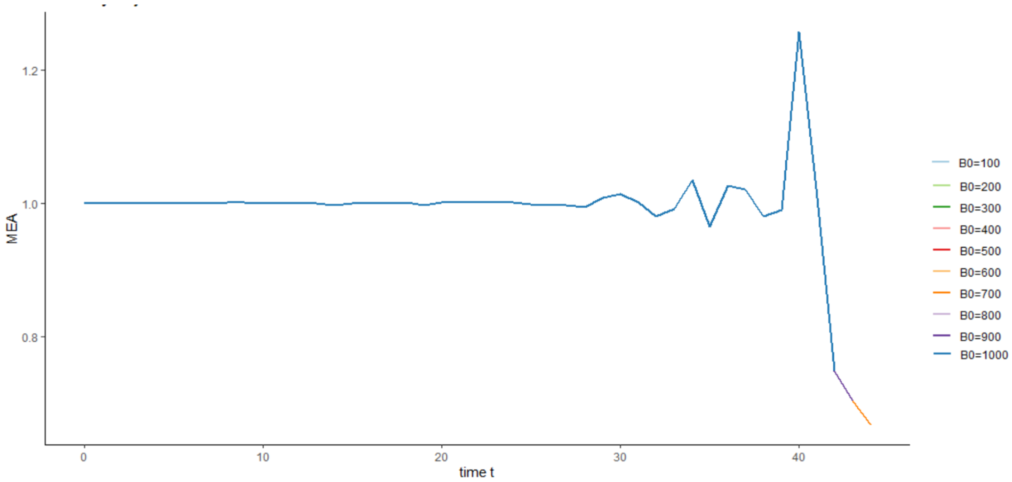
We can immediately recognize that the survival patterns generated are totally different with respect to those presented in the previous section. Up to time $t = 40$ there is at least one members in all the groups; at time $t = 41$ the 4^{th} group exhausts and at time $t = 44$ there is still at least one individual in groups 1, 2, 5, 6 and 7. It is therefore evident that these survival paths are more optimistic: the case examined before projected that at time $t = 44$ there was only one individual alive of the first group; now there are 5 members still alive at time $t = 44$, all belonging to different groups. In order to understand the behavior of this GSA plan I report here the evolution of benefit payouts:

Figure 3.15: Varying contribution GSA benefit payout - Second Simulation



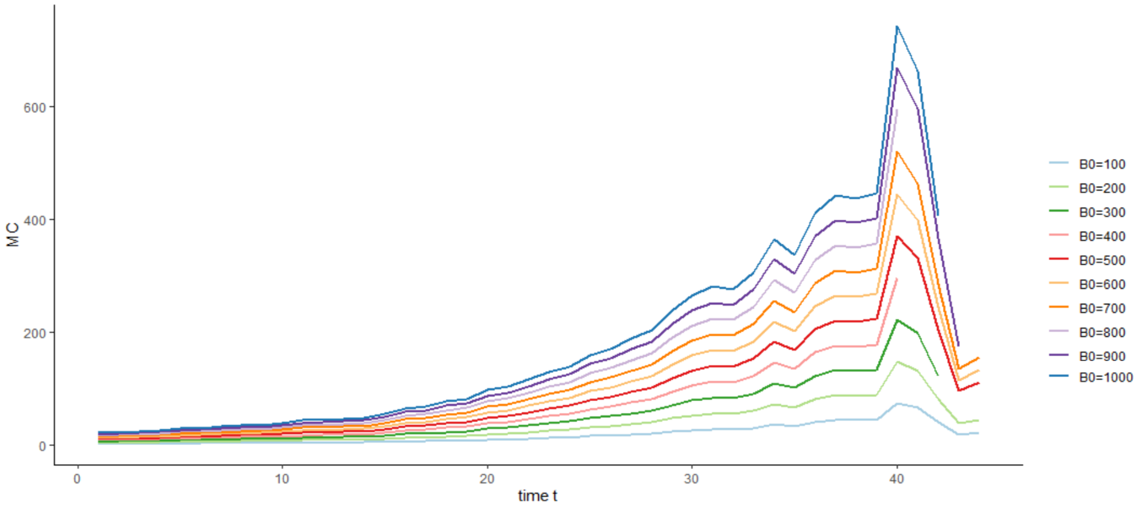
We can see again that the trend across all groups is consistent: the payments all vary in the same proportions. Further, as in the previous case, it is clear that benefits stop when there is no longer any individual alive in a group. It is evident that up to $t = 40$, the pattern is slightly increasing, providing members of the GSA plan with benefits higher than annuities. After, at time $t = 42$, instead the pattern becomes decreasing: the amount of individual still alive in the entire pool is too high, mortality credits allocated among survival members are not so significant and benefits decrease. We can emphasize this effect by looking at the evolution of mortality experience adjustments.

Figure 3.16: Varying contribution GSA mortality adjustments - Second Simulation



It is evident that after time $t = 42$, MEA_t are significantly lower than 1. Probably, as already mentioned, the amount of members alive is too high, the percentage of fund of individual still alive is higher than expected survival rates, i.e. $\frac{\frac{\sum F_{i,t}}{A_t}}{F_t} > p_{x+t-1}$.
This is made also cleared if we look at the evolution of mortality credits, which are not as increasing as before:

Figure 3.17: Varying contribution GSA mortality credits - Second Simulation



We can see that mortality credits decrease in later years. However, there are some groups which are more affected and other less. For instance, the seventh group, suffers a significant drop in mortality credits, while the first group, is not as impacted as the others. Obviously this is completely reflected in the evolution of payouts: any decline in mortality credits leads to a decline in benefits.

For the sake completeness, I also report for this simulation the expected discounted present values of the GSA plan and of the annuity.

B_0	GSA	Annuity	% Difference
100	3024.10	3022.23	0.06
200	6048.21	6044.47	0.06
300	8978.33	8904.37	0.83
400	11700.74	11631.22	0.59
500	15120.54	15111.18	0.06
600	18144.65	18133.42	0.06
700	21168.76	21155.65	0.06
800	23401.48	23262.44	0.59
900	27130.81	27006.49	0.46
1 000	29927.78	29681.26	0.83

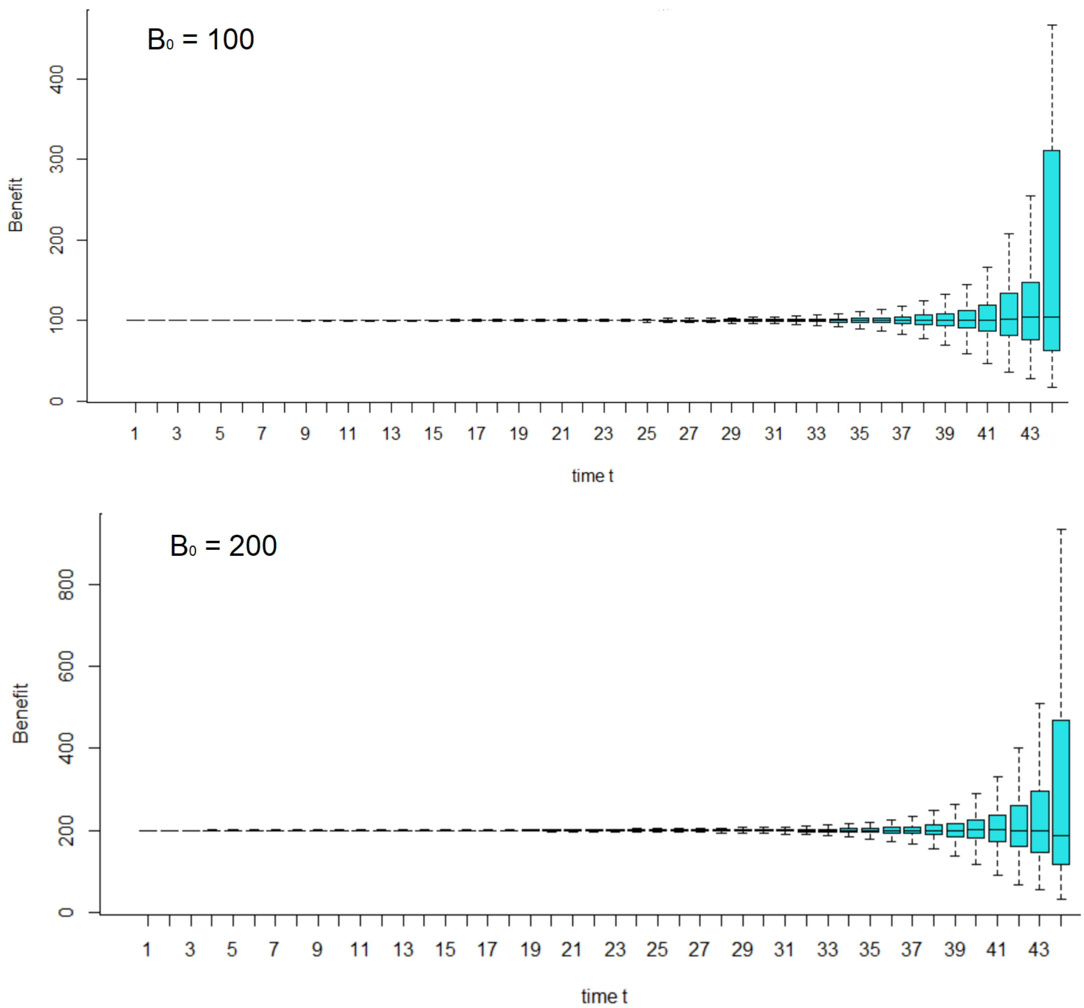
Even in this simulation, the EPDVs of the GSA plan are greater than those of annuities. Any individual, regardless of group, would prefer to enter into a

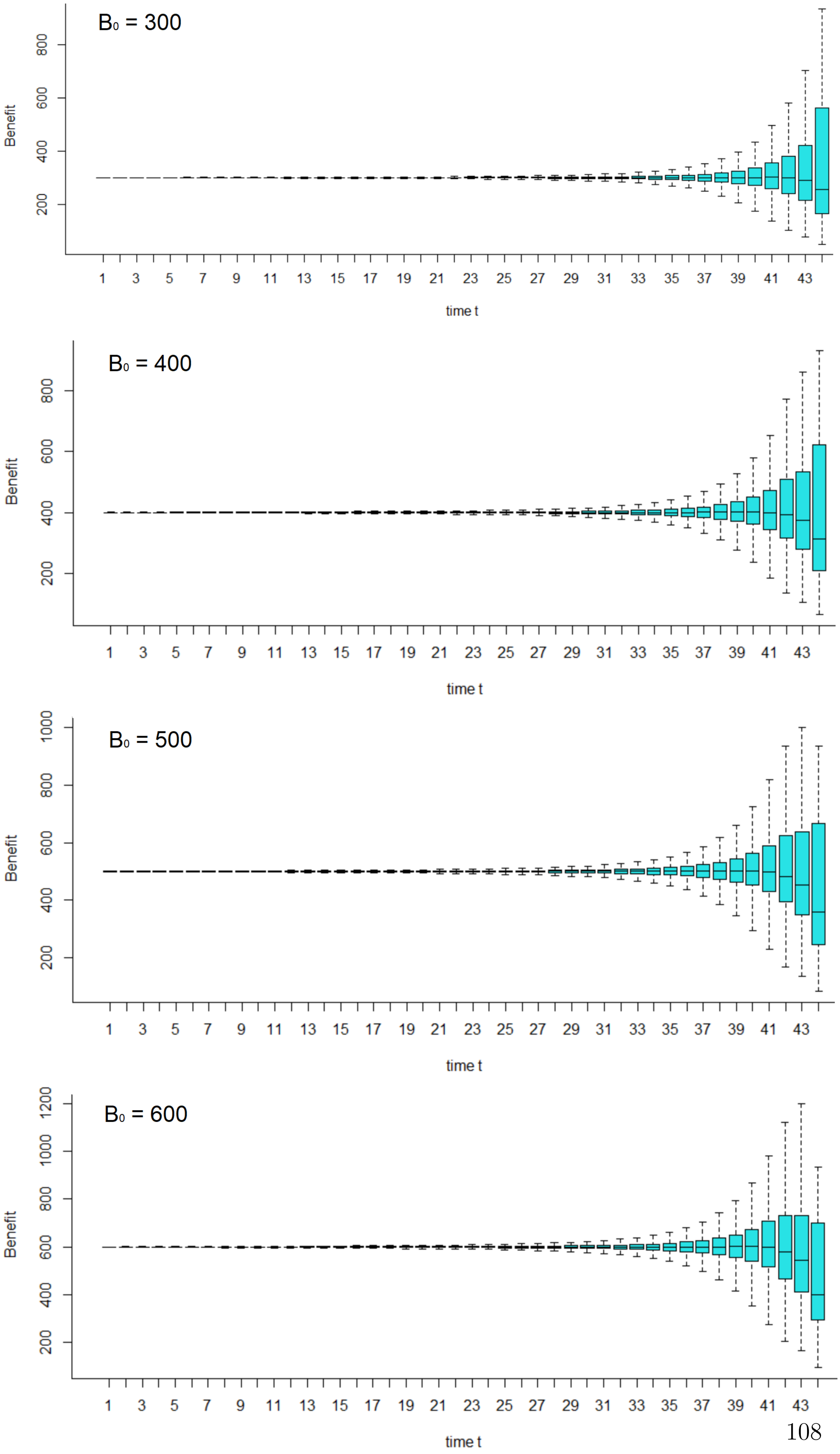
varying contribution GSA plan rather than buy an annuity.
However, this is again a specific scenario generated by this simulation, not a generic trend. In fact, we can notice the differences with respect to the annuities are minimal compared to the previous simulation. Probably if the survival pattern would have been even more optimistic, the result would have been the contrary: annuity would have been preferred to the GSA plan. In order to really understand the plan it is necessary to perform many simulations.

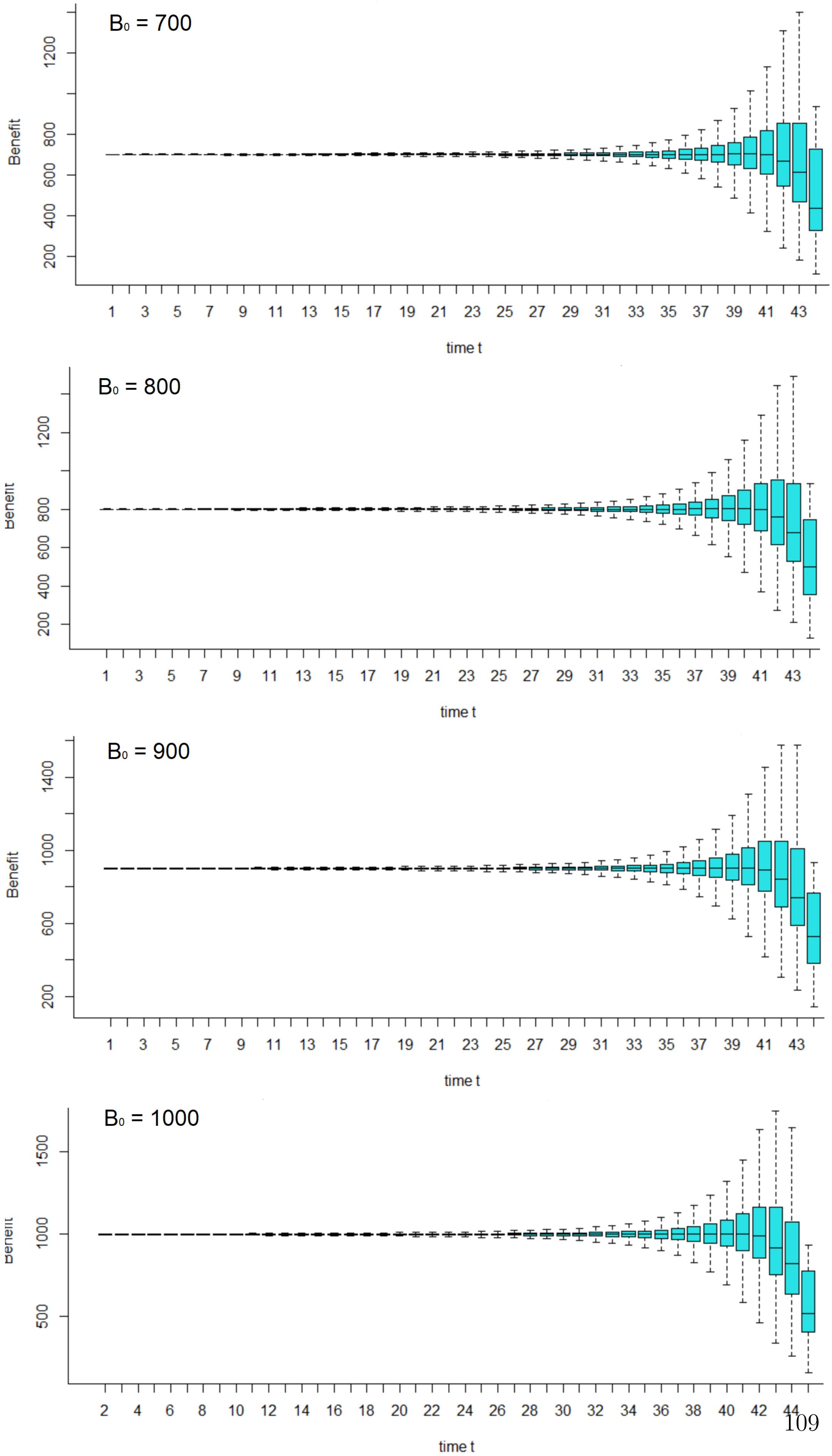
Overall Trend

In this section I report the results obtained from 100 000 simulations of a varying contribution GSA plan. I repeated all the steps presented in the previous sections 100 000 times and I collect the values into a matrix.
I first show the overall trend of benefits for each homogeneous group, through the following box-plots, and then I will comment on the outcomes.

Figure 3.18: Varying contribution GSA plan - Overall Trend




B₀ = 300
Benefit
200
400
600
800
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 400
Benefit
200
400
600
800
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 500
Benefit
200
400
600
800
1000
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 600
Benefit
200
400
600
800
1000
1200
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t


B₀ = 700
Benefit
200
400
600
800
1200
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 800
Benefit
200
400
600
800
1200
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 900
Benefit
200
600
1000
1400
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43
time t
B₀ = 1000
Benefit
500
1000
1500
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44
time t

Comparing all the above graphs there are some considerations that need to be made.
First of all it is evident that up to time $t = 40$ the overall trend of benefit payouts is more or less the same among groups. Obviously the variability of payments increases over time, but in general the consumption received by GSA members varies almost uniformly around the initial B_0. Thus, it is not possible to say with certainty whether a variable contribution GSA plan is preferable to an annuity until this time: in fact, the two products appear to be very similar. The strange thing happens when $t > 40$: we can see that for members with a small initial B_0, benefit tends to be statistically higher than the initial benefit considered. For instance, if we look at the distribution of benefits when $B_0 = 200$, we can notice that at time $t = 44$, the payouts can reach up to a level of 1 000. On the other hand, for members with large initial B_0, benefits tends to be statistically lower than the initial benefit considered. For example, if we look at the distribution of benefits when $B_0 = 900$, we can see that at time $t = 44$ the maximum level reached is 900.
We understand that for members whose initial benefit B_0 is small, it would be preferable to enter into a varying contribution GSA plan if they expect to live long. Whereas for members with a large initial benefit B_0, it would be preferable, especially if they survive a lot, not to enter into a variable contribution GSA plan.
The general idea, also explained by Donnelly (2015) [8], is that poorer members of the scheme, benefit subsidies from richer members. This scheme does not appear to be equitable among individuals: poorer people receive, in very late years, high payouts derived from richer people. The money goes from the pockets of those who are richer, to the pockets of those who are poorer. Poorer people benefit and enjoy a lot by sharing funds with richer people. If we compare the evolution of benefits when $B_0 = 100$ in the case of varying contribution and in the case of constant contribution, we understand at once how much can be convenient for poorer people to join a pool with richer people. The maximum level of payout reached in case of constant contribution, in a pool where all members contributes equal amount F_0, is only 150. Whereas the maximum level of payout reached in case of varying contribution, in a pool where all members contributes different amounts, is 600. This concept is summarized in a very precise way by a theorem stated by Donnelly(2015) [8].

Theorem. *Consider a GSA plan consisting only of two groups, A and B, with the same number of members in each group, i.e.* $l_x = l_{A,x} = l_{B,x}$ *. Each member of Group A has initial wealth* $F_0^A = 1$ *and each member of Group B has initial wealth* $F_0^B > 0$*. Fix a member in Group B, whom we call Bob.*
For $F_0^B > 1$*, i.e.* $F_0^B > F_0^A$*:*

- *Bob's expected consumption in the GSA plan is less than his initial wealth* F_0^B*;*
- *Bob's expected consumption conditional upon his survival to time 1 in the heterogeneous GSA is less than in a homogeneous GSA, in which all*

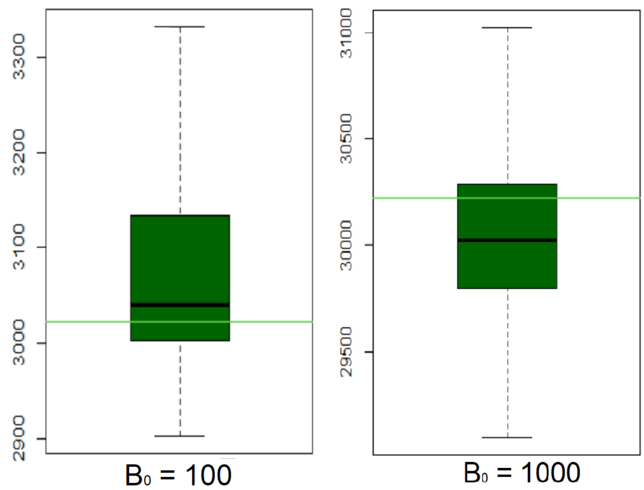
surviving members have the same amount of wealth as Bob.

For $F_0^B < 1$, *i.e* $F_0^B < F_0^A$

- *Bob's expected consumption in the GSA plan is greater than his initial wealth* F_0^B*;*
- *Bob's expected consumption conditional upon his survival to time 1 in the heterogeneous GSA is less than in a homogeneous GSA, in which all surviving members have the same amount of wealth as Bob.*

Watching carefully the simulations that I performed is very easy to visualize what has been expressed in the theory by Donnelly. Further, I report in the following graph the expected present discounted value (i.e. the money worth of Mitchell [21]) of all the simulations made, conditioned to the survivorship at the time $t = 44$. As already done in the constant contribution case I report with a box-plot the general behaviour of the GSA money's worth and with a thin green line the benchmark level of an annuity. In the left panel I consider the poorest members, those with an initial payout $B_0 = 100$, and in the right panel I consider the richest individuals, those with an initial benefit $B_0 = 1000$.

Figure 3.19: EPDV - General trend conditional upon survival at time $t = 44$.



It is evident that for poorer members, the EPDV of the GSA plan is higher than that of the annuity in most cases. On the other hand, for richest individuals the exact reverse applies: the EPDV of the annuity is higher than that of the GSA plan. Again, it is clear that, conditional upon survival to time $t = 44$, poorest members benefit more from joining the pool.

In this type of tontine there is therefore a form of discrimination, in the sense of Milevsky and Salisbury (2016) [20]: the richest are forced to pay part of the benefits to poorest. But this can also be viewed as a form of solidarity, in which there are people who pay more and substitute for others.
The same argument can be made in a GSA plan in which people of different ages are allowed to enter in the pool, i.e. when we consider inhomogeneous cohorts. Older people are discriminated against younger people. The young benefit from subsidies from the old.

Chapter 4

Conclusions

Longevity risk is one of the major issues facing retirees. To mitigate the impact and severity of this challenge, individuals may choose to transfer their risks directly to insurance companies and annuitize their accumulated pension pot. Historically, and especially under pay-as-you-go or generally public pension schemes, individuals were forced to hedge against longevity risk through annuity products.

In this thesis, I looked at other potential options available to retirees. In particular, I focused my attention on modern tontines, which are very innovative and interesting products that can be seen as viable sustainable alternatives. In reviewing the various tontines proposed in the literature, a very interesting feature common to almost all products emerged: almost all modern tontines have returns and costs in line with fair annuity products. Tontine benefit payments are more volatile, but in general the payout patterns of the two contracts are really similar. However, insurance companies do not actually sell fair annuities in the marketplace. Insurers, in order to sustain their costs and risks, charge loadings to fair premiums. Premiums received by insurers are therefore higher than those assumed in fair annuities, or consistently, benefits received by pensioners are lower. The insurance company bears longevity risks of retirees and must be compensated for this.
In contrast, in a tontine there is a sharing mechanism: there is no an outside financial institution that assumes additional risks and guarantees certain incomes. Therefore, because of the risk-pooling mechanism among members, for the same level of benefits, the costs of a tontine will certainly be lower than an annuity sold in the market. Especially for the Group Self-Annuitization plan, for which I performed a simulation, this insight was evident. The GSA plan with constant contribution turned out to be in line with a fair annuity; the payout dynamic of the GSA plan reflects fair annuities and the money's worth of the two products was similar. But if we considered actual annuities sold in the market, with higher costs or lower payments, the GSA plan would turn out to perform better in almost every simulation.

The important aspect to be emphasized is that the tontines proposed in this work, and in particular the GSA plan, do not need an insurance company to be sold. Through the internal mechanism of risk sharing, the inside management of funds and the distribution of dividends among survivors, the tontines are able to be managed also by other types of financial providers, at a much lower cost. External providers, if present, would have to deal with more managerial and operational aspects. They would not have to take on the risks of all policyholders and provide them with a certain monthly return until death. In addition, they do not need to set aside solvency capital and meet supervisory requirements imposed by the authorities.
In a tontine, the pooling mechanism is the essential core.

Tontines should therefore be considered as appealing ways to fund later years and solve a lot of the problems to which many countries are subjected. Tontines may have considerable appeal in countries that have adopted public pay-as-you-go or defined benefit pension systems. There are indeed sustainability and funding challenges that can only be addressed through new products, such as tontine or pooled funds, that innovate the way longevity risk is managed.

References

[1] Andréasson, J. G., Shevchenko, P. V., and Novikov, A. (2017). *Optimal consumption, investment and housing with means-tested public pension in retirement*, Insurance: Mathematics and Economics, 75, 32-47.

[2] ANIA (2020). *La previdenza sociale e il valore della garanzia.*

[3] Bayraktar, E., and Young, V. R. (2007). *Minimizing the probability of lifetime ruin under borrowing constraints*, Insurance: Mathematics and Economics, 41(1), 196-221.

[4] Bernhardt, T., and Donnelly, C. (2018). *Pension decumulation strategies: A state of the art report.*

[5] Bruhn, K., and Steffensen, M. (2013). *Optimal smooth consumption and annuity design*, Journal of Banking and Finance, 37(8), 2693-2701.

[6] Data obtained through the Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at https://www.mortality.org

[7] Di Giacinto, M., Federico, S., Gozzi, F., and Vigna, E. (2014). *Income drawdown option with minimum guarantee*, European Journal of Operational Research, 234(3), 610-624.

[8] Donnelly, C. (2015). Actuarial fairness and solidarity in pooled annuity funds. ASTIN Bulletin: The Journal of the IAA, 45(1), 49-74.

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

[10] Financial Conduct Authority (2018). *Retirement Outcomes Review Final report*, annex 4: Data collection and analysis.

[11] Financial Conduct Authority (FCA), *Data Bulletin: Issue 14 – September 2018*

[12] Fullmer, R. K. (2009). *A Framework for Portfolio Decumulation*, Journal of Investment Consulting, 10(1), 63-71.

[13] Gerrard, R., Haberman, S., and Vigna, E. (2006). *The management of decumulation risks in a defined contribution pension plan*, North American Actuarial Journal, 10(1), 84-110.

[14] Guyton, J. T., and Klinger, W. J. (2006). *Decision rules and maximum initial withdrawal rates*, JOURNAL OF FINANCIAL PLANNING-DENVER-, 19(3), 48.

[15] Insurance Europe (2020). *Pension Survey.*

[16] McKeever, K. (2009). *A short history of tontines*, Fordham J. Corp. & Fin. L., 15, 491.

[17] Milevsky, M. A., and Huang, H. (2011). *Spending retirement on planet vulcan: The impact of longevity risk aversion on optimal withdrawal rates*, Financial Analysts Journal, 67(2), 45-58.

[18] Milevsky, M. A., and Salisbury, T. S. (2015). *Optimal retirement income tontines*, Insurance: Mathematics and economics, 64, 91-105.

[19] Milevsky, M. A., and Salisbury, T. S. (2015). *Optimal retirement income tontines*, Insurance: Mathematics and economics, 64, 91-105.

[20] Milevsky, M. A., and Salisbury, T. S. (2016). *Equitable retirement income tontines: Mixing cohorts without discriminating*, ASTIN Bulletin: The Journal of the IAA, 46(3), 571-604.

[21] Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., and Brown, J. R. (1999). *New evidence on the money's worth of individual annuities*, American economic review, 89(5), 1299-1318.

[22] Money and Pensions Service, *The Pensions Advisory Service*, https://www.pensionsadvisoryservice.org.uk/about-pensions/pension-reform/freedom-and-choice

[23] Munk, C. (2008). *Portfolio and consumption choice with stochastic investment opportunities and habit formation in preferences*, Journal of Economic Dynamics and Control, 32(11), 3560-3589.

[24] New York Times (2017). *When Others Die, Tontine Investors Win*

[25] Office for National Statistics. *Employee workplace pensions in the UK: 2018 provisional and 2017 revised results*, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/workplacepensions

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

[27] Sabin, M. J. (2010). *Fair tontine annuity*, Available at SSRN 1579932.

[28] Scott, J. S., Sharpe, W. F., and Watson, J. G. (2009). *The 4% Rule - At What Price?* , Journal of Investment Management, 3, 31.

[29] Serrano, A.S. and Peltonen, T. (2020). *Pension schemes in the European Union: challenges and implications from macroeconomic and financial stability perspectives*, ESRB Occasional Paper Series,17.

[30] Stamos, M. Z. (2008). *Optimal consumption and portfolio choice for pooled annuity funds*, Insurance: Mathematics and Economics, 43(1), 56-68.

[31] Tversky, A., and Kahneman, D. (1992). *Advances in prospect theory: Cumulative representation of uncertainty*, Journal of Risk and uncertainty, 5:297-323.

[32] Van Bilsen, S., Laeven, R. J., and Nijman, T. E. (2017). *Consumption and portfolio choice under loss aversion and endogenous updating of the reference level*

[33] Young, V. R. (2004). *Optimal investment strategy to minimize the probability of lifetime ruin*, North American Actuarial Journal, 8:106-126.

[34] https://qsuper.qld.gov.au/our-products/superannuation/lifetime-pension

[35] https://tontine.com

[36] https://www.conservateur.fr/nos-produits/tontine

Appendices

Code Constant Contribution GSA plan - One Simulation

Definition of variables:

```
male1910 = read_excel("male1910.xlsx") #mortality table
R = 0.01 #interest rate
B_0 = 100 #initial benefit set also equal to 200,300,400,500,
600,700,800,900 and 1000
l_x = 100000 #initial pool size
x = 65 #age of members at time 0

p_x = rep(0,45) #EXPECTED survival rates taken from mortality table
for( i in 1:45)
{ p_x[i] = male1910$lx[x+i+1]/male1910$lx[x+i] }

a_x = rep(0,46) #Annuity factors
for (t in 0:45) #First annuity factor
{
  c = ((1+R)^-t) *(male1927$lx[x+t+1]/male1927$lx[x+1])
  a_x[1] = a_x[1] + c
}

for (i in 1:45) #calculating other annuity factors recursively
{ a_x[i+1] = (a_x[i]-1)*(1+R)/ (male1927$lx[x+i+1]/male1927$lx[x+i]) }
```

Simulation of actual survivors:

```
Actual_l_x = rep(0,46) #ACTUAL number of survivors
Actual_l_x[1] <- l_x

for (i in 1:45)
{
  Actual_l_x[i+1] = rbinom(1, Actual_l_x[i] , p_x[i] )
}
```

Simulation of benefit payouts:

```
Fund_t <- rep(0,46) #total fund
IndividualFund_t <- rep(0,46) #individual fund
B_t <- rep(0,46) #benefit payout
Fund_t[1] = B_0 * a_x[1]*l_x
IndividualFund_t[1] = B_0 * a_x[1]
B_t[1] = B_0

for (i in 2:46)
{
  if (Actual_l_x[i] !=0)
  {Fund_t[i] = (Fund_t[i-1]-B_t[i-1]*Actual_l_x[i-1])*(1+R)
  IndividualFund_t[i] = Fund_t[i]/Actual_l_x[i-1]}
  if (Actual_l_x[i] ==0)
  {Fund_t[i] =0
  IndividualFund_t[i] =0}

  B_t[i] = Fund_t[i] /(Actual_l_x[i]*a_x[i])
}
```

Derivation of mortality adjustments and mortality credits:

```
MEA_t <- rep(0,45) #mortality adjustments

for (i in 2:45)
{
  MEA_t[i-1] = B_t[i]/B_t[i-1]
}

MC_t <- rep(0,45) #mortality credits
for (t in 2:45)
{
  MC_t[t-1] = (Actual_l_x[t-1]-Actual_l_x[t])/(Actual_l_x[t])*
 IndividualFund_t[t]/a_x[t]
}
```

Derivation of expected present discounted value of GSA plan:

```
p_x = c(1, p_x)
EPDV_GSA = 0
for ( i in 0:45)
{
   if(is.nan(B_t[i+1]))
     {break}
   c = p_x[i+1]*B_t[i+1]*((1+R)^(-i))
   EPDV_GSA = EPDV_GSA + c
}

EPDV_Annuity = 0
for ( i in 0:45)
{
   if(is.nan(B_t[i+1]))
   {break}
   d = p_x[i+1]*B_0*((1+R)^(-i))
   EPDV_Annuity = EPDV_Annuity + d
}
```

Charts:

```
Benefits <- data.frame(cbind(c(0:45), rep(100,46), B_t[1:46]))
ggplot() + theme_classic() +
  geom_line(data=Benefits, aes(x=X1, y=X3, col='GSA'), lwd=1)+
  geom_line(data=Benefits, aes(x=X1, y=X2, col='Annuity'), lwd=1, lty=2)+
  ylab('Benefits') + xlab('time t')

MEA <- data.frame(cbind(c(1:45), rep(1,45), MEA_t[1:45]))
ggplot() + theme_classic() +
  geom_line(data=MEA, aes(x=X1, y=X3), col ="blue3",lwd=1)+
  geom_line(data=MEA, aes(x=X1, y=X2), col ="red3", lwd=1, lty=2)+
  ylab('MEA') + xlab('time t')

MC <- data.frame(cbind(c(1:45),  MC_t[1:45]))
ggplot(data=MC, aes(x=X1, y=X2)) + theme_classic() +
  geom_point(col ="blue3", lwd=2)+
   ylab('MC') + xlab('time t')
```

Code Constant Contribution GSA plan - General Trend

100 000 simulations:

```
N= 100000
Actual_l_x = matrix(0, N, 46) #ACTUAL number of survivors
Actual_l_x[,1] <- l_x
Fund_t <- matrix(0,N,46) #Fund value at time t
Fund_t[,1] = B_0 * a_x[1]*l_x
IndividualFund_t <- matrix(0,N,46) #individual fund at time t
IndividualFund_t[,1] = B_0 * a_x[1]
B_t <- matrix(0,N,46) #benefit at time t
B_t[,1] = B_0
MEA_t <- matrix(1,N,46) #mortality adjustment
MC_t <- matrix(0,N,45) #mortality credit

for (i in 1:N)
{
 for (j in 2:46)
 {
  Actual_l_x[i,j] = rbinom(1, Actual_l_x[i,j-1], p_x[j-1])
  if (Actual_l_x[i,j] != 0)
  {Fund_t[i,j] = (Fund_t[i,j-1]-B_t[i,j-1]*Actual_l_x[i,j-1])*(1+R)
   IndividualFund_t[i,j] = Fund_t[i,j]/Actual_l_x[i,j-1]}
  else
  { Fund_t[i,j] = 0
    IndividualFund_t[i,j] =0
  }

   B_t[i,j] = Fund_t[i,j]/(a_x[j]*Actual_l_x[i,j])

  MEA_t[i,j-1] = B_t[i,j]/B_t[i,j-1]
  MC_t[i,j-1] = (Actual_l_x[i,j-1]-Actual_l_x[i,j])/(Actual_l_x[i,j])*
                              IndividualFund_t[i,j]/a_x[j]}
}


B_t_new <- data.frame(B_t)
colnames(B_t_new) <- c(1:46)
boxplot(B_t_new[,1:45], col = 5 , xlab ='time t', ylab = 'Benefit',
        outline = FALSE, main = '', lwd =0.5)


MC_new <- data.frame(MC_t)
colnames(MC_new) <- c(1:45)
boxplot(MC_new[,1:43], col = 10 , xlab ='time t', ylab = 'Mortality credits',
        outline = FALSE, main = '', lwd =0.5)


MEA_new <- data.frame(MEA_t)
colnames(MEA_new) <- c(1:45)
boxplot(MEA_new[,1:43], col = 15 , xlab ='time t', ylab = 'Mortality Adj',
                                    outline = FALSE, main = '', lwd =0.5)
```

EPDVs:

```
EPDV_GSA = rep(0,N)
for (j in 1:N){
for ( i in 0:45)
{
  if(is.nan(B_t[j,i+1]))
  {break}
  c = p_x[i+1]*B_t[j,i+1]*((1+R)^(-i))
  EPDV_GSA[j] = EPDV_GSA[j] + c
}}

EPDV_Annuity1 = 0 #confronto con pool in cui muoiono tutti a 109 anni
EPDV_Annuity2 = 0 #confronto con pool in cui muoiono tutti a 108 anni
EPDV_Annuity3 = 0 #confronto con pool in cui muoiono tutti a 107 anni

for ( i in 0:44)
{
  d = p_x[i+1]*B_0*((1+R)^(-i))
  EPDV_Annuity1 = EPDV_Annuity1 + d
}


for ( i in 0:43)
{
  d = p_x[i+1]*B_0*((1+R)^(-i))
  EPDV_Annuity2 = EPDV_Annuity2 + d
}

for ( i in 0:42)
{
  d = p_x[i+1]*B_0*((1+R)^(-i))
  EPDV_Annuity3 = EPDV_Annuity3 + d
}

boxplot(EPDV_GSA[which(is.finite(B_t[,45]))],
              outline = FALSE, col="darkgreen")
abline(h= EPDV_Annuity1, col=11)

boxplot(EPDV_GSA[which(is.finite(B_t[,44]) & is.nan(B_t[,45])) ],
              outline = FALSE,   col="darkgreen")
abline(h= EPDV_Annuity2,  col=11)

boxplot(EPDV_GSA[which(is.finite(B_t[,43]) & is.nan(B_t[,44]))],
                              outline = FALSE,   col="darkgreen")
abline(h= EPDV_Annuity3,  col=11)
```

Code Varying Contribution GSA plan - One Simulation

Definition of variables:

```
male1910 = read_excel('male1910.xlsx')
R = 0.01 #interest rate
B_01 = 100 #initial benefit group 1
B_02 = 200 #initial benefit group 2
B_03 = 300
B_04 = 400
B_05 = 500
B_06 = 600
B_07 = 700
B_08 = 800
B_09 = 900
B_10 = 1000

l_x1 = 10000 #initial pool size of all groups
x = 65 #age of members at time 0

p_x = rep(0,45) #EXPECTED survival rates
for( i in 1:44)
 { p_x[i] = male1910$lx[x+i+1]/male1910$lx[x+i] }

a_x = rep(0,46) #Annuity factors
for (t in 0:45)# First Annuity factor
{
 c = ((1+R)^-t) *(male1927$lx[x+t+1]/male1927$lx[x+1])
 a_x[1] = a_x[1] + c
 }

for (i in 2:46) #others recursively
{    a_x[i] = (a_x[i-1]-1)*(1+R)/ (male1927$lx[x+i]/male1927$lx[x+i-1]) }
```

Simulation of actual survivors:

```
Actual_l_x1 = rep(0,46) #ACTUAL number of survivors group1
Actual_l_x2 = rep(0,46) #ACTUAL number of survivors group2
Actual_l_x3 = rep(0,46) #ACTUAL number of survivors group3
Actual_l_x4 = rep(0,46) #ACTUAL number of survivors group4
Actual_l_x5 = rep(0,46) #ACTUAL number of survivors group5
Actual_l_x6 = rep(0,46) #ACTUAL number of survivors group6
Actual_l_x7 = rep(0,46) #ACTUAL number of survivors group7
Actual_l_x8 = rep(0,46) #ACTUAL number of survivors group8
Actual_l_x9 = rep(0,46) #ACTUAL number of survivors group9
Actual_l_x10 = rep(0,46) #ACTUAL number of survivors group10


Actual_l_x1[1] <- l_x1, Actual_l_x2[1] <- l_x1, Actual_l_x3[1] <- l_x1
Actual_l_x4[1] <- l_x1, Actual_l_x5[1] <- l_x1, Actual_l_x6[1] <- l_x1
Actual_l_x7[1] <- l_x1, Actual_l_x8[1] <- l_x1, Actual_l_x9[1] <- l_x1
Actual_l_x10[1] <- l_x1
```

```
for (i in 1:45)
{
   Actual_l_x1[i+1] = rbinom(1,Actual_l_x1[i], p_x[i])
   Actual_l_x2[i+1] = rbinom(1,Actual_l_x2[i], p_x[i])
   Actual_l_x3[i+1] = rbinom(1,Actual_l_x3[i], p_x[i])
   Actual_l_x4[i+1] = rbinom(1,Actual_l_x4[i], p_x[i])
   Actual_l_x5[i+1] = rbinom(1,Actual_l_x5[i], p_x[i])
   Actual_l_x6[i+1] = rbinom(1,Actual_l_x6[i], p_x[i])
   Actual_l_x7[i+1] = rbinom(1,Actual_l_x7[i], p_x[i])
   Actual_l_x8[i+1] = rbinom(1,Actual_l_x8[i], p_x[i])
   Actual_l_x9[i+1] = rbinom(1,Actual_l_x9[i], p_x[i])
   Actual_l_x10[i+1] = rbinom(1,Actual_l_x10[i], p_x[i])
}
```

Simulation of benefit payouts

```
Fund_t <- rep(0,46) #TOTAL fund of the entire pool
IndividualFund1_t <- rep(0,46) #Individual Fund group 1
IndividualFund2_t <- rep(0,46) #Individual Fund group 2
IndividualFund3_t <- rep(0,46) #Individual Fund group 3
IndividualFund4_t <- rep(0,46) #Individual Fund group 4
IndividualFund5_t <- rep(0,46) #Individual Fund group 5
IndividualFund6_t <- rep(0,46) #Individual Fund group 6
IndividualFund7_t <- rep(0,46) #Individual Fund group 7
IndividualFund8_t <- rep(0,46) #Individual Fund group 8
IndividualFund9_t <- rep(0,46) #Individual Fund group 9
IndividualFund10_t <- rep(0,46) #Individual Fund group 10

TotalIndividualFund1_t <- rep(0,45) #F hat
TotalIndividualFund2_t <- rep(0,45)
TotalIndividualFund3_t <- rep(0,45)
TotalIndividualFund4_t <- rep(0,45)
TotalIndividualFund5_t <- rep(0,45)
TotalIndividualFund6_t <- rep(0,45)
TotalIndividualFund7_t <- rep(0,45)
TotalIndividualFund8_t <- rep(0,45)
TotalIndividualFund9_t <- rep(0,45)
TotalIndividualFund10_t <- rep(0,45)

B_t1 <- rep(0,46) #Benefit group 1
B_t2 <- rep(0,46) #Benefit group 2
B_t3 <- rep(0,46) #Benefit group 3
B_t4 <- rep(0,46) #Benefit group 4
B_t5 <- rep(0,46) #Benefit group 5
B_t6 <- rep(0,46) #Benefit group 6
B_t7 <- rep(0,46) #Benefit group 7
B_t8 <- rep(0,46) #Benefit group 8
B_t9 <- rep(0,46) #Benefit group 9
B_t10 <- rep(0,46) #Benefit group 10
B_t <- rep(0,46) #TOTAL benefit entire pool

Fund1_t[1] = B_01 * a_x[1]*l_x1 , Fund2_t[1] = B_02 * a_x[1]*l_x1
Fund3_t[1] = B_03 * a_x[1]*l_x1 , Fund4_t[1] = B_04 * a_x[1]*l_x1
Fund5_t[1] = B_05 * a_x[1]*l_x1 , Fund6_t[1] = B_06 * a_x[1]*l_x1
Fund7_t[1] = B_07 * a_x[1]*l_x1 , Fund8_t[1] = B_08 * a_x[1]*l_x1
Fund9_t[1] = B_09 * a_x[1]*l_x1 , Fund10_t[1] = B_10 * a_x[1]*l_x1
```

```
Fund_t[1] = Fund1_t[1] + Fund2_t[1] + Fund3_t[1] + Fund4_t[1]
        +Fund5_t[1] +Fund6_t[1] +Fund7_t[1] +Fund8_t[1] +Fund9_t[1] +Fund10_t[1]

IndividualFund1_t[1] = B_01 * a_x[1], IndividualFund2_t[1] = B_02 * a_x[1]
IndividualFund3_t[1] = B_03 * a_x[1], IndividualFund4_t[1] = B_04 * a_x[1]
IndividualFund5_t[1] = B_05 * a_x[1], IndividualFund6_t[1] = B_06 * a_x[1]
IndividualFund7_t[1] = B_07 * a_x[1], IndividualFund8_t[1] = B_08 * a_x[1]
IndividualFund9_t[1] = B_09 * a_x[1], IndividualFund10_t[1] = B_10 * a_x[1]

B_t1[1] = B_01, B_t2[1] = B_02, B_t3[1] = B_03
B_t4[1] = B_04, B_t5[1] = B_05, B_t6[1] = B_06
B_t7[1] = B_07, B_t8[1] = B_08, B_t9[1] = B_09, B_t10[1] = B_10
B_t[1] = (B_01 + B_02 + B_03 + B_04 + B_05 + B_06
          + B_07 + B_08 + B_09 + B_10)*l_x1

denom <- rep(0,46) #funds of survivors

for (i in 2:46)
{
  Fund_t[i] = (B_t[i-1])*(a_x[i-1] - 1)*(1+R)
  B_t[i] = Fund_t[i] /a_x[i]

  if (Actual_l_x1[i]!=0)
  { TotalIndividualFund1_t[i-1] =  B_t1[i-1]*a_x[i-1]
    IndividualFund1_t[i] = (TotalIndividualFund1_t[i-1] - B_t1[i-1] )*(1+R)}
  else
  {IndividualFund1_t[i] = 0}

  if (Actual_l_x2[i]!=0)
  { TotalIndividualFund2_t[i-1] =  B_t2[i-1]*a_x[i-1]
  IndividualFund2_t[i] = (TotalIndividualFund2_t[i-1] - B_t2[i-1] )*(1+R)}
  else
  {IndividualFund2_t[i] = 0}

  if (Actual_l_x3[i]!=0)
  { TotalIndividualFund3_t[i-1] =  B_t3[i-1]*a_x[i-1]
  IndividualFund3_t[i] = (TotalIndividualFund3_t[i-1] - B_t3[i-1] )*(1+R)}
  else
  {IndividualFund3_t[i] = 0}

  if (Actual_l_x4[i]!=0)
  { TotalIndividualFund4_t[i-1] =  B_t4[i-1]*a_x[i-1]
  IndividualFund4_t[i] = (TotalIndividualFund4_t[i-1] - B_t4[i-1] )*(1+R)}
  else
  {IndividualFund4_t[i] = 0}

  if (Actual_l_x5[i]!=0)
  { TotalIndividualFund5_t[i-1] =  B_t5[i-1]*a_x[i-1]
  IndividualFund5_t[i] = (TotalIndividualFund5_t[i-1] - B_t5[i-1] )*(1+R)}
  else
  {IndividualFund5_t[i] = 0}

  if (Actual_l_x6[i]!=0)
  { TotalIndividualFund6_t[i-1] =  B_t6[i-1]*a_x[i-1]
```

```
  IndividualFund6_t[i] = (TotalIndividualFund6_t[i-1] - B_t6[i-1] )*(1+R)}
  else
  {IndividualFund6_t[i] = 0}

  if (Actual_l_x7[i]!=0)
  { TotalIndividualFund7_t[i-1] = B_t7[i-1]*a_x[i-1]
  IndividualFund7_t[i] = (TotalIndividualFund7_t[i-1] - B_t7[i-1] )*(1+R)}
  else
  {IndividualFund7_t[i] = 0}

  if (Actual_l_x8[i]!=0)
  { TotalIndividualFund8_t[i-1] = B_t8[i-1]*a_x[i-1]
  IndividualFund8_t[i] = (TotalIndividualFund8_t[i-1] - B_t8[i-1] )*(1+R)}
  else
  {IndividualFund8_t[i] = 0}

  if (Actual_l_x9[i]!=0)
  { TotalIndividualFund9_t[i-1] = B_t9[i-1]*a_x[i-1]
  IndividualFund9_t[i] = (TotalIndividualFund9_t[i-1] - B_t9[i-1] )*(1+R)}
  else
  {IndividualFund9_t[i] = 0}

  if (Actual_l_x10[i]!=0)
  { TotalIndividualFund10_t[i-1] = B_t10[i-1]*a_x[i-1]
  IndividualFund10_t[i] = (TotalIndividualFund10_t[i-1] - B_t10[i-1] )*(1+R)}
  else
  {IndividualFund10_t[i] = 0}

  denom[i] =IndividualFund1_t[i]*Actual_l_x1[i]+IndividualFund2_t[i]*Actual_l_x2[i]
     + IndividualFund3_t[i]*Actual_l_x3[i]+IndividualFund4_t[i]*Actual_l_x4[i]
     + IndividualFund5_t[i]*Actual_l_x5[i]+IndividualFund6_t[i]*Actual_l_x6[i]
     + IndividualFund7_t[i]*Actual_l_x7[i]+IndividualFund8_t[i]*Actual_l_x8[i]
     + IndividualFund9_t[i]*Actual_l_x9[i]+IndividualFund10_t[i]*Actual_l_x10[i]

   B_t1[i] = B_t[i]*(IndividualFund1_t[i])/(denom[i])
   B_t2[i] = B_t[i]* (IndividualFund2_t[i])/(denom[i])
   B_t3[i] = B_t[i]* (IndividualFund3_t[i])/(denom[i])
   B_t4[i] = B_t[i]* (IndividualFund4_t[i])/(denom[i])
   B_t5[i] = B_t[i]* (IndividualFund5_t[i])/(denom[i])
   B_t6[i] = B_t[i]* (IndividualFund6_t[i])/(denom[i])
   B_t7[i] = B_t[i]* (IndividualFund7_t[i])/(denom[i])
   B_t8[i] = B_t[i]* (IndividualFund8_t[i])/(denom[i])
   B_t9[i] = B_t[i]* (IndividualFund9_t[i])/(denom[i])
   B_t10[i] = B_t[i]* (IndividualFund10_t[i])/(denom[i])
}
```

Derivation of mortality credits and mortality adjustment:

```
MEA1_t <- rep(1,46) #mortality adjustment group 1
MEA2_t <- rep(1,46)  #mortality adjustment group 2
MEA3_t <- rep(1,46)  #mortality adjustment group 3
MEA4_t <- rep(1,46)  #mortality adjustment group 4
MEA5_t <- rep(1,46)  #mortality adjustment group 5
MEA6_t <- rep(1,46)  #mortality adjustment group 6
MEA7_t <- rep(1,46)  #mortality adjustment group 7
```

```
MEA8_t <- rep(1,46)  #mortality adjustment group 8
MEA9_t <- rep(1,46)  #mortality adjustment group 9
MEA10_t <- rep(1,46)#mortality adjustment group 10

for (i in 2:46)
{
  MEA1_t[i] = B_t1[i]/B_t1[i-1]
  MEA2_t[i] = B_t2[i]/B_t2[i-1]
  MEA3_t[i] = B_t3[i]/B_t3[i-1]
  MEA4_t[i] = B_t4[i]/B_t4[i-1]
  MEA5_t[i] = B_t5[i]/B_t5[i-1]
  MEA6_t[i] = B_t6[i]/B_t6[i-1]
  MEA7_t[i] = B_t7[i]/B_t7[i-1]
  MEA8_t[i] = B_t8[i]/B_t8[i-1]
  MEA9_t[i] = B_t9[i]/B_t9[i-1]
  MEA10_t[i] = B_t10[i]/B_t10[i-1]
}

MC1_t <- B_t1 - IndividualFund1_t/a_x
MC2_t <- B_t2 - IndividualFund2_t/a_x
MC3_t <- B_t3 - IndividualFund3_t/a_x
MC4_t <- B_t4 - IndividualFund4_t/a_x
MC5_t <- B_t5 - IndividualFund5_t/a_x
MC6_t <- B_t6 - IndividualFund6_t/a_x
MC7_t <- B_t7 - IndividualFund7_t/a_x
MC8_t <- B_t8 - IndividualFund8_t/a_x
MC9_t <- B_t9 - IndividualFund9_t/a_x
MC10_t <- B_t10 - IndividualFund10_t/a_x
```

Derivation of expected present discounted value (EPDV):

```
EPDV_GSA1 = 0
for ( i in 0:45)
{ if(is.nan(B_t1[i+1]))
  {break}
  c = p_x[i+1]*B_t1[i+1]*((1+R)^(-i))
  EPDV_GSA1 = EPDV_GSA1 + c}
EPDV_Annuity1 = 0
for ( i in 0:45)
{ if(is.nan(B_t1[i+1]))
  {break}
  d = p_x[i+1]*B_01*((1+R)^(-i))
  EPDV_Annuity1 = EPDV_Annuity1 + d}

EPDV_GSA2 = 0
for ( i in 0:45)
{ if(is.nan(B_t2[i+1]))
  {break}
  c = p_x[i+1]*B_t2[i+1]*((1+R)^(-i))
  EPDV_GSA2 = EPDV_GSA2 + c}
EPDV_Annuity2 = 0
for ( i in 0:45)
{ if(is.nan(B_t2[i+1]))
  {break}
  d = p_x[i+1]*B_02*((1+R)^(-i))
  EPDV_Annuity2 = EPDV_Annuity2 + d}
```

```
EPDV_GSA3 = 0
for ( i in 0:45)
{ if(is.nan(B_t3[i+1]))
  {break}
  c = p_x[i+1]*B_t3[i+1]*((1+R)^(-i))
  EPDV_GSA3 = EPDV_GSA3 + c}
EPDV_Annuity3 = 0
for ( i in 0:45)
{ if(is.nan(B_t3[i+1]))
  {break}
  d = p_x[i+1]*B_03*((1+R)^(-i))
  EPDV_Annuity3 = EPDV_Annuity3 + d}

EPDV_GSA4 = 0
for ( i in 0:45)
{ if(is.nan(B_t4[i+1]))
  {break}
  c = p_x[i+1]*B_t4[i+1]*((1+R)^(-i))
  EPDV_GSA4 = EPDV_GSA4 + c}
EPDV_Annuity4 = 0
for ( i in 0:45)
{ if(is.nan(B_t4[i+1]))
  {break}
  d = p_x[i+1]*B_04*((1+R)^(-i))
  EPDV_Annuity4 = EPDV_Annuity4 + d}

EPDV_GSA5 = 0
for ( i in 0:45)
{ if(is.nan(B_t5[i+1]))
  {break}
  c = p_x[i+1]*B_t5[i+1]*((1+R)^(-i))
  EPDV_GSA5 = EPDV_GSA5 + c}
EPDV_Annuity5 = 0
for ( i in 0:45)
{ if(is.nan(B_t5[i+1]))
  {break}
  d = p_x[i+1]*B_05*((1+R)^(-i))
  EPDV_Annuity5 = EPDV_Annuity5 + d}

EPDV_GSA6 = 0
for ( i in 0:45)
{ if(is.nan(B_t6[i+1]))
  {break}
  c = p_x[i+1]*B_t6[i+1]*((1+R)^(-i))
  EPDV_GSA6 = EPDV_GSA6 + c}
EPDV_Annuity6 = 0
for ( i in 0:45)
{ if(is.nan(B_t6[i+1]))
  {break}
  d = p_x[i+1]*B_06*((1+R)^(-i))
  EPDV_Annuity6 = EPDV_Annuity6 + d}
```

```
EPDV_GSA7 = 0
for ( i in 0:45)
{ if(is.nan(B_t7[i+1]))
  {break}
  c = p_x[i+1]*B_t7[i+1]*((1+R)^(-i))
  EPDV_GSA7 = EPDV_GSA7 + c}
EPDV_Annuity7 = 0
for ( i in 0:45)
{ if(is.nan(B_t7[i+1]))
  {break}
  d = p_x[i+1]*B_07*((1+R)^(-i))
  EPDV_Annuity7 = EPDV_Annuity7 + d}

EPDV_GSA8 = 0
for ( i in 0:45)
{ if(is.nan(B_t8[i+1]))
  {break}
  c = p_x[i+1]*B_t8[i+1]*((1+R)^(-i))
  EPDV_GSA8 = EPDV_GSA8 + c}
EPDV_Annuity8 = 0
for ( i in 0:45)
{ if(is.nan(B_t8[i+1]))
  {break}
  d = p_x[i+1]*B_08*((1+R)^(-i))
  EPDV_Annuity8 = EPDV_Annuity8 + d}

EPDV_GSA9 = 0
for ( i in 0:45)
{ if(is.nan(B_t9[i+1]))
  {break}
  c = p_x[i+1]*B_t9[i+1]*((1+R)^(-i))
  EPDV_GSA9 = EPDV_GSA9 + c}
EPDV_Annuity9 = 0
for ( i in 0:45)
{ if(is.nan(B_t9[i+1]))
  {break}
  d = p_x[i+1]*B_09*((1+R)^(-i))
  EPDV_Annuity9 = EPDV_Annuity9 + d}

EPDV_GSA10 = 0
for ( i in 0:45)
{ if(is.nan(B_t10[i+1]))
  {break}
  c = p_x[i+1]*B_t10[i+1]*((1+R)^(-i))
  EPDV_GSA10 = EPDV_GSA10 + c}
EPDV_Annuity10 = 0
for ( i in 0:45)
{ if(is.nan(B_t10[i+1]))
  {break}
  d = p_x[i+1]*B_10*((1+R)^(-i))
  EPDV_Annuity10 = EPDV_Annuity10 + d}
```

Charts:

```
benefits <- data.frame(cbind(c(0:44),B_t1,B_t2, B_t3, B_t4, B_t5,
                              B_t6, B_t7, B_t8, B_t9, B_t10))[1:45,]
colnames(benefits) <- c("time", "Group1", "Group2",  "Group3",  "Group4",
        "Group5"  , "Group6","Group7" , "Group8", "Group9" , "Group10")
ggplot() + theme_classic()+
  geom_line(data = benefits, aes(x = time, y = Group1, col='B0=100'), lwd=1) +
  geom_line(data = benefits, aes(x = time, y = Group2, col='B0=200'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group3, col='B0=300'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group4, col='B0=400'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group5, col='B0=500'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group6, col='B0=600'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group7, col='B0=700'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group8, col='B0=800'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group9, col='B0=900'), lwd =1) +
  geom_line(data = benefits, aes(x = time, y = Group10, col='B0=1000'), lwd =1) +
  ylab('Benefits') + xlab('time t') +  scale_color_manual(name = 'Group')+
  ggtitle('GSA benefit') + scale_color_brewer(palette = "Paired")


MEA <- data.frame(cbind(c(0:44),MEA1_t,MEA2_t, MEA3_t,
                          MEA4_t,MEA5_t, MEA6_t,MEA7_t,MEA8_t,MEA9_t,MEA10_t))[1:45,]
colnames(MEA) <- c("time", "Group1", "Group2",  "Group3",  "Group4",
        "Group5"  , "Group6","Group7" , "Group8", "Group9" , "Group10")
ggplot() + theme_classic()+
  geom_line(data = MEA, aes(x = time, y = Group1, col='B0=100'), lwd=1) +
  geom_line(data = MEA, aes(x = time, y = Group2, col='B0=200'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group3, col='B0=300'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group4, col='B0=400'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group5, col='B0=500'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group6, col='B0=600'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group7, col='B0=700'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group8, col='B0=800'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group9, col='B0=900'), lwd =1) +
  geom_line(data = MEA, aes(x = time, y = Group10, col='B0=1000'), lwd =1) +
  ylab('MEA') + xlab('time t') +  scale_color_manual(name = 'Group')+
  ggtitle('Mortality Adjustments') + scale_color_brewer(palette = "Paired")

MC <- data.frame(cbind(c(0:44),MC1_t,MC2_t, MC3_t,
                          MC4_t,MC5_t, MC6_t, MC7_t,MC8_t, MC9_t, MC10_t))[2:45,]
colnames(MC) <-c("time", "Group1", "Group2",  "Group3",  "Group4",
        "Group5"  , "Group6","Group7" , "Group8", "Group9" , "Group10")
ggplot() + theme_classic()+
  geom_line(data = MC, aes(x = time, y = Group1, col='B0=100'), lwd=1) +
  geom_line(data = MC, aes(x = time, y = Group2, col='B0=200'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group3, col='B0=300'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group4, col='B0=400'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group5, col='B0=500'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group6, col='B0=600'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group7, col='B0=700'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group8, col='B0=800'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group9, col='B0=900'), lwd =1) +
  geom_line(data = MC, aes(x = time, y = Group10, col='B0=1000'), lwd =1) +
  ylab('MC') + xlab('time t') +  scale_color_manual(name = 'Group')+
  ggtitle('Mortality Credits') + scale_color_brewer(palette = "Paired")
```

Code Varying contribution GSA plan - General Trend

100 000 Simulations:

```
N=100000
Actual_l_x1 = matrix(0,N, 46) #ACTUAL number of survivors group1
Actual_l_x2 = matrix(0,N, 46) #ACTUAL number of survivors group2
Actual_l_x3 = matrix(0,N, 46) #ACTUAL number of survivors group3
Actual_l_x4 = matrix(0,N, 46) #ACTUAL number of survivors group4
Actual_l_x5 = matrix(0,N, 46)#ACTUAL number of survivors group5
Actual_l_x6 = matrix(0,N, 46) #ACTUAL number of survivors group6
Actual_l_x7 = matrix(0,N, 46) #ACTUAL number of survivors group7
Actual_l_x8 = matrix(0,N, 46)#ACTUAL number of survivors group8
Actual_l_x9 = matrix(0,N, 46) #ACTUAL number of survivors group9
Actual_l_x10 = matrix(0,N, 46) #ACTUAL number of survivors group10


Actual_l_x1[,1] <- l_x1, Actual_l_x2[,1] <- l_x1, Actual_l_x3[,1] <- l_x1
Actual_l_x4[,1] <- l_x1, Actual_l_x5[,1] <- l_x1, Actual_l_x6[,1] <- l_x1
Actual_l_x7[,1] <- l_x1, Actual_l_x8[,1] <- l_x1, Actual_l_x9[,1] <- l_x1
Actual_l_x10[,1] <- l_x1

for (i in 1:N){
for (j in 2:46)
{
  Actual_l_x1[i,j] = rbinom(1,Actual_l_x1[i,j-1], p_x[j-1])
  Actual_l_x2[i,j] = rbinom(1,Actual_l_x2[i,j-1], p_x[j-1])
  Actual_l_x3[i,j] = rbinom(1,Actual_l_x3[i,j-1], p_x[j-1])
  Actual_l_x4[i,j] = rbinom(1,Actual_l_x4[i,j-1], p_x[j-1])
  Actual_l_x5[i,j] = rbinom(1,Actual_l_x5[i,j-1], p_x[j-1])
  Actual_l_x6[i,j] = rbinom(1,Actual_l_x6[i,j-1], p_x[j-1])
  Actual_l_x7[i,j] = rbinom(1,Actual_l_x7[i,j-1], p_x[j-1])
  Actual_l_x8[i,j] = rbinom(1,Actual_l_x8[i,j-1], p_x[j-1])
  Actual_l_x9[i,j] = rbinom(1,Actual_l_x9[i,j-1], p_x[j-1])
  Actual_l_x10[i,j] = rbinom(1,Actual_l_x10[i,j-1], p_x[j-1])
}}

Fund1_t <- matrix(0,N,46), Fund2_t <- matrix(0,N,46), Fund3_t <- matrix(0,N,46),
Fund4_t <- matrix(0,N,46), Fund5_t <- matrix(0,N,46), Fund6_t <- matrix(0,N,46),
Fund7_t <- matrix(0,N,46), Fund8_t <- matrix(0,N,46), Fund9_t <- matrix(0,N,46),
Fund10_t <- matrix(0,N,46), Fund_t <- matrix(0,N,46)

IndividualFund1_t <- matrix(0,N,46), IndividualFund2_t <- matrix(0,N,46)
IndividualFund3_t <- matrix(0,N,46), IndividualFund4_t <- matrix(0,N,46)
IndividualFund5_t <- matrix(0,N,46), IndividualFund6_t <- matrix(0,N,46)
IndividualFund7_t <- matrix(0,N,46), IndividualFund8_t <- matrix(0,N,46)
IndividualFund9_t <- matrix(0,N,46), IndividualFund10_t <- matrix(0,N,46)

TotalIndividualFund1_t <- matrix(0,N,45), TotalIndividualFund2_t <- matrix(0,N,45)
TotalIndividualFund3_t <- matrix(0,N,45), TotalIndividualFund4_t <- matrix(0,N,45)
TotalIndividualFund5_t <- matrix(0,N,45), TotalIndividualFund6_t <- matrix(0,N,45)
TotalIndividualFund7_t <- matrix(0,N,45), TotalIndividualFund8_t <- matrix(0,N,45)
TotalIndividualFund9_t <- matrix(0,N,45), TotalIndividualFund10_t <- matrix(0,N,45)
```

```
B_t1 <- matrix(0,N,46), B_t2 <- matrix(0,N,46), B_t3 <- matrix(0,N,46),
B_t4 <- matrix(0,N,46), B_t5 <- matrix(0,N,46), B_t6 <- matrix(0,N,46),
B_t7 <- matrix(0,N,46), B_t8 <- matrix(0,N,46), B_t9 <- matrix(0,N,46),
B_t10 <- matrix(0,N,46), B_t <- matrix(0,N,46)

Fund1_t[,1] = B_01 * a_x[1]*l_x1, Fund2_t[,1] = B_02 * a_x[1]*l_x1
Fund3_t[,1] = B_03 * a_x[1]*l_x1, Fund4_t[,1] = B_04 * a_x[1]*l_x1
Fund5_t[,1] = B_05 * a_x[1]*l_x1, Fund6_t[,1] = B_06 * a_x[1]*l_x1
Fund7_t[,1] = B_07 * a_x[1]*l_x1, Fund8_t[,1] = B_08 * a_x[1]*l_x1
Fund9_t[,1] = B_09 * a_x[1]*l_x1, Fund10_t[,1] = B_10 * a_x[1]*l_x1
Fund_t[,1] = Fund1_t[,1] + Fund2_t[,1] + Fund3_t[,1] + Fund4_t[,1]
+Fund5_t[,1] +Fund6_t[,1] +Fund7_t[,1] +Fund8_t[,1] +Fund9_t[,1] +Fund10_t[,1]

IndividualFund1_t[,1] = B_01 * a_x[1], IndividualFund2_t[,1] = B_02 * a_x[1]
IndividualFund3_t[,1] = B_03 * a_x[1], IndividualFund4_t[,1] = B_04 * a_x[1]
IndividualFund5_t[,1] = B_05 * a_x[1], IndividualFund6_t[,1] = B_06 * a_x[1]
IndividualFund7_t[,1] = B_07 * a_x[1], IndividualFund8_t[,1] = B_08 * a_x[1]
IndividualFund9_t[,1] = B_09 * a_x[1], IndividualFund10_t[,1] = B_10 * a_x[1]

B_t1[,1] = B_01, B_t2[,1] = B_02, B_t3[,1] = B_03, B_t4[,1] = B_04,
B_t5[,1] = B_05, B_t6[,1] = B_06, B_t7[,1] = B_07, B_t8[,1] = B_08,
B_t9[,1] = B_09, B_t10[,1] = B_10

B_t[,1] = (B_01 + B_02 + B_03 + B_04 + B_05 + B_06
              + B_07 + B_08 + B_09 + B_10)*l_x1
denom <- matrix(0,N,46)

for (i in 1:N){
for (j in 2:46)
{
  Fund_t[i,j] = (B_t[i,j-1])*(a_x[j-1] - 1)*(1+R)
  B_t[i,j] = Fund_t[i,j] /a_x[j]

  if (Actual_l_x1[i,j]!=0)
  { TotalIndividualFund1_t[i,j-1] =  B_t1[i,j-1]*a_x[j-1]
  IndividualFund1_t[i,j] = (TotalIndividualFund1_t[i,j-1] - B_t1[i,j-1] )*(1+R)}
  else
  {IndividualFund1_t[i,j] = 0}

  if (Actual_l_x2[i,j]!=0)
  { TotalIndividualFund2_t[i,j-1] =  B_t2[i,j-1]*a_x[j-1]
  IndividualFund2_t[i,j] = (TotalIndividualFund2_t[i,j-1] - B_t2[i,j-1] )*(1+R)}
  else
  {IndividualFund2_t[i,j] = 0}

  if (Actual_l_x3[i,j]!=0)
  { TotalIndividualFund3_t[i,j-1] =  B_t3[i,j-1]*a_x[j-1]
  IndividualFund3_t[i,j] = (TotalIndividualFund3_t[i,j-1] - B_t3[i,j-1] )*(1+R)}
  else
  {IndividualFund3_t[i,j] = 0}

  if (Actual_l_x4[i,j]!=0)
  { TotalIndividualFund4_t[i,j-1] =  B_t4[i,j-1]*a_x[j-1]
  IndividualFund4_t[i,j] = (TotalIndividualFund4_t[i,j-1] - B_t4[i,j-1] )*(1+R)}
```

```
else
{IndividualFund4_t[i,j] = 0}

if (Actual_l_x5[i,j]!=0)
{ TotalIndividualFund5_t[i,j-1] = B_t5[i,j-1]*a_x[j-1]
IndividualFund5_t[i,j] = (TotalIndividualFund5_t[i,j-1] - B_t5[i,j-1] )*(1+R)}
else
{IndividualFund5_t[i,j] = 0}

if (Actual_l_x6[i,j]!=0)
{ TotalIndividualFund6_t[i,j-1] = B_t6[i,j-1]*a_x[j-1]
IndividualFund6_t[i,j] = (TotalIndividualFund6_t[i,j-1] - B_t6[i,j-1] )*(1+R)}
else
{IndividualFund6_t[i,j] = 0}

if (Actual_l_x7[i,j]!=0)
{ TotalIndividualFund7_t[i,j-1] = B_t7[i,j-1]*a_x[j-1]
IndividualFund7_t[i,j] = (TotalIndividualFund7_t[i,j-1] - B_t7[i,j-1] )*(1+R)}
else
{IndividualFund7_t[i,j] = 0}

if (Actual_l_x8[i,j]!=0)
{ TotalIndividualFund8_t[i,j-1] = B_t8[i,j-1]*a_x[j-1]
IndividualFund8_t[i,j] = (TotalIndividualFund8_t[i,j-1] - B_t8[i,j-1] )*(1+R)}
else
{IndividualFund8_t[i,j] = 0}

if (Actual_l_x9[i,j]!=0)
{ TotalIndividualFund9_t[i,j-1] = B_t9[i,j-1]*a_x[j-1]
IndividualFund9_t[i,j] = (TotalIndividualFund9_t[i,j-1] - B_t9[i,j-1] )*(1+R)}
else
{IndividualFund9_t[i,j] = 0}

if (Actual_l_x10[i,j]!=0)
{ TotalIndividualFund10_t[i,j-1] = B_t10[i,j-1]*a_x[j-1]
IndividualFund10_t[i,j] = (TotalIndividualFund10_t[i,j-1] - B_t10[i,j-1] )*(1+R)}
else
{IndividualFund10_t[i,j] = 0}

denom[i,j] = IndividualFund1_t[i,j]*Actual_l_x1[i,j] +
             IndividualFund2_t[i,j]*Actual_l_x2[i,j] +
             IndividualFund3_t[i,j]*Actual_l_x3[i,j] +
             IndividualFund4_t[i,j]*Actual_l_x4[i,j] +
             IndividualFund5_t[i,j]*Actual_l_x5[i,j] +
             IndividualFund6_t[i,j]*Actual_l_x6[i,j] +
             IndividualFund7_t[i,j]*Actual_l_x7[i,j] +
             IndividualFund8_t[i,j]*Actual_l_x8[i,j] +
             IndividualFund9_t[i,j]*Actual_l_x9[i,j] +
             IndividualFund10_t[i,j]*Actual_l_x10[i,j]



B_t1[i,j] = B_t[i,j]*(IndividualFund1_t[i,j])/(denom[i,j])
B_t2[i,j] = B_t[i,j]* (IndividualFund2_t[i,j])/(denom[i,j])
B_t3[i,j] = B_t[i,j]* (IndividualFund3_t[i,j])/(denom[i,j])
```

```
  B_t4[i,j] = B_t[i,j]* (IndividualFund4_t[i,j])/(denom[i,j])
  B_t5[i,j] = B_t[i,j]* (IndividualFund5_t[i,j])/(denom[i,j])
  B_t6[i,j] = B_t[i,j]* (IndividualFund6_t[i,j])/(denom[i,j])
  B_t7[i,j] = B_t[i,j]* (IndividualFund7_t[i,j])/(denom[i,j])
  B_t8[i,j] = B_t[i,j]* (IndividualFund8_t[i,j])/(denom[i,j])
  B_t9[i,j] = B_t[i,j]* (IndividualFund9_t[i,j])/(denom[i,j])
  B_t10[i,j] = B_t[i,j]* (IndividualFund10_t[i,j])/(denom[i,j])
}}

chart1 <- data.frame(B_t1)
boxplot(chart1[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart2 <- data.frame(B_t2)
boxplot(chart2[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart3 <- data.frame(B_t3)
boxplot(chart3[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart4 <- data.frame(B_t4)
boxplot(chart4[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart5 <- data.frame(B_t5)
boxplot(chart5[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart6 <- data.frame(B_t6)
boxplot(chart6[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart7 <- data.frame(B_t7)
boxplot(chart7[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart8 <- data.frame(B_t8)
boxplot(chart8[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart9 <- data.frame(B_t9)
boxplot(chart9[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
chart10 <- data.frame(B_t10)
boxplot(chart10[,1:46], col = 5 , xlab ='time t', ylab = 'Benefit',
             outline = FALSE, main = '', lwd =0.5)
```

EPDVs:

```
EPDV_GSA1 = rep(0,N)
for (j in 1:N){
for ( i in 0:45)
  { if(is.nan(B_t1[j,i+1]))
    {break}
    c = p_x[i+1]*B_t1[j,i+1]*((1+R)^(-i))
    EPDV_GSA1[j] = EPDV_GSA1[j] + c }}

EPDV_Annuity1 = 0 #benchmark annuity EPDV conditional upon survival at time t=44
for ( i in 0:45)
{ d = p_x[i+1]*B_01*((1+R)^(-i))
  EPDV_Annuity1 = EPDV_Annuity1 + d}
```

```
EPDV_GSA10 = rep(0,N)
for (j in 1:N){
for ( i in 0:45)
  { if(is.nan(B_t10[j,i+1]))
    {break}
    c = p_x[i+1]*B_t10[j,i+1]*((1+R)^(-i))
    EPDV_GSA10[j] = EPDV_GSA10[j] + c }}

EPDV_Annuity10 = 0 #benchmark annuity EPDV conditional upon survival at time t=44
for ( i in 0:45)
{ d = p_x[i+1]*B_10*((1+R)^(-i))
  EPDV_Annuity10 = EPDV_Annuity10 + d}


boxplot(EPDV_GSA1[which(is.finite(B_t1[,45]))], outline = FALSE, col="darkgreen")
abline(h= EPDV_Annuity1,col=11)
boxplot(EPDV_GSA10[which(is.finite(B_t1[,45]))], outline = FALSE, col="darkgreen")
abline(h= EPDV_Annuity10,col=11)
```

Ringraziamenti

Vorrei dedicare qualche riga a coloro che hanno contribuito alla realizzazione della mia tesi di laurea e che mi hanno supportata durante tutto il percorso di studi.

Innanzitutto, ringrazio la mia relatrice Vigna Elena, per la sua immensa disponibilità, per i suoi indispensabili consigli, e per le conoscenze trasmesse durante l'intera stesura dell'elaborato.
Grazie anche al mio correlatore Lods Bertrand per il suo sostegno.

Ringrazio infinitamente i miei genitori, senza i quali non avrei potuto intraprendere questo percorso. Grazie mamma e grazie papà per avermi sostenuta in ogni momento, per aver creduto in me e per la vostra immensa pazienza. Ve ne sarò per sempre grata.

Grazie di cuore a mio fratello, Mauro. Per essere stato il mio mentore, la mia linea guida e la mia spalla. Gran parte di ciò che sono lo devo sicuramente a te.

Grazie a tutta la mia famiglia per la loro vicinanza, stima e per il loro affetto. In particolare grazie ai miei nonni, per avermi cresciuta e amata in tutti questi anni.

Ringrazio immensamente Fra, per essere cresciuto insieme a me. Grazie per essere stato la mia certezza, la mia ancora e la mia felicità. Grazie per avermi resa una persona migliore.

Ringrazio Ale e Gio, le mie amiche del cuore. Grazie per essermi sempre state accanto e per avermi spronata a fare sempre del mio meglio.

Ringrazio Checco e Nic, i miei compagni di avventura. Grazie per tutte le esperienze condivise. Siete stati la mia salvezza.

Ringrazio Giu, Mario e Edo, la mia family per scelta. Grazie per le mille partite di burraco serali, per le risate, la spensieratezza e la leggerezza che mi regalate quotidianamente.

Ringrazio Dadi, Enri e anche Bongo, i miei super compagni del cuore. Grazie per aver condiviso con me questo percorso di studi, lo avete reso sicuramente migliore.

Grazie anche a tutti gli altri miei amici e colleghi di studio, che hanno rallegrato questi anni rendendoli più colorati e meno faticosi.