H. Lüthy, P. L. Keller, K. Binswanger, B. Gmür, Zürich

Adaptive Algorithmic Annuities

Introduction

One of the reasons for the relative illiquidity of the annuity market is the fact that the future mortality development is unknown. The consensus amongst a majority of specialists is that mortality will continue to improve so that, using prudent assumptions, annuities become quite expensive. Also, longevity risk is not a risk which can be mitigated by writing a large book of annuity business.

The systemic nature of the longevity risk has as a secondary effect that the reinsurance market is very thin since few reinsurers are prepared to take such a nondiversifiable risk. Hence especially small and medium sized companies, find it difficult to enter the annuity market since they find it difficult to take on the whole longevity risk.

The second systematic risk in the context of annuities is the technical interest rate which in fact is a guaranteed minimal return on the savings. Intending to managing this risk, insurance companies created unit linked products, where the investment risk is born by the policyholder. Consequently, a company selling unit linked annuities has no longer the possibility to cross-subsidise the two systematic risks and is hence even more exposed to longevity risk.

A further complicating factor are the widely divergent perceptions of different market participants regarding the future development of mortality. On one side there are agents who need to sell the annuities and have an interest of using a high mortality in the calculation basis as possible. On the other side are regulators, rating agencies and reinsurers who adopt a more conservative approach.

Since no convergence in opinion regarding future mortality is on the horizon at the moment, one possible way out of this quandary would be to make the annuity less dependent on mortality. In the following we present a way of achieving this.

The main reason to design a new annuity product is the recognition that the main source of longevity risk is the lock-in of assumptions at the beginning of the contract.

Even when experience shows that the assumptions were incorrect, the company (or the policyholder) can not change the basis anymore.

We propose to use frequent estimates of the actual mortality to adjust the benefits to the policyholder. There is no unique way in adjusting the benefit payments nor in estimating the mortality. The former can be chosen such that less or more of

the longevity risk is transferred to the policyholder, the latter has to be done in a reasonably simple and transparent way. We would like to emphasise that the main trust of this paper is not mathematical sophistication but rather a proposal for an annuity product which we feel can be sold in the market and can get approval by the regulators.
The main ingredients of a 'mortality linked' product are:

The mortality parameter model,
the parameter estimation procedure,
the benefit adaptation scheme.

In the following we present a number of suggestions on how to chose a simple mortality model and how to estimate the relevant parameters. We then show how the benefits can be adjusted depending on the current best mortality estimate and how the effects on profitability and risks retained are with respect to the longevity risk remaining with the insurer.
We propose two different ways of adjusting the benefits resulting in different risk profiles for the insurer. We also show that the annuity does not constitute a Tontine scheme but rather that there is still some risk left to the insurer. However the risk is reduced to a large degree to a model risk which is smaller and of less consequence than the longevity risk we started from.
For further information look also at Brown (2000), Dyson (1969), James (1999), Murthi (1999), Poterba (1998) and Wadsworth (2001).

Mortality Model

Since we have to predict the future mortality, we are forced to chose a model for the development of the mortality in time. Even assuming no mortality improvement at all constitutes the choice of a model, namely one where the future mortality equals the present one. In view of the fact that the product has to be kept reasonably simple and also has to find approval with the regulators, the Nolfi assumption on future mortality seems to be an adequate choice. Under the Nolfi assumption, mortality is supposed to improve exponentially over time for all ages, i. e. for each age x the mortality $q(x,t)$ t years in the future is given by

$$q(x,t) = q(x,0) \cdot \exp(-l(x) \cdot t)$$

where $l(x)$ are constant, age specific parameters. A well-known example is for instance the ERM/F2000 table of Switzerland, which we used in our sample calculations.

It would also be possible to choose other models. However we felt that the Nolfi model was the least controversial and also the one used by many practitioners. In addition, the dynamic adjustment of the parameters will make the choice of a specific model less important.

If we imagine a portfolio of n persons, we will observe each year u a certain number $N(u)$ of deaths. Most likely this number will be smaller or higher than predicted under the Nolfi Model above. We therefore propose to make the model in a simple way adjustable according to the experience. We have however to consider that the number of claims will be quite small for a reasonably sized portfolio and the scarcity (and perhaps also the quality) of the data will not warrant a constant adjustment of all $l(x)$ for each year u, e. g. the determination of $l(x, u)$. However reducing the problem to a one-dimensional one would make it statistically much more feasible. We propose to use the following approach

$$q(x, t, u) = q(x, 0) \cdot \exp(-l(x) \cdot \lambda(u) \cdot t), \qquad \lambda(0) = 1 \quad \text{(Model A)}$$

where $\lambda(u)$ is estimated on a, say, yearly basis.
For ease of notation, we will assume that the information is always known since time $u = 0$. Therefore we write in the following $q(x, t)$ for $q(x, t, u)$, where $u = 0$. Similarly, the other expressions should be interpreted.
Of course, depending on the size of the portfolio, it might make sense to estimate more parameters. Alternatively, we might have another view on how the model should be adjusted. For instance, we could also imagine the yearly adjustment to be not on the decreased $l(x)$ but on the base mortality $q(x, 0)$, i. e.

$$q(x, t) = a(t) \cdot q(x, 0) \cdot \exp(-l(x) \cdot t) \qquad \text{(Model B)}$$

We also could easily combine (Model A) and (Model B) to obtain

$$q(x, t) = a(t) \cdot q(x, 0) \cdot \exp(-l(x) \cdot \lambda(t) \cdot t) \qquad \text{(Model C)}$$

In this paper we consider in more detail (Model A). This is because we have the opinion that it is in many cases more difficult to measure the speed of mortality improvement rather than the baseline mortality at time $t = 0$. This is however not a central feature of the new annuity product but rather a choice in our example.

This model actually encompasses a wide spectrum of possible scenarios for a cohort of annuitants. For instance setting $\lambda(t) = \infty$ would model the case of zero mortality at year t. However the power of a model of course does not lie in its ability of postdiction but rather in its ability to predict the future.

Parameter Estimation

There are many choices in how to estimate the mortality improvement and it is not the intention of this paper to discuss the merits of one scheme or the other. It has to be born in mind that the estimation scheme has likely to be fixed in the contract and should therefore not be too complicated and ad-hoc. For our examples we used a simple Maximum Likelihood Estimator. Consider first the simplest case where the portfolio at the beginning of year t consist of $S(t)$ persons aged $(x+t)$, i. e. the annuity started running with a homogenous portfolio of $S(0)$ persons aged x. During year t, $D(t)$ persons die. Hence the empirical mortality was $q(x+t,t) = D(t)/S(t)$. The MLE for $\lambda(t)$ can then easily be calculated using the relation

$$D(t)/S(t) = q(x+t,0) \cdot \exp(-l(x+t) \cdot \hat{\lambda}(t) \cdot t$$

or

$$\hat{\lambda}(t) = -\frac{\log\left(\dfrac{D(t)}{S(t) \cdot q(x+t,0)}\right)}{l(x+t) \cdot t}$$

We would like to emphasise that this is of course not the only way to estimate $\lambda(t)$. The MLE scheme above considers the experience of only one year. It is most exposed to stochastic variability, however it is also the fastest reacting scheme possible. If experience from a number of years is collected we can imagine for instance a "moving window" approach where the number of deaths during the last w years is considered. Then we have the relation

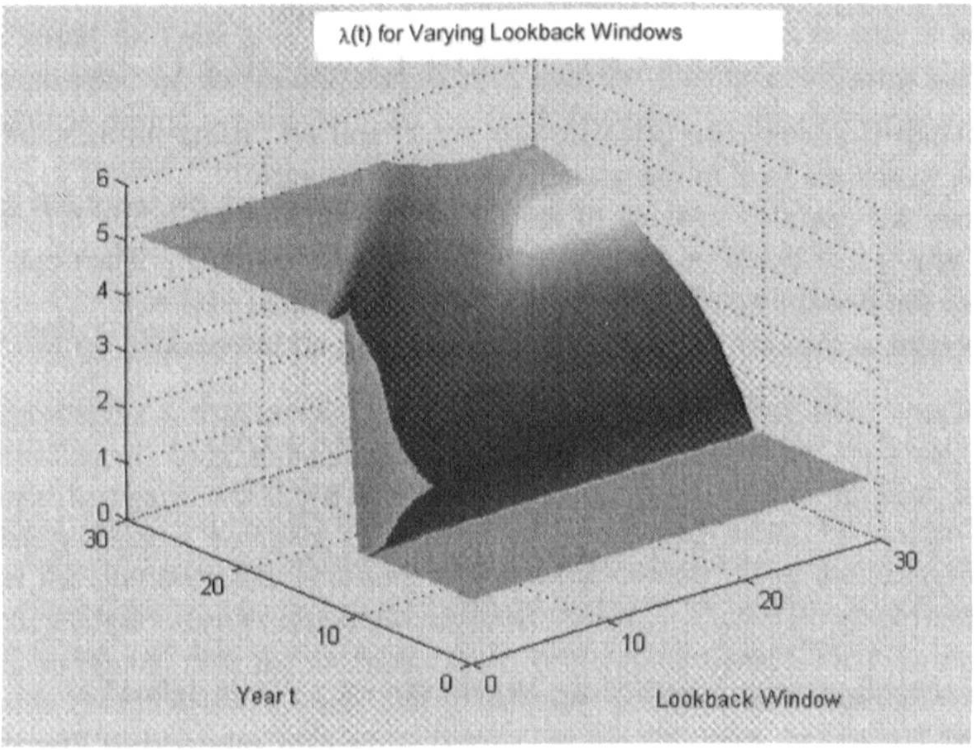
$$\prod_{j=t-w+1}^{t} [1 - q(x+j,0) \cdot \exp(-l(x+j) \cdot \hat{\lambda}(t) \cdot j)]$$
$$= 1 - \sum_{j=t-w+1}^{t} D(j)/S(t-w+1)$$

The above can not explicitly be solved for $\hat{\lambda}(t)$ but the Newton scheme converges quite fast. Another possibility would be for instance to exponentially weigh the past years.

The longer the window, the more of the past experience is included in the estimation and the less exposed to fluctuation is the estimate of $\lambda(t)$. If we assume that the adapted Nolfi model (Model A) is indeed the correct one, then taking a window as wide as possible would be most optimal. However in case that the mortality improvement is different from our assumptions, the lowered volatility of the estimate is bought with a decreased speed in adjustment. As an example, we consider the following scenario where the true $\lambda(t)$ is of the form

$$\lambda(t) = \begin{cases} 1, & t \leq 10, \\ 5, & t > 10. \end{cases}$$

In this case the MLE with window 1 year, follows $\lambda(t)$ quite closely whereas the estimator with maximal window, taking into account the data from the beginning, is slow to adjust.



Benefit Adjustment

Here also, there is no unique way of adjusting the benefit depending on the mortality estimation. In the following, we present two different ways of making the benefit payment dependent on mortality. The first method is the more efficient in the sense that more of the risk is transferred to the policyholder. The second method is conceptually simpler hence it might be easier to sell.

Method 1

Consider $a(x,t)$, the annuity immediate for a person aged x at year t. Let $r(t)$ be the payout at time t. Obviously, $a(x,t)$ is an estimate and hence depends on the available information. Assume that we are at the moment time $s \leq t$. Then the expected value of an annuity at time $t \geq s$ of a person aged x can be written concisely as $(a(x,t) \mid F_s)$, where F_s denotes all information up to time s. Normally we are most interested in $(a(x,t) \mid F_t)$, i. e. in the expected future payout at time t measured at time t corresponding to the most recent best estimate.
We consider first the well known recursion relation for $a(x+t,t)$:

$$a(x+t,t) \cdot r(t) \cdot (1+i) - (1-q(x+t,t)) \cdot r(t)$$
$$= (1-q(x+t,t)) \cdot a(x+t+1,t+1) \cdot r(t+1)$$

In a standard annuity, we calculate $a(x+t,t)$ and $r(t)$ using information up to time 0, hence we lock in our assumptions for the future.
However we can also imagine to update our information on $a(x+t,t)$ as well as on $q(x+t,t)$ based on the most up-do-date information F_t. Then one has to give up the possibility of guaranteeing a constant benefit $r(t) \equiv r$.
The benefit at the end of year $t+1$, incorporating all information up to t would be

$$r(t+1) = r(t) \frac{a(x+t,t) \cdot (1+i) - (1-q(x+t,t))}{a(x+t+1,t+1) \cdot (1-q(x+t,t))}$$

Method 2

In the second presented method the benefit at year $t+1$ is defined by

$$r(t+1) = r(1) \cdot (a(x,0) \mid F_0) \,/\, (a(x,t) \mid F_t)$$

In other words, the payment at the end of the $t+1$ year is chosen such that it would correspond to the annual benefit if the information at time t would have been known already at time 0.
This is also an effective methodology to shed mortality risk, however less so than the first proposed algorithm. It has the advantage of utmost simplicity but is less efficient than the first method since it depends only on the estimation procedure at time t, not however on the actual past. It will be less efficient if the mortality improvement changes discontinuously. Nevertheless, it is a strong contender due to its conceptual simplicity.

Remark: The situation where policyholders participate in returns in excess of the technical interest rate can easily be incorporated in the above framework. Let the true return during year t be $i(t)$ and the technical interest rate be i. Then the excess profit during year t would be

$$a(x+t,t) \cdot r(t) \cdot (i(t) - i)_+$$

We then assume that the above is a single premium for an annuity starting at age $(x+t+1)$ for a first-year benefit of

$$r'(t+1) = a(x+t,t) \cdot r(t) \cdot (i(t) - i)_+ / a(x+t+1, t+1).$$

Then this annuity generates a stochastic stream of benefits $r'(t+1+j)$, $j = 1, 2, \ldots$ as the initial annuity. This argument can then be made for each excess profit at years $t = 1, 2, \ldots$ and the effective benefit at year t will be a superposition of the initial benefit $r(t)$ and the benefit emanating from the profits made during years $t' = 1, 2, \ldots, t-1$. However, in the following we only consider annuities without profit-participation in order not to distract from the core of the topic we are interested in.

Risk Implications

It is (probably) a truism that risk cannot be destroyed, but only transformed and repackaged. In a standard annuity, the full mortality risk is born by the insurance company while the annuitant is guaranteed (in the simplest case) a completely riskless payment of, say, 1 per year until death. In our proposed product the mortality risk is shared between the annuitant and the company and the policyholder receives a random payment stream. It should be emphasised that for this additional risk, the annuitant needs to be compensated. First, he now has an upside potential in the sense that if mortality is worse than expected then his annuity will increase (as long as he belongs to the survivors, of course). Second, the premium should contain a smaller loading since the insurance company takes on less risk.

Residual Risk

Even for the proposed annuity schemes, the insurance company is still exposed to residual mortality risk due mainly to two sources: the risk of random fluctuations in the number of deaths and the risk that the actual development of mortality does not correspond to the model, even though the model risk is now greatly reduced.

To see the effects on the profitability of the product due to the residual risks, we have to stress that the results are determined also by the estimator we use for $\lambda(t)$. Depending on the lookback window we use (how many years of experience to use for the estimate) and on the estimator (MLE, robust, etc) the results will differ. On the one extreme is the case where there is no lookback, where the estimator is the most nimble but also the most volatile. The other extreme is an estimator where we take into account the whole past history, sacrificing reaction for stability. It is instructive to see what would happen in the case of a small portfolio with few deaths. In a year without deaths the predictor would assume that $q(x,t) = 0$ or equivalently, $\lambda(t) = \infty$. Hence the payout would simply be the remaining reserve times the technical interest rate (for Method 1) or 1 times the technical interest rate for Method 2. An alternative, less naive estimator would automatically extend the lookback window to the last year t' where a death occurred to obtain an estimate for $\lambda(t)$.

Stochastic Risk

Random fluctuations in the number of deaths lead to estimates for $\lambda(t)$ which are off from the true value and consequently the payouts $r(t)$ will also be randomly varying.

Model Risk

More important is the fact that the development of true mortality might not follow our simple Nolfi model. Actually, it's quite likely that this will be the case. The question is more how far away from reality the model will be and what the effects on profitability will be. To obtain an idea of this question we simulated a number of extreme cases to see what the effects were. The most extreme, and probably the most unrealistic one is the case where true mortality is zero, e. g. nobody will die.

Consider the last case with zero lookback and no deaths at all. Then the payout at the end of year 1 is the given reserve multiplied by the technical interest rate. If actual interest equals technical one, the profit ratio of the annuity with Method 1 comes out as nearly 1 even in this extreme case.

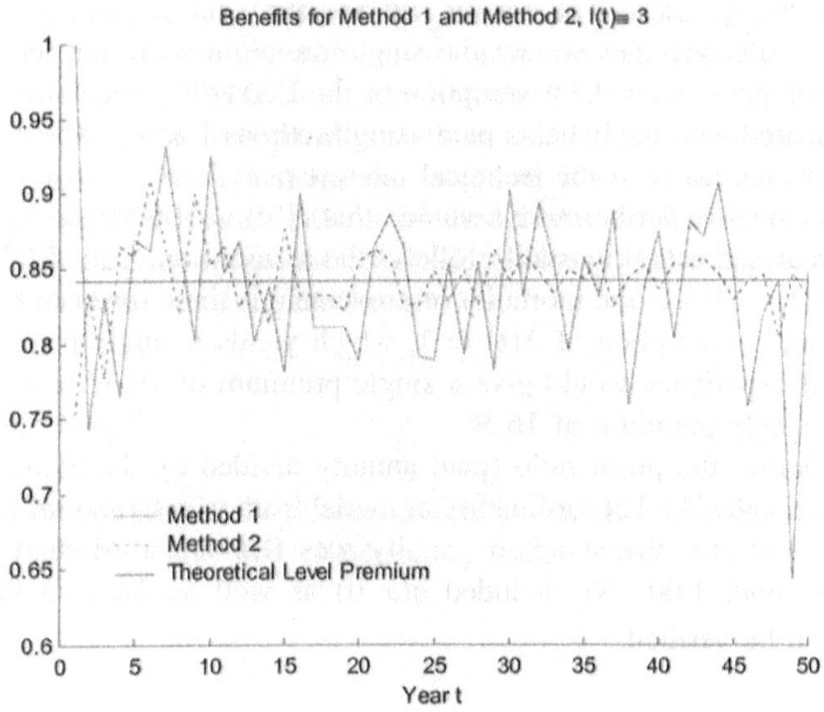
However it can happen that mortality first develops unfavourably, causing big benefit payments. Later drops in mortality can then cause a loss for the issuers even under Method 1.

Examples

In the following, we keep a number of assumptions fixed. We always consider single premium annuities where the premium is paid at age 65 and the annuity starts at the same age. The benefits are paid at the end of the year. We assume that the number of deaths during a year follows a Poisson law. We also assume a technical interest rate of 2.5 %. The benefit will be yearly and payable at the end of the year and in all examples we set the single premium such that an annual benefit of 1 can be paid under the assumption of the ERM2000 generation table. This is then compared with the benefits paid using Methods 1 and 2. The discount rate is always set identically to the technical interest rate.
In the example below, we furthermore assumed that $S(0) = 10000$, the lookback window is one year and actual mortality follows the assumption of the ERM2000 tables but with $\lambda(t) = 3$, i. e. the mortality improvement is three times as fast. We start with the initial assumption of $\lambda(t) = 1$, which yields a single premium of 16.56. The actual experience would give a single premium of 19.67 or a benefit of 0.842 for the single premium of 16.56.
In the example below, the profit ratio (paid annuity divided by the theoretically correct benefit, i. e. calculated according to our model from which random samples have been generated) for the standard annuity was 0.844, for Method 1 and Method 2 it was both 1.00. We included $q(x, 0)$ as well as $\lambda(x)$ so that the values of $\hat{\lambda}(t)$ can be verified.

Year t	$S(t)$ BoY	$D(t)$ EoY	$q(x+t,0)$	$l(x+t)$	$\hat{\lambda}(t)$	Benefit Method 1	Benefit Method 2
1	10000	74	0.00833	0.02672	1.000	1.000	0.753
2	9926	78	0.00926	0.02648	4.445	0.741	0.835
3	9848	74	0.01031	0.02622	3.104	0.831	0.777
4	9774	85	0.01149	0.02593	4.019	0.765	0.865
5	9689	87	0.01281	0.02563	2.682	0.870	0.858
10	9205	104	0.02232	0.02382	2.207	0.926	0.852
15	8662	149	0.03880	0.02151	3.430	0.780	0.876
20	7964	182	0.06638	0.01893	3.300	0.789	0.855
30	5722	289	0.17725	0.01392	2.714	0.903	0.842
40	2538	285	0.38416	0.00997	2.862	0.874	0.836
50	454	92	0.63120	0.00713	2.936	0.848	0.829

Graphically, the evolution of the benefits under Method 1 and 2 looks as follows. Note that the standard deviation of benefits under Method 1 is on average 0.074 for the above parameters and 0.048 for Method 2. Hence the policy holder is, as expected, exposed to more risk.



In the following, we show the behaviour of the annuities for different scenarios. We compare Method 1 and Method 2 with the standard annuity. To this end we determine the profit ratio of the benefits.

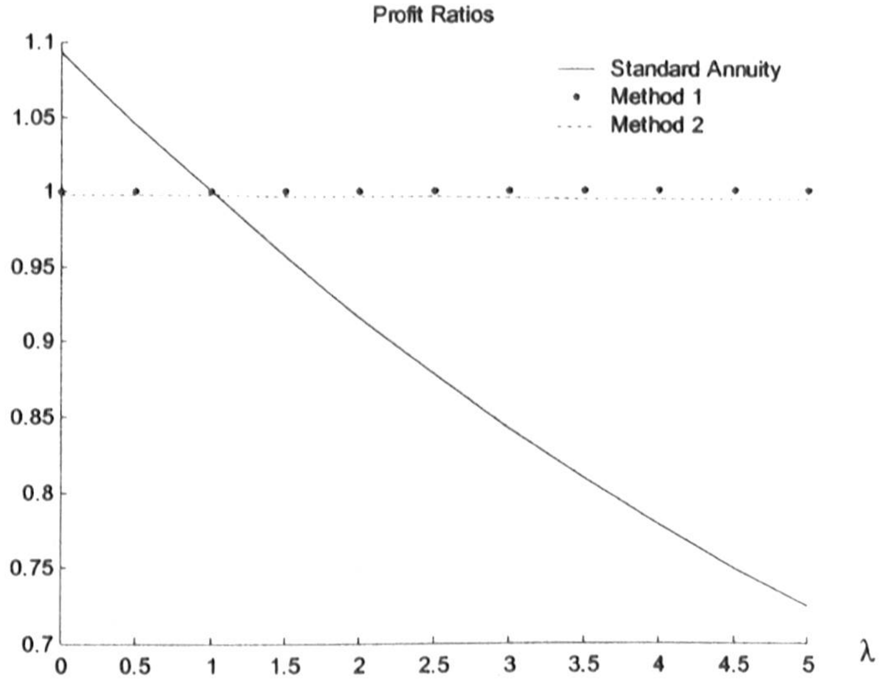
We consider two classes of scenarios. The first one is when mortality actually follows the assumption of our model, i. e. when $\lambda(t) = \lambda =$ constant. The second, much bigger class, is the case when mortality follows a path outside the scope of our model.

Behaviour Without Model Risk

Recall that our model was a generalisation of the Nolfi assumption:

$$q(x,t) = q(x,0) \cdot \exp(-l(x) \cdot \lambda(t) \cdot t)$$

In the following figure we show the profit ratios depending on the mortality improvement factor $\lambda(t) = \lambda$. We generated $n = 1000$ simulations and assumed that the number of deaths are Poisson distributed.



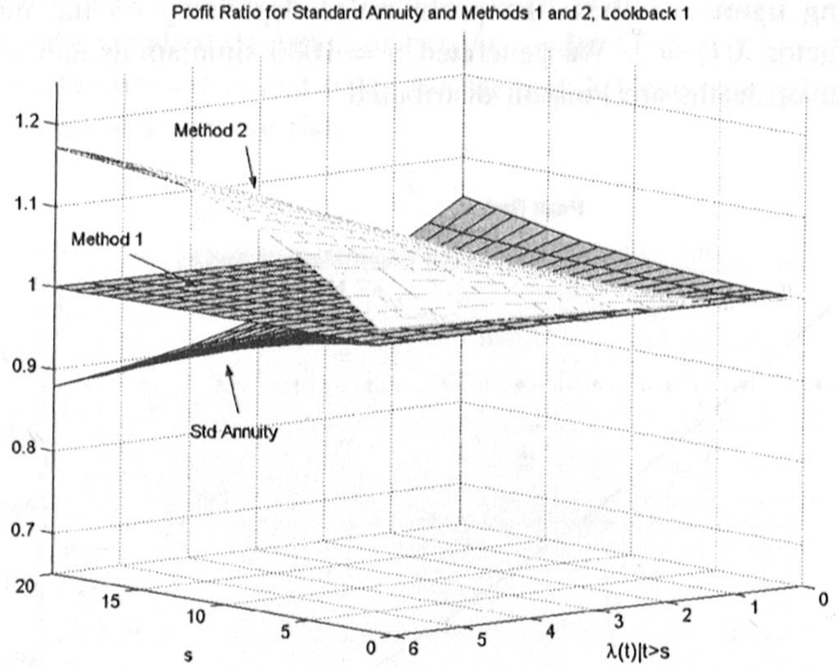
As can be seen above, the expected profit ratio for both Method 1 and 2 are nearly perfectly 1.0.

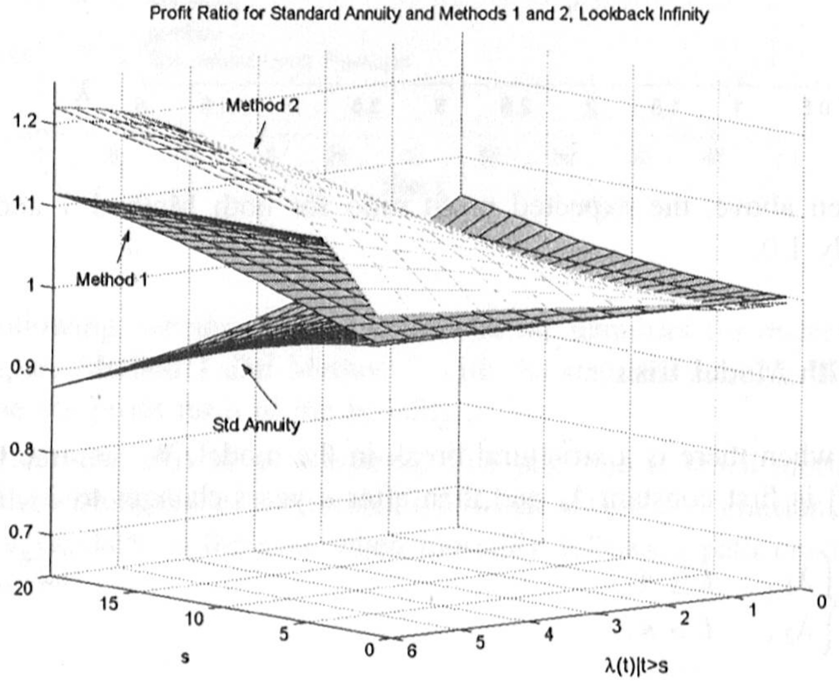
Behaviour With Model Risk

First we look when there is a structural break in the model. We assume that the parameter $\lambda(t)$ is first constant λ_1 and then after s years changes to λ_2, i. e.

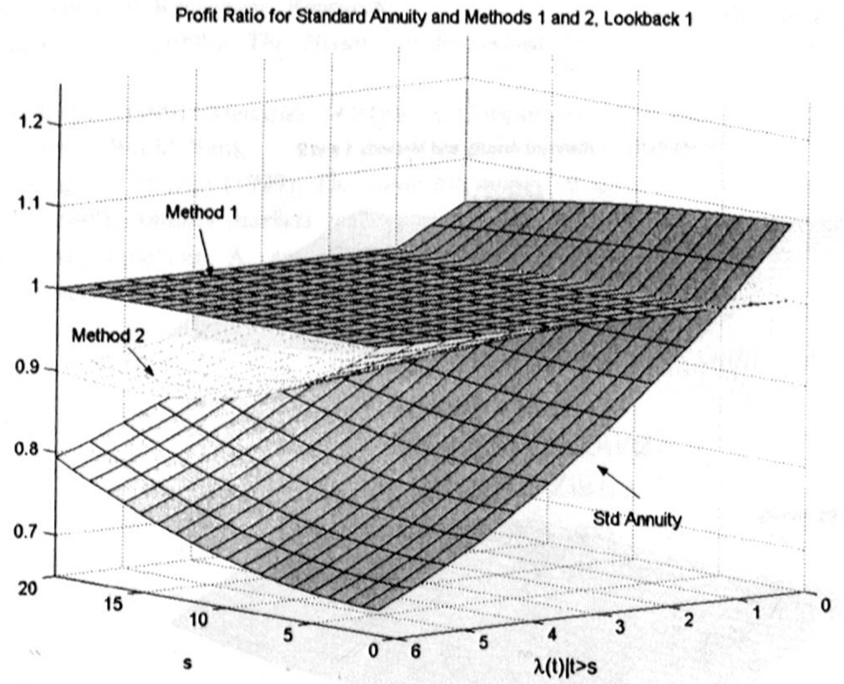
$$\lambda(t) = \begin{cases} \lambda_1 , & t \leq s , \\ \lambda_2 , & t > s . \end{cases}$$

We consider first the case when λ_2 equals 1 and λ_1 is going from 0 to 6. We let s go from 1 to 20 years. The figures below show the profit ratio for the standard annuity as well as for Methods 1 and 2 for a lookback of 1 and for a lookback of infinity.


Profit Ratio for Standard Annuity and Methods 1 and 2, Lookback 1
Method 2
Method 1
Std Annuity
1.2
1.1
1
0.9
0.8
0.7
20
15
10
5
0
s
6
5
4
3
2
1
0
λ(t)|t>s

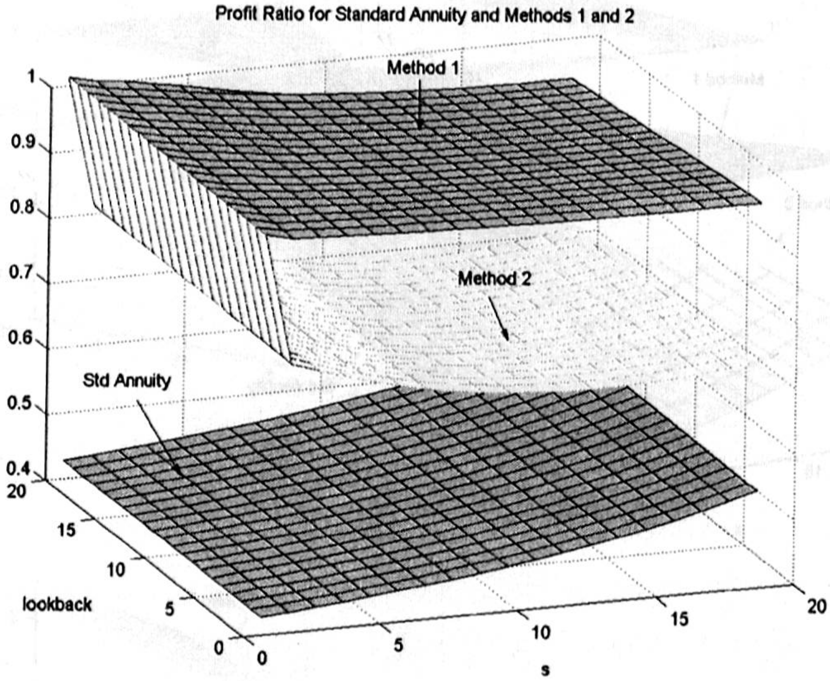

Profit Ratio for Standard Annuity and Methods 1 and 2, Lookback Infinity
Method 2
Method 1
Std Annuity
1.2
1.1
1
0.9
0.8
0.7
20
15
10
5
0
s
6
5
4
3
2
1
0
λ(t)|t>s

Next we consider the case when $\lambda_1 = 1$ and λ_2 is going from 0 to 6.



The above graphs show clearly that the standard annuity can not react on systematic changes in the parameters λ_1 and λ_2. On the other hand Method 1 and, with some delay, also Method 2 perfectly deal with the situation.

Next we consider the case when mortality will drop to zero after s years. Whilst somewhat unrealistic, it is instructive to see how Methods 1 and 2 cope with such extreme cases. Again we vary s from 1 to 20 years and we use lookback windows from 1 to 20 years.



Under this extreme scenario a standard annuity would leave an insurance company with a loss ratio of around 50 %, Method 2 would lead to a loss ratio of around 30 % and Method 1 with loss ratio of only around 5 % over the whole period.

References

Brown, J. R. (2000) *How should we insure longevity risk in pensions and social security?*; Harvard University Center of Retirement Research

Dyson, EJW et. al. (1969); *The History of Individual Annuity Contracts*; Insurance Institute of London

James and Vittas (1999), *Annuities Markets in Comparative Perspective: Do Consumer Get Their Money's worth?*; World Bank

Murthi, Orszag and Orszag (1999); *The value for money of annuities: theory, experience and policy*

Poterba, J. (1998); *Annuity markets and retirement security*; MIT

Wadsworth, M., Findlater, A. and Boardman T. (2001); *Reinventing Annuities*, The Staple Inn Actuarial Society

Prof. Dr. Herbert Lüthy
Swiss Re
Mythenquai 50/60
CH-8022 Zürich
E-mail address: Herbert_Luethy@swissre.com

Summary

We present a form of annuity with reduced mortality and longevity risk for the insurer. This is achieved by comparing the mortality experience of the annuity portfolio with the expected experience regularly, and adjusting the benefits accordingly. We show that this can be done in a way that the longevity risk is substantially reduced.

The intention of this paper is to introduce a new form of annuity insurance that will widen the spectrum of products offered to annuitants. It will enable policyholders to obtain cheaper annuities, since the risk for the insurance company is reduced, and also to benefit from a favourable mortality experience.

We also feel that a less risky product (from the insurer's viewpoint) will enlarge the field of issuers and thereby lead to a more competitive and efficient market for annuities.

Finally, we are convinced that an unbundling of the different risks (mortality, interest rate, etc.) is of general benefit, since in this way annuities can be tailored to the consumer's needs. Starting with an annuity without mortality or investment guarantees one would add in a modular way any additional features.

Zusammenfassung

Wir stellen eine Art der Rente vor, die das Sterblichkeits- und Langleberisiko für den Emittenten reduziert. Dies wird erreicht, indem man regelmässig die Mortalitätserfahrung des Rentenportfeuilles mit der erwarteten Erfahrung vergleicht und die Leistungen entsprechend anpasst. Dadurch kann das Langleberisiko substantiell vermindert werden.

Die Absicht dieses Artikels ist es, eine neue Form einer Rente vorzustellen, um die Palette von Rentenprodukten zu erweitern. Der Versicherungsnehmer kann dadurch kostengünstigere Produkte erwerben, da das Risiko für die Versicherung verringert ist und der Käufer von der Mortalitätsentwicklung profitieren kann.
Wir sind der Meinung, dass ein für die Versicherung risikoärmeres Produkt die Anzahl der Rentenanbieter vergrössern wird und dadurch ein kompetitiverer und effizienterer Markt entsteht. Wir sind auch überzeugt davon, dass die Zerlegung der Risiken (Sterblichkeit, Zins, usw.) zum Vorteil aller gereichen wird. Auf diese Art wird es möglich, dem Kunden massgeschneiderte Produkte zu offerieren: Angefangen von einer Rente ohne Sterblichkeit- und Renditengarantie können eben diese Garantien im Baukastensystem dazugekauft werden und einzeln bewertet werden.

Résumé

Nous présentons un type de rente avec un risque de mortalité et de longévité réduit pour l'émetteur. Nous atteignons cet objectif, d'abord en comparant l'espérance de la mortalité et la mortalité réelle d'un même portefeuille, ensuite en adaptant les prestations. Ceci peut être fait de telle sorte que le risque de longévité soit substantiellement réduit.
Ce document a pour but de présenter une nouvelle forme de rente qui viendra peut-être élargir l'offre actuelle des diverses rentes. Le preneur d'assurance pourra alors se voir octroyer des rentes moins onéreuses, dans la mesure où le risque sera moins grand pour la compagnie d'assurances et où l'assuré pourra bénéficier d'une évolution favorable de la mortalité.
Nous espérons également qu'un produit à composante risque réduite contribuera à augmenter le nombre de pourvoyeurs de rentes, avec pour résultat, un marché de la rente plus compétitif et plus efficace. En outre, nous sommes persuadés de l'intérêt que présente la partition du risque en ses composants (risque inhérent à la mortalité, au taux d'intérêt, ...) tant pour l'assureur que pour le preneur d'assurance : les rentes pourraient être ainsi adaptées au mieux aux attentes des clients potentiels dans la mesure où l'on partirait d'une rente sans garantie du risque de mortalité ni garantie financière, à laquelle seules les garanties souhaitées par le prospect seraient ajoutées une à une.