Insurance: Mathematics and Economics 29 (2001) 299–318
INSURANCE MATHEMATICS & ECONOMICS
www.elsevier.com/locate/econbase

Mortality derivatives and the option to annuitise

Moshe A. Milevsky [a,*], S. David Promislow [b]

[a] *Finance Area, Schulich School of Business, York University, Toronto, Ont., Canada M3J 1P3*
[b] *Department of Mathematics and Statistics, York University, Toronto, Ont., Canada M3J 1P3*

Received 1 November 2000; received in revised form 1 May 2001; accepted 29 August 2001

Abstract

Most US-based insurance companies offer holders of their tax-sheltered savings plans (VAs), the long-term option to *annuitise* their policy at a pre-determined rate over a pre-specified period of time. Currently, there is approximately one trillion dollars invested in such policies, with guaranteed annuitisation rates, in addition to any guaranteed minimum death benefit. The insurance company has essentially granted the policyholder an option on *two* underlying stochastic variables; future interest rates and future mortality rates. Although the (put) option on interest rates is obvious, the (put) option on mortality rates is not. Motivated by this product, this paper attempts to value (options on) mortality-contingent claims, by stochastically modelling the future *hazard-plus-interest* rate. Heuristically, we treat the underlying life annuity as a defaultable coupon-bearing bond, where the default occurs at the exogenous time of death. From an actuarial perspective, rather than considering the force of mortality (hazard rate) at time t for a person now age x, as a number $\mu_x(t)$, we view it as a random variable forward rate $\tilde{\mu}_x(t)$, whose expectation is the force of mortality in the classical sense ($\mu_x(t) = E[\tilde{\mu}_x(t)]$). Our main qualitative observation is that *both* mortality and interest rate risk can be hedged, and the option to annuitise can be priced by locating a replicating portfolio involving insurance, annuities and default-free bonds. We provide both a discrete and continuous-time pricing framework.

Keywords: Annuity; Financial options; Longevity risk; Hazard rates

"*If life spans increase during the accumulation period, the guaranteed purchase rates may produce larger monthly payments than would the current or immediate annuity rates when the policyholder annuitises. On the other hand, if the rates being offered to individuals who purchase immediate annuities are more favourable than the guaranteed purchase rates, the policyholder can annuitise using the immediate annuity rates. In effect, the policyholder is in a win/win situation . . .*" Source: National Association for Variable Annuities, 1997.

1. Introduction and motivation

Most US-based insurance companies offer holders of their tax-sheltered savings plans — also known as variable annuity contracts — the long-term option to annuitise their policy at a pre-determined rate over a pre-specified

* Corresponding author. Tel.: +1-416-736-2100; fax: +1-416-225-5035.
E-mail addresses: milevsky@yorku.ca (M.A. Milevsky), promis@yorku.ca (S. David Promislow).

PII: S0167-6687(01)00093-2

period of time. If annuity rates at the time of annuitisation are more favourable than the contractually specified value, the policyholder can demand current rates or go elsewhere. Conceptually, this call option on annuity purchase factors can be viewed as the right, but not the obligation, to purchase a fixed immediate life annuity, for a deterministic strike price during the life of the contract. The company has essentially granted the policyholder an option on two underlying stochastic variables; future interest rates and future mortality rates. Interest rates may decrease — relative to the guaranteed return — thus putting the insurance company at investment risk. As well, mortality patterns can improve leaving the insurance company exposed to unanticipated longevity risk. We believe this risk is not trivial since at least one trillion dollars in retirement savings are invested in these products, according to Moody Investor Services. [1] The option to annuitise — and the implicit risk exposure — should be contrasted with a spot transaction of purchasing a single premium immediate annuity (SPIA) or a forward transaction of purchasing a single premium deferred annuity (SPDA); both of these immediately lock-in irreversible long-term rates. In practice, insurance companies tend to protect themselves by guaranteeing low interest rates combined with 'aggressive' mortality improvement projections for the annuitant population.

In this paper, we attempt to value mortality-contingent claims, by stochastically modelling the *hazard-plus-interest* rate. Heuristically, we treat the underlying life annuity as a defaultable coupon-bearing bond, where the default occurs at the exogenous time of death. In practice, the option to annuitise is an American-style contingent claim on a corporate bond. From an actuarial perspective, rather than considering the force of mortality (hazard rate) at time t for a person now age x, as a number $\mu_x(t)$, we view it as a random variable forward rate $\tilde{\mu}_x(t)$, whose expectation is the force of mortality in the classical sense ($\mu_x(t) = E[\tilde{\mu}_x(t)]$). In our model, we will assume that the forward term structure of interest-plus-mortality is modelled in continuous time by using a Gompertz specification for the *expected* force of mortality.

Our main qualitative observation can be summarised as follows. Aside from the interest rate risk — which is well understood in the financial literature — there are two distinct categories of mortality risk that an insurance company faces when selling an option to annuitise. The first category we label as 'small sample' risk. It reflects the chance that any particular option holder is healthier than average, anti-selection not withstanding. When faced with such a client, the insurance company is confronted with a payment stream that is longer than originally expected based on annuitant mortality rates. However, actuarial theory has long established that this particular risk can be eliminated — and therefore should not be priced — by selling enough identical policies and taking advantage of the law of large numbers. If enough policies are sold, the realisation will converge to the expected. The second risk is more subtle. It is the risk that the insurance company overestimated the population's force of mortality. This longevity risk cannot be hedged by appealing to the law of large numbers and selling more annuities or options to annuitise. Of course, the insurance company prices immediate annuities with an expected (dynamically projected) improvement rate. And thus, aggregate mortality may indeed be worse or better than expected. We, therefore, argue that this second type of risk can be hedged by selling more life insurance policies. The risks offset each other. If mortality rates improve, then the life insurance book will show unexpected profits.

Table 1 provides an indication of this second type of risk, by displaying the change in *Society of Actuaries* individual annuitant (period) mortality tables over the last three decades. The symbol ${}_{(t-55)}p_{55}$ denotes the probability that a 55-year-old will survive to age t. Clearly, mortality has been improving. We seek to model this phenomena using a stochastic process.

In practice, the insurance company may not be able to sell insurance to the same group that is buying annuities — the young tend to buy life insurance, while the old purchase annuities — but this is a mere question of implementation. (And, insurance to pay estate taxes is marketed at older individuals.) In complete markets, the value of the option is the price of the residual risk that cannot be eliminated by selling more endowments and life insurance policies. We will shortly present a binomial model that should help explain this concept.

[1] See Milevsky and Posner (2001) for a description of the GMDB options which are embedded in Variable Annuity products.

Table 1
Period annuitant mortality tables: change over time in the value of $(t-55)P55$

t	1971 IAM		1983 IAM		1996 IAM	
	F	M	F	M	F	M
55	1.00	1.00	1.00	1.00	1.00	1.00
60	0.976	0.952	0.982	0.966	0.985	0.974
65	0.938	0.886	0.956	0.919	0.962	0.937
70	0.889	0.799	0.914	0.848	0.926	0.880
75	0.812	0.682	0.849	0.742	0.899	0.791
80	0.689	0.530	0.745	0.596	0.775	0.663
85	0.504	0.353	0.586	0.415	0.628	0.496
90	0.281	0.181	0.379	0.234	0.427	0.313
95	0.103	0.056	0.181	0.100	0.221	0.154
100	0.026	0.007	0.059	0.028	0.082	0.055

1.1. *Agenda*

The remainder of this paper is organised as follows. Section 2 presents a discrete time model which should provide the basic intuition for the hedging argument. Section 3 provides a continuous-time model of the hazard-plus-interest rate, which is based on the Cox et al. (1985) specification for interest and a Gompertz specification for mortality. This is followed by some numerical examples. Section 4 concludes the paper. Finally, although we have motivated this investigation by alluding to the *options to annuitise* that are embedded in VA contracts, we must make absolutely clear that this paper is for the most part concerned with European-style options on a mortality-contingent claim that pays one lump sum upon surviving a pre-specified period. This is also known as an endowment policy — in contrast to a traditional life annuity — and is quite similar to a zero-coupon bond. The options contained in VA policies are: (i) American style, (ii) on coupon-bearing bonds, and (iii) with a variable notional principal, all of which complicate the modelling substantially.

2. Discrete time

In this section, we illustrate the basic ideas by means of some simple discrete time examples covering a small number of periods. Our main objective is to illustrate that *mortality improvement risk* can be hedged, and therefore the option to purchase a mortality contingent claim has a unique and quantifiable price. We assume throughout the section that all activity takes place on a yearly basis, and insurance benefits will be paid at the end of the year of death. We clearly ignore expenses, profits and other administrative charges and thus assume that everything is presented on a net basis. We assume in addition that mortality rates used for life insurance are the same as those used for annuities. Of course, in practice this is not true, but it does not impact the hedging argument other than by virtue of the ratio. Once again, in order that our theory should parallel the well-developed results for bond options, we will concentrate at first on the case of options on pure endowments. These are contracts that provide a single payment at a specified time, if the purchaser is then alive. A life annuity can of course be viewed as a basket of several pure endowments. There is a complication, however, in that the sum of option prices for all these pure endowments will in general be a strict upper bound to the option price for the annuity. In the latter case, the option holder must exercise either all or none of the pure endowment options. It is possible, however, that if interest and mortality move in opposite directions (relative to their effect on annuity prices) some pure endowment options would be in the money at expiry, and some would not. We will elaborate with examples in Section 2.5.

2.1. General notation

We begin with set Ω, representing all possible states of nature, equipped with a filtration to denote the information flow. This is an increasing sequence $\mathcal{F}_n$ of σ-algebras, where $\mathcal{F}_0$ is the trivial algebra $\{\Omega, \emptyset\}$. We are given an interest rate term structure of default-free discount bonds. That is, for each discrete $k \leq n$ we have random variables $D(k, n)$ which are $\mathcal{F}_k$ measurable, together with a risk neutral probability measure Q on Ω such that:

$$D(k, n) = D(k, k+1)E_Q[D(k+1, n)|\mathcal{F}_k] \quad \forall k < n. \tag{1}$$

The term $D(k, n)$ represents the price, at time k, of a 1 unit zero-coupon bond maturing at time n. Throughout the paper we assume that the risk neutral measure, Q, is used for pricing.

In addition to the default-free market, we need a 'term structure' for mortality rates. This is most easily expressed in terms of survival probabilities which are analogous to the bond prices as given above. For each $k \leq n$ let $p_x(k, n)$ denote the probability that an individual in our cohort — who is currently aged x — will survive to time n conditional upon surviving to time k. In the traditional actuarial approach this is a constant and would be written in standard notation as ${}_{n-k}p_{[x]+k}$. Consistent with our new treatment of mortality as given in the introduction, $p_x(k, n)$ is no longer a constant but rather a $\mathcal{F}_k$ measurable — random variable. Moreover, we postulate as an axiom the analogous relationship we used for bond prices. That is, for the same probability measure Q as above.

$$p_x(k, n) = p_x(k, k+1)E_Q[p_x(k+1, n)|\mathcal{F}_k] \quad \forall k < n. \tag{2}$$

Under the traditional approach when probabilities are constant, the expectation sign can be removed and this reduces to a well-known actuarial identity.

One way to see that formula (2) is reasonable is to notice that at an interest rate of 0, $p_x(k, n)$ will denote the price at time k of a 1 unit pure endowment maturing at time n. More generally, see formula (4).

To gain further insight into Eq. (2), consider the following example. Suppose we have an individual age 60, and we assume that there is a probability of 0.8 that this person will survive to age 61. In other words, $p_{60}(0, 1) = 0.8$ (a constant, being $\mathcal{F}_0$ measurable). We do not, however, know what $p_{60}(1, 2)$ is, as this could depend on random events. This is analogous to stochastic interest rate models, where the interest rate for the current period is known, but the rates for the following periods are not known.

Suppose that two possible states of nature could materialise over the following year. In the first case, a hoped for medical breakthrough materialises and the actual value of $p_{60}^{(1)}(1, 2) = 0.8$. In the second, the medical breakthrough does not materialise and the value of $p_{60}^{(2)}(1, 2) = 0.7$. (The superscripts are used to 'count' the possible state of nature realisations.) Now the traditional mortality basis used by actuaries would show a single number for the probability that a person will live to age 62, conditional upon reaching age 61. Suppose that this was 0.75. Our point of view is that the actuary would then be in effect assigning equal probabilities to each of the two possible states, and that the 0.75 represents the expectation $E_Q[p_{60}(1, 2)]$.

Note by Eq. (2) we would then have that $p_{60}(0, 2) = (0.8)(0.75) = 0.6$. Looking at this from another point of view, if we start with the figure of 0.6 for $p_{60}(0, 2)$, we can use (2) to recover

$$E_Q[p_{60}(1, 2)] = \frac{p_{60}(0, 2)}{p_{60}(0, 1)} = 0.75.$$

This shows that if one postulates probabilities for survival of more than 1 year at time 0, one is in effect choosing the probabilities of the various states of mortality improvement. Under our approach, the traditional rates used by actuaries are really "forward rates" exactly analogous to a forward interest rate implied by existing bond prices.

We further postulate, quite naturally, that interest rates and mortality rates are independent. That is, for all $k \leq n, r \leq s$, the variable $D(k, n)$ and $p_x(r, s)$ are independent.

Finally, for all $k \leq n$ we let

$$\Lambda_x(k, n) = D(k, n)p_x(k, n), \tag{3}$$

which is $\mathcal{F}_k$ measurable, and denotes our pure endowment contract. In other words, $\Lambda_x(k, n)$ is the (random variable) price at time k that would be paid by an individual (currently age x) surviving to that time, for a contract paying 1 unit at time n, if alive. It follows from formulas (1) and (2) and the independence that

$$\Lambda_x(k, n) = \Lambda_x(k, k+1) E_Q[\Lambda_x(k+1, n)|\mathcal{F}_k] \quad \forall k < n. \tag{4}$$

Our ultimate objective is to price a call option to acquire a pure endowment at time k, that pays 1 unit at time n, if alive. Our notation is $C_x(k, n|\Lambda)$, where Λ is the strike price of the call option.

2.2. Replicating and hedging with deterministic interest

In this section, we use a simple example to show that an option on a mortality contingent claim can be replicated by the purchaser, or hedged by the issuer, of the option. We will start the analysis by assuming that interest rates are deterministic.

Suppose that at time 0 an individual, now age 60, would like to receive a single payment of 1 unit at age 62 if they are then alive. This is a two period pure endowment contract. Suppose they purchase this from an insurance company which is assuming mortality as given in the example above. That is $p_{60}(0, 1) = 0.8$ and $p_{60}(1, 2) = \{0.8 \text{ or } 0.7\}$ with equal probabilities. See Fig. 1 for details, where we display $q = 1 - p$, the probability of death. We assume a constant interest rate of 0, so the endowment price is 0.6.

Now, an alternative for the individual would be to wait a year and purchase the pure endowment contract if they are then alive. Why buy it now if they forfeit nothing for death during the year? Of course, instead of paying 0.6, they will pay either $\Lambda_{60}^{(2)}(1, 2) = (1 - 0.30) = 0.7$ or $\Lambda_{60}^{(1)}(1, 2) = (1 - 0.20) = 0.8$, assuming they are alive and depending on which state of nature occurs. Suppose that as another alternative the insurer offers them the option to purchase this contract at age 61 at a guaranteed price of 0.75. That is, if the improvement occurs, they will only have to pay 0.75 rather than 0.8. If the improvement does not occur, they will not exercise the option and instead, they will pay the prevailing price of 0.7. The intrinsic value of the option at maturity will then be either 0.05 or 0. Our intention is to 'price' the option to purchase the pure endowment contract. Using our previous notation, we are looking for $C_{60}(1, 2|0.75)$

We can calculate this in a standard way by deriving a replicating portfolio. The basic difference between this situation and the usual one with bond options is that we must consider mortality contingent assets, where the payoff depends on whether a certain individual is alive or not. A fundamental set of assets that we consider for this purpose are the underlying n year pure endowments, which we will just denote by $\mathbf{E}_n$. That is, $\mathbf{E}_n$ pays 1 unit at time n if the individual is alive at time n, and it pays zero if the individual dies prior to time n. We will show that in this example the option as described is equivalent to the portfolio which consists of 0.5 units of a 2 year pure endowment combined with a short-sale of 0.35 units of a 1 year pure endowment. Symbolically, let $\mathbf{P}$ denote the option. (We use

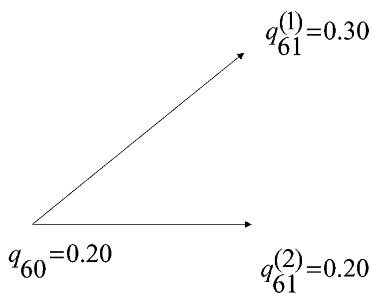


Fig. 1. One period model for stochastic mortality rates.

boldface characters to denote contingent claims as opposed to a numerical quantity.) Then

$$\mathbf{P} = 0.5\mathbf{E}_2 - 0.35\mathbf{E}_1. \tag{5}$$

To see this, suppose an individual holds the portfolio on the right-hand side above, and is alive at time 1. They must pay 0.35 at time 1 to discharge the liability on the 1 year pure endowment. If the mortality improvement has occurred, they can purchase the additional 0.5 units of income for a price of 0.4, making a total outlay of 0.75. If the improvement has not occurred, they purchase the additional 0.5 units of income for a price of 0.3, making a total outlay of 0.7. They have, therefore, replicated the option. The no arbitrage value of the option must then be the time zero cost of this portfolio which is $0.5(0.6) - 0.35(0.8) = 0.02$.

There is of course a practical problem with this replicating procedure as described, since it is not normally feasible to directly sell short a pure endowment on one's life. This is easily remedied, however, since such a short sale can be manufactured. Precisely, let $\mathbf{B}_n$ denote a 1 year, 1 unit zero-coupon bond. Let $\mathbf{I}_n$ denote a life insurance contract, which pays 1 unit at time n provided the individual dies within n years. This of course is not a standard contract, but at deterministic interest it is equivalent to a n year term insurance policy with a variable death benefit equal to the price of zero-coupon bond maturing at time n. Then

$$\mathbf{E}_n = \mathbf{B}_n - \mathbf{I}_n \tag{6}$$

as each side of the above equation denotes a contingent asset which pays 1 unit if and only if the individual is alive at the end of n years.

A short position of $\mathbf{E}_1$ is, therefore, the same as a long position in $\mathbf{I}_1 - \mathbf{B}_1$, which is nothing more than an *insured loan*. In our example, the individual borrows 0.35, pays 0.07 to insure, leaving proceeds of 0.28. Together with the 0.02 option price this purchases a 0.5 unit pure endowment at age 62. It is of interest to compare this with the replicating portfolio in the classical Cox–Ross–Rubinstein one period binomial framework. At time zero we buy a certain number of units of the underlying security as a hedge, and finance this by selling bonds short, arranging at the same time insurance to cover the short position in the event of death in the first year.

We now look at matters from the point of view of the insurer who writes the option, and show how they can hedge their position. The insurer, of course, has a built in way to hedge annuity options, namely by selling life insurance. Indeed, it is intuitively obvious, and follows formally from Eq. (6), that an insurer can take a long position in pure endowments by selling insurance, and thereby hedge against possible mortality improvements in pure endowments that they have sold. Similarly, they can hedge against *options* on pure endowments by selling a ratio of H units of insurance for each 1 unit of option. In our particular 1 year example it is clear that $H = \frac{1}{2}$, and for each unit of option sold, the insurer should sell $\frac{1}{2}$ a unit of 2 year term insurance. The resulting portfolio, from the combination of these, will be $-\mathbf{P} - 0.5\mathbf{I}_2$ (in view of the zero interest rates) and by substituting from (5) and (6) this equals $0.35\mathbf{E}_1 - 0.5\mathbf{B}_2$. This eliminates the term $\mathbf{E}_2$ and therefore hedges the risk arising from the uncertain mortality in year 2.

This must be interpreted in the proper way, however, since the contracts are not on the same lives. The purpose of the example is to isolate the risk inherent in the option and not the *small sample risk* that we discussed in the introduction. We postulate that this latter risk is hedged in the usual way, by selling a sufficiently large number of contracts. In other words, we assume that if the insurer has a long position of N units of $\mathbf{E}_2$ and an effective short position of N units of $\mathbf{E}_2$, then the mortality experience will be the same among both groups of lives and the insurer will have a net payment of zero on the totality of these contracts.

We can easily write down general formulas for this one period, deterministic interest case as follows. Let

$$H = \frac{\Lambda_{60}^{(1)}(1,2) - \Lambda}{\Lambda_{60}^{(1)}(1,2) - \Lambda_{60}^{(2)}(1,2)}, \tag{7}$$

provided that $\Lambda_{60}^{(2)}(1,2) \leq \Lambda \leq \Lambda_{60}^{(1)}(1,2)$. We define $H = 1$ if $\Lambda < \Lambda_{60}^{(2)}(1,2)$, and $H = 0$ if $\Lambda > \Lambda_{60}^{(1)}(1,2)$. (This latter condition is not possible in our present one period example, but we will encounter it later when we complicate the model.) The option portfolio is given by

$$\mathbf{P} = H[\mathbf{E}_2 - \Lambda_{60}^{(2)}(1,2)\mathbf{E}_1]. \tag{8}$$

Taking expectations

$$C_{60}(1,2|\Lambda) = H[\Lambda_{60}(0,2) - \Lambda_{60}^{(2)}(1,2)\Lambda_{60}(0,1)]. \tag{9}$$

The insurer hedges the option by writing H units of insurance for every unit of option income. From (4)

$$\Lambda_{60}(0,2) = \Lambda_{60}(0,1)E_Q\Lambda_{60}(1,2),$$

and we can substitute for H in Eq. (9) to represent the price of the call option as

$$C_{60}(1,2|\Lambda) = \left(\frac{E_Q[\Lambda_{60}(1,2)] - \Lambda_{60}^{(2)}(1,2)}{\Lambda_{60}^{(1)}(1,2) - \Lambda_{60}^{(2)}(1,2)}\right)\Lambda_{60}(0,1)(\Lambda_{60}^{(1)}(1,2) - \Lambda). \tag{10}$$

We interpret this as follows. Let π be the probability under Q of state 1 occurring. Then, assuming zero interest, $E_Q[\Lambda_{60}(1,2)] = \pi p_{60}^{(1)}(1,2) + (1-\pi)p_{60}^{(2)}(1,2)$, $\Lambda_{60}^{(1)}(1,2) = p_{60}^{(1)}(1,2)$, and $\Lambda_{60}^{(2)}(1,2) = p_{60}^{(2)}(1,2)$, so the first term in the last line reduces to just π, the probability under Q of mortality improvement. This is then multiplied by the expected payoff from the call option, and discounted back to time 0 with interest and survivorship.

The extension of the procedure above to periods of longer than 1 year is straightforward. The replicating and hedging procedures can easily be handled by working backwards through the resulting lattice, or alternatively we can calculate an expected value, which will of course involve many more paths. We will consider briefly a two period example with zero interest

Suppose mortality rates at age 60 and 61 are as above. Suppose further, that if the value of q_{61} at age 61 is 0.20, then the value of q_{62} applicable at time 2 for an individual now age 60 is either 0.25 or 0.35. If the value of q_{61} at age 62 is 0.30, the corresponding value of q_{62} is either 0.35 or 0.45. See Fig. 2 for details. Prices for pure endowments maturing at time 3 are given as follows.

$$\Lambda_{60}(0,3) = 0.392, \qquad \Lambda_{60}^{(1)}(1,3) = 0.560, \qquad \Lambda_{60}^{(2)}(1,3) = 0.420.$$

Prices for pure endowments maturing at time 2 are the same as they were in the one period example. That is

$$\Lambda_{60}^{(1)}(1,2) = 0.8, \qquad \Lambda_{60}^{(2)}(1,2) = 0.7.$$

Our assumptions on the movement of mortality rates, together with the postulated prices, clearly imply that we are assigning equal probabilities to each of the two possible occurrences at each stage. This in turn will mean an equal probability for each of the four possibilities at age 62. Consider a contract that provides an option to purchase at age 62, for a price of 0.60, an annuity which pays 1 unit at age 63 if then alive.

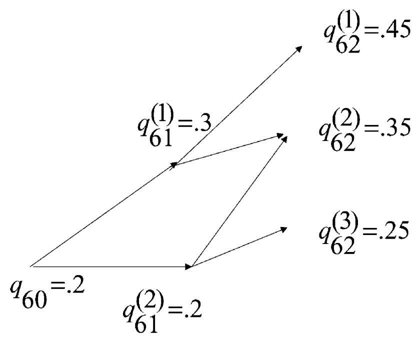


Fig. 2. Two period model for stochastic mortality rates.

From Eqs. (7) and (8), we can calculate our time 1 portfolios. This will be $\mathbf{E}_3 - 0.6\mathbf{E}_2$ if improvement occurs, or $0.5\mathbf{E}_3 - 0.275\mathbf{E}_2$ if not. Our time zero portfolio will then be of the form $\alpha\mathbf{E}_3 - \beta\mathbf{E}_2$ where, in order to be self-financing we need the time 1 value of this portfolio to be equal to the value of the portfolios given above in each of the two mortality realisations. That is

$$\alpha\Lambda^{(1)}(1,3) - \beta\Lambda^{(1)}(1,2) = \Lambda^{(1)}(1,3) - 0.6\Lambda^{(1)}(1,2),$$

and

$$\alpha\Lambda^{(2)}(1,3) - \beta\Lambda^{(2)}(1,2) = 0.5\Lambda^{(2)}(1,3) - 0.275\Lambda^{(2)}(1,2).$$

Solving, $\alpha = 0.75$, $\beta = 0.425$ and the option price is

$$0.75(0.392) - 0.425(0.600) = 0.039.$$

Alternatively, the option price can be computed directly by taking an expectation using the risk neutral probabilities as given. At age 62, the value of the pure endowment will be either 0.750, 0.650, 0.650, or 0.550, each with equal probability. The value of the option is computed as in Eq. (10), except we must now discount back on four paths rather than two. This value will be

$$\tfrac{1}{4}[(0.8)(0.8)0.15 + (0.8)(0.8)0.05 + (0.8)(0.7)0.05 + (0.8)(0.7)0] = 39$$

as we had above.

2.3. Hedging and replication with stochastic interest

We now consider the more realistic case where interest rates are stochastic. We will consider only the one period case here, leaving the details of the multi-period extension to the reader. We will redo the one period example of the previous section, except instead of zero interest, we postulate interest rates by the following prices for 1 unit zero-coupon bonds. The price of a 1 year bond is 0.8. The price of a 2 year bond is 0.56. One year from now the price of the bond maturing at time two is either 0.8 or 0.6. We assume further that at time 0, the price of the contract paying 1 unit at age 62 if the purchaser is alive is 0. 336, and the price of the contract paying 1 unit at age 61 if the purchaser is alive is 0.640. See Fig. 3 for details. These prices are then consistent with risk neutral probabilities of

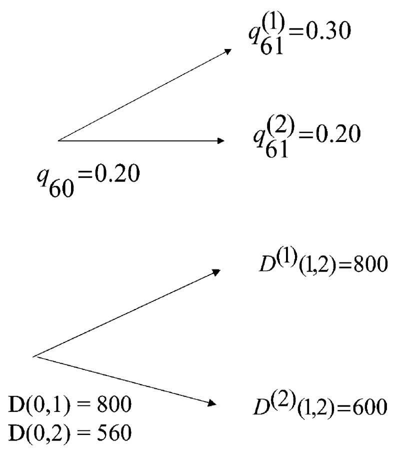


Fig. 3. Stochastic mortality rates and interest rates.

$\frac{1}{2}$ each, for the up or down movement in both interest or mortality. Consider a contract which gives the individual the option to purchase a 1 unit, 1 year pure endowment at time 1, for a strike price of Λ.

In view of the four possibilities for interest and mortality in year 2, we will need four, rather than two assets to replicate. The individual will now take long positions in both $\mathbf{E}_2$ and $\mathbf{E}_1$ and finance this by two distinct types of insured loans.

Type 1: The loan bears interest at the prevailing rates, and the borrower purchases variable insurance to provide for repayment in the event of death within 2 years. Hence, a unit of this asset provides for a payment at time 2 if the individual is alive. The amount of this payment is 1/0.64 if $D(1,2) = 0.8$, or 1/0.48 if $D(1,2) = 0.6$.

Type 2: This is a 1 year loan, which promises to repay at time 1, a bond maturing at time 2. Once again, the borrower purchases insurance to repay the loan in the event of death in the first year. A unit of this asset then provides at time 1, a 1 unit bond maturing at time 2, if the individual is alive.

Our initial replicating portfolio will consist of G units of $\mathbf{E}_1$, J units of $\mathbf{E}_2$, K units of the type 1 loan and L units of the type 2 loan.

One method of solving for G, J, K and L is to compute the prices of each asset at time 1 according to which of the states one is in, set the value equal to 0, and solve the four equations. We will, however, describe an alternate method, which better explains the logic behind the procedure and which is more adaptable to an extension to the multi-period case.

As in the deterministic interest case, the key quantity for the replicating–hedging procedure is the quantity H defined as in Eq. (7). The difference here is that H now is a two-valued random variable, since its value will depend on the interest rate applicable to year 2.

As before, one replicates the option by buying H units of income now, and $1 - H$ at time 1. Even though one does not know the value of H at time 0, this can be accomplished through the mechanism of the loans. At time 0, the individual buys H units of $\mathbf{E}_2$, plus an additional Δ units of $\mathbf{E}_2$, where Δ is sufficient to repay the type 1 loan. Although H and Δ both depend on the interest rate in year 2, we can do this if the *sum* of the two is deterministic. In other words, we require that

$$J = H + \Delta = H - \frac{K}{D(0,1)D(1,2)}. \tag{11}$$

Define $\Lambda' = \min\{\Lambda_{60}^{(2)}(1,2), \Lambda\}$. That is, Λ' is the amount that will be paid at time 1 to acquire 1 unit of a pure endowment at time 2.

In addition to the income at time 2, the individual buys sufficient units of $\mathbf{E}_1$, so that at time 1, the pure endowment payment together with Λ' will purchase the additional $(1 - H)$ units of pure endowment and also repay the type 2 loan. In other words, we require that

$$G + \Lambda' = (1 - H)\Lambda(1,2) - D(1,2)L. \tag{12}$$

For a fuller understanding of what is happening here, we analyse the case where the interest rate in year 2 is the lower of the two possibilities. This will cause the value of the option to go up, requiring a greater amount of hedging, so H will increase. The lower interest, however, will result in a lower payment necessary to discharge the type 1 loan, and Δ will decrease. The opposite is true for the type 2 loan, as the bonds which must be provided at time 1 to repay the loan will increase in value. This will compensate for the decrease in $1 - H$ and we can then solve for G and L as deterministic quantities in Eq. (12).

For a particular example, suppose the strike price is 0.525. Then $H = 1$ when $D(1,2) = 0.8$, or $H = 0$ when $D(1,2) = 0.6$. From Eq. (11)

$$J = 1 - \frac{K}{0.64} = -\frac{K}{0.48},$$

which is solved to give

$$J = 4, \qquad K = -1.920.$$

From Eq. (12)

$$G + 0.525 = -0.8L, \qquad G + 0.480 = 0.480 - 0.6L.$$

(For the second equation above, we need only choose one of the possible mortality realisations. It is easy to see from the definition of H that using the other realisation will simply add a constant to each side of the equation.) Our two equations are solved immediately to give

$$G = 1.575, \qquad L = -2.625.$$

Since the initial price vector of $\mathbf{E}_2, \mathbf{E}_1$, type 1 loan, type 2 loan, respectively, is (0.64, 0.336, 0.6, 0.448). We can then calculate the option value as

$$1.575(0.64) + 4(0.336) - 1.92(0.6) - 2.625(0.448) = 0.024.$$

Of course, if we know the probability measure Q, we do not need to compute the replicating portfolio to calculate the option price, but can simply take an appropriate expectation as we did at the end of the previous section. The prices for the pure endowment at time 1 will be either 0.64, 0.56, 0.48, or 0.42, each equally likely under Q. The 1 year discount factor at time 0 for interest and mortality is $(0.8)(0.8) = 0.64$, giving the option value of

$$\tfrac{1}{4}(0.64)[115 + 35 + 0 + 0] = 0.024$$

as calculated above.

Hedging by the insurer can be done in the same way as we showed with deterministic interest. The insurer must arrange to hold H units of insurance on lives aged 61 for every unit of option sold at age 60. This will ensure that the total expected payments on the annuities and insurance will then be independent of the mortality in year 2. The difference here is that this must be variable insurance, to account for the two possible values of H. For example with the strike price of 0.525, they would have to sell insurance where the death benefit at time 2 was 1, if $D(1, 2) = 0.8$, or 0 if $D(1, 2) = 0.6$. While this may not be directly achievable, it could be manufactured by selling a policy with a death benefit of 4, and then reinsuring with an indexed policy which pays 3 if the bond price is 0.8 or 4 if the bond price is 0.6.

To clarify the procedure here, we will redo the example with a strike price of 0.435. We now have $H = 1$ if $D(1, 2) = 0.8$, and $H = 0.75$ if $D(1, 2) = 0.6$. which gives

$$J = 1 - \frac{K}{0.64} = 0.75 - \frac{K}{0.48},$$

so that

$$J = 1.75, \qquad K = -0.48.$$

Moreover,

$$G + 0.465 = -0.8L = 0.25(0.48) - 0.6L$$

so that

$$G = 0.045, \qquad L = -0.6.$$

The price of the option is then

$$0.045(0.64) + 1.75(0.336) - 0.48(0.6) - 1.8(0.448) = 0.06.$$

As a check, calculating this by the expectation method yields

$$\tfrac{1}{4}(0.64)[205 + 125 + 45 + 0] = 0.06.$$

The deterministic interest case discussed in the previous section can be obtained as a special case of the above. In that case $D(1,2)$ takes only one value. We, therefore, can solve Eq. (11) by taking $K = 0$ and $J = H$. Similarly, we solve Eq. (12) by taking $G = 0$. Choosing the realisation where mortality improves, we have

$$L = \frac{\Lambda - (1-H)\Lambda^1(1,2)}{D(1,2)} = \frac{H\Lambda^2(1,2)}{D(1,2)},$$

which is exactly that which we obtained in formula (8). (We obtain the last equality above by using the definition of H.)

To summarise, each unit of the option is replicated by purchasing H units of endowment income now and $1-H$ later, where the variable nature of H is handled by taking out the appropriate types of loans to finance the purchase. The insurer does not have to worry about the $1-H$ units purchased at time 1, since the appropriate price will be paid for these. They do, however, need to hedge against the H units bought at time 0, for which they have guaranteed the unknown second year mortality. They do this by selling H units of life insurance.

2.4. General formula for the option in discrete time

We can now give a general formula for the option to purchase, at time k, a contract paying 1 unit at time n if the individual is alive at that time, for a strike price of Λ. The option price will just be the expectation under Q of the difference between the market price and the strike price at time of exercise (if positive), discounted with both interest and mortality back to time 0. It is given by

$$C_x(k,n|\Lambda) = E_Q\left[\prod_{i=1}^{k}\Lambda_x(i-1,i)\max[\Lambda_x(k,n)-\Lambda,0]\right].$$

2.5. Annuity options vs. pure endowment options

In this section, we provide examples, as indicated above, to show that we cannot in general consider an option on an annuity as equivalent to a basket of options on pure endowments.

Consider such a basket of options, one for each integer n, $k+1 \le n \le N$ (where N is the last duration at which members of the particular cohort will be living). The nth contract gives the purchaser the option to purchase at time k, a 1 unit pure endowment maturing at time n for a strike price of K_n. Let $K = \sum_{n=k+1}^{N} K_n$. Now the nth option will be exercised if and only if $K_n \le \Lambda(k,n)$ while the annuity option with strike price K will be exercised if and only if $K \le \sum_{n=k+1}^{N}\Lambda(k,n)$. In practice, we would normally be given the strike price K for the annuity. The question is then, can we find a sequence K_n summing to K so that all or none of the individual pure endowments will be exercised. If we can find such a sequence then we could indeed reduce the pricing of annuity options to that of pricing pure endowment options. In the simplified lattice model where there are only finitely many outcomes at each time, it is easy to derive conditions for this to hold. Fix k and consider the option to buy at time k, a 1 unit life annuity beginning at time $k+1$. Consider the possible outcomes, numbered $1, 2, \dots, s$ for interest and mortality at time k. Let $\Lambda^i(k,n)$ be the price at time k of a 1 unit pure endowment payable at time n assuming outcome i. Let

$$a^i(k) = \sum_{n=k+1}^{N}\Lambda^i(k,n), \tag{13}$$

which is the price for a 1 unit life annuity at time k assuming outcome i at time k.

We can partially order outcomes at time k by declaring that

$$i \le j \text{ if } \Lambda^i(k,n) \le \Lambda^j(k,n) \tag{14}$$

for all $n = k+1, \dots, N$.

Theorem. *Annuity options with exercise date k can be priced by pure endowment options if and only if the above ordering is linear*.

Proof. Suppose the ordering is linear. Choose the minimal i such that $a^i(k) \geq K$, and then choose K_n so that $K_n \leq \Lambda^i(k,n)$ and $\sum_{k=1}^{N} K_n = K$. If an outcome j, with $a^j(k) \geq K$ occurs, then $j \geq i$ and the annuity option will be exercised as will all the pure endowment options. If an outcome j with $a^j(k) < K$ occurs then necessarily $j < i$ and none of the pure endowment options will be exercised. □

Conversely, suppose the ordering is not linear. Then we can find indices i, j, m, n such that

$$\Lambda^i(k,n) > \Lambda^j(k,n), \qquad \Lambda^j(k,m) > \Lambda^i(k,m). \tag{15}$$

Suppose that $a^i(k) \leq a^j(k)$, and let $K = a^i(k)$. Consider a basket of pure endowment options with strike prices (K_n) summing to K. For a strike price of K the annuity option will be exercised if outcome i or j occurs, but we claim that not all the pure endowment options will be exercised for both outcomes i and j. Note first that in order for both the options payable at time m and n to be exercised for both outcomes we would need

$$K_m \leq \Lambda^i(k,m), \tag{16}$$

and

$$K_n \leq \Lambda^j(k,n) < \Lambda^i(k,n). \tag{17}$$

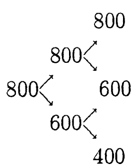
This in turn implies that

$$\sum_{r\neq n,m} K_r > \sum_{r\neq n,m} \Lambda^i(k,r),$$

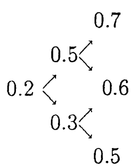
and therefore, that at least one pure endowment option for times other than m, n will not be exercised in outcome i.

2.5.1. Example

Suppose that the bond prices $D(0,1)$, $D(1,2)$ $D(2,3)$ are given by



with equal probabilities of $\frac{1}{2}$ on the up and down movements, and the lattice of q's is given by



with equal probabilities of $\frac{1}{2}$ on the up and down movements.

An option is given to purchase at time 1, a 1000 unit — 2 year annuity with the first payment at time 2 for a strike price of 497. At time 1 there are four possible events depending on whether bond prices are high or low, and whether the q's are high or low. We need only consider two of these, namely outcome 1, high bond prices and high values of q, or outcome 2, low bond prices and low values of q. We have

$$\Lambda^1(1, 2) = 400 < \Lambda^2(1, 2) = 420 \tag{18}$$

but

$$\Lambda^1(1, 3) = 98 > \Lambda^2(1, 3) = 94.50 \tag{19}$$

so the condition is not met.

In the example above, the changes in mortality have a monotone property. If we have at any time, a value of q that is higher than another, then all possible future values arising from rate will be higher than those arising from the other. If interest were deterministic we would achieve the linear ordering, but the problem is that interest rates in this example have the same monotone behaviour. If we modify this by asserting a mean reverting nature to interest rates then it becomes easier to achieve the linear ordering. Suppose for example that the bond prices have the same values as above, but for each of the two outcomes at time 1, the probability that the prices will be 600 at time 2 is $\frac{3}{4}$ rather than $\frac{1}{2}$. Now we get the same values for outcome 1 and also

$$91 = \Lambda^1(1, 3) < \Lambda^2(1, 3) = 103.95 \tag{20}$$

and we do have linear ordering (it is clear that the other two outcomes will give sequences larger than outcome 2 and smaller than outcome 1).

In summary, our option prices on pure endowments can be added to serve as an upper bound for the price of an option to purchase a life annuity but they normally cannot be used to price the option exactly.

3. Continuous-time model

In this section, we develop a continuous-time model of the option-to-annuitise (OTA), by assuming that it is driven by two independent 'short rates'; the instantaneous interest rate and hazard rate. We first derive the No Arbitrage values of the pure-endowment (annuity bond) contracts and then derive option prices on the underlying security. In our example, we assume a CIR process for the short rate, and a generalised mean-reverting Gompertz (MRG) process for the hazard rate.

3.1. General framework, notation and terminology

Following the approach taken by Duffie and Singleton (1997) for the pricing of defaultable bonds, we model a hazard-plus-interest rate process, and denote it by

$$\xi_t = r_t + h_t, \tag{21}$$

where r_t is the instantaneous risk-free rate of interest, and h_t is the hazard rate. In contrast to the corporate bond pricing literature, we quite naturally assume that r_t is independent of h_t. We also ignore recovery rates, since in the personal insurance context, default is complete.

3.1.1. Default-free bonds

In the absence of mortality (default) risk, the time $-t$ price of a zero-coupon bond, which matures at $T > t$, is denoted by, and equal to

$$D_t(T) := E_Q\left[\mathrm{e}^{-\int_t^T r_u\,\mathrm{d}u}|\mathcal{F}_t\right]. \tag{22}$$

Our notation differs slightly from the previous section, simply to emphasise the continuous-time nature of the computation. One again, the risk neutral expectation $E_Q[\cdot]$, should be contrasted with the real world expectation

$E_P[\cdot]$. See Musiela and Rutkowski (1997, Chapter 12) for details on the relationship between the two possible expectations. Clearly, in the trivial (deterministic, constant) case that $r_t = r$, Eq. (22) collapses to $D_t(T) = \exp\{-r(T-t)\}$. On the other hand, when r_t can be represented by a continuous-time stochastic process, Eq. (22) reduces to evaluating the Laplace transform of the random variable $\int_t^T r_t\, dt$. Quite commonly, the expression $D_t(T)$ is also written as

$$D_t(T) = e^{-y_t(T)(T-t)}, \tag{23}$$

where $y_t(T)$ denotes the time $-t$ yield to maturity of a zero-coupon bond. The family $\{y_t(T)\, \forall T \geq t\}$ is referred to as the yield curve, or term structure of interest rates. See Musiela and Rutkowski (1997, Chapter 12) or Hull (2000, Chapter 17) for an in-depth discussion. With some slight abuse of notation, we use $D_t(s,T)$ to denote the time $-t$ forward price of a zero-coupon bond maturing at time $T \geq s$. (A commitment is made at time t to purchase a zero-coupon bond at time $s \geq t$.) Applying the fundamental pricing equation, or a simple cost of carry argument, we have that:

$$D_t(s,T) = E_Q\left[e^{-\int_s^T r_u\, du} | \mathcal{F}_t\right] = E_Q\left[e^{\int_t^s r_u\, du}\, e^{-\int_t^T r_u\, du} | \mathcal{F}_t\right] = \frac{D_t(T)}{D_t(s)} \quad \forall T \geq s \geq t. \tag{24}$$

Finally, the time $-t$ instantaneous forward rate for time s, is defined equal to

$$f_t(s) = \lim_{T \downarrow s} \frac{\ln[D_t(s)] - \ln[D_t(T)]}{T-s} = \lim_{T \downarrow s} \frac{y_t(T)(T-t) - y_t(s)(s-t)}{T-s}. \tag{25}$$

This leads to the well-known forward-based pricing equation for a zero-coupon bond price:

$$D_t(T) = e^{-\int_t^T f_t(u)\, du}, \tag{26}$$

which we will generalise to the annuity context.

3.1.2. Mortality functions

Using the same framework, the probability of survival to time T — conditional on being alive at time $-t$ — is denoted and equal to

$$p_t(T) := E_Q\left[e^{-\int_t^T h_u\, du} | \mathcal{F}_t\right]. \tag{27}$$

Eq. (27) implicitly assumes we are focusing on one cohort group, all born in a particular year. In the event we want to distinguish between different cohorts, we will attach the birth year y, to the probability $p_t(T|y)$. Now, to classically trained actuaries, Eq. (27) may seem like an odd way to represent survival functions. But, in fact, this generalisation is consistent with, and has embedded within it, traditional practice. For example, in the trivial case that the hazard (failure) rate is constant, $h_t = h$, the probability of survival is given by: $p_t(T) = \exp\{-h(T-t)\}$. This constant force of mortality assumption is synonymous with a future lifetime that is exponentially distributed. In the more general case that the hazard rate itself is stochastic — which is the essence of our paper — the probability of survival can be represented via the Laplace transform of the integral of the hazard rate. Actuaries might be more familiar with the instantaneous force of mortality μ_s as the underlying pricing factor. Using that approach, $p_t(T) = \exp\{-\int_t^T \mu_u\, du\}$. However, as we argued in the introduction to the paper, we prefer to define the force of mortality, analogous to a *forward interest rate*, where

$$\mu_t(s) = \lim_{T \downarrow s} \frac{\ln[p_t(s)] - \ln[p_t(T)]}{T-s} \quad \forall T \geq s \geq t. \tag{28}$$

In this framework, the instantaneous force of mortality is defined directly from the probability of survival, which in turn is driven by the stochastic hazard rate. In fact, a simple application of Jensen's inequality reveals that:

$$p_t(T) = \mathrm{e}^{-\int_t^T \mu_u\,\mathrm{d}u} = E\left[\mathrm{e}^{-\int_t^T h_u\,\mathrm{d}u}|\mathcal{F}_t\right] \geq \mathrm{e}^{-\int_t^T E[h_u]\,\mathrm{d}u}. \tag{29}$$

Practically speaking, we are arguing that the instantaneous force-of-mortality curve (in insurance) is equivalent to the instantaneous forward-rate curve (in finance). All claims can be priced 'off' these curves.

The idea of a stochastic hazard rate does not seem to have previously appeared in the actuarial literature. However, related concepts have been employed by population biologists in studying human aging. See Woodbury and Manton (1977) or Yashin et al. (1985).

3.1.3. *Mortality contingent-claim curve*

As in the discrete time model, we start with pure endowment contracts, which can then be joined (summed) to form a life annuity. This is analogous to creating coupon bearing bonds using a mixture of zero-coupon contracts. In continuous time, the price of a pure endowment contract, with maturity T, is denoted and equal to

$$\begin{aligned}\Lambda_t(T) &= E_Q\left[\mathrm{e}^{-\int_t^T h_u\,\mathrm{d}u}|\mathcal{F}_t\right] E_Q\left[\mathrm{e}^{-\int_t^T r_u\,\mathrm{d}u}|\mathcal{F}_t\right] = E_Q\left[\mathrm{e}^{-\int_t^T (h_u+r_u)\,\mathrm{d}u}|\mathcal{F}_t\right] \\ &= E_Q\left[\mathrm{e}^{-\int_t^T \xi_u\,\mathrm{d}u}|\mathcal{F}_t\right].\end{aligned} \tag{30}$$

Intuitively, the pure endowment price is equal to the product of the default-free bond price and the probability of survival. Naturally, $\Lambda_t(T) \to 0$, when $T \to \infty$. Also, as in the case of survival probability, the notation $\Lambda_t(T|y)$ will be used when a particular cohort is being identified. Analogous with the (default-free) zero-coupon bond, we define $\Lambda_t(s, T)$ to represent the time $-t$ forward price of a pure endowment maturing at time $T \geq s$. In the context of a life annuity, a commitment is made at time t to purchase a pure endowment at time $s \geq t$. We emphasise that $\Lambda_t(s, T)$ is not exactly a deferred annuity (in the classical actuarial sense), since the acquisition payment is not made now, but rather at time $s > t$, contingent on survival. In any event, applying the fundamental pricing equation, we have that:

$$\Lambda_t(s, T) = E_Q\left[\mathrm{e}^{-\int_s^T \xi_u\,\mathrm{d}u}|\mathcal{F}_t\right] = \frac{\Lambda_t(T)}{\Lambda_t(s)} \quad \forall T \geq s \geq t. \tag{31}$$

Following the same line of reasoning, the time $-t$ price of a call option to acquire a pure endowment with maturity T, at time $s \leq T$, for a fixed price of Λ, is denoted by $C_t(s, T, \Lambda)$. The payoff at maturity is

$$C_s(s, T, \Lambda) = \max[\Lambda_s(T) - \Lambda, 0]. \tag{32}$$

The complete-market (or No Arbitrage) price of this call option is

$$\begin{aligned}C_t(s, T, \Lambda) &= E_Q\left[\mathrm{e}^{-\int_t^s \xi_u\,\mathrm{d}u} \max[\Lambda_s(T) - \Lambda, 0]|\mathcal{F}_t\right] \\ &= E_Q\left[\mathrm{e}^{-\int_t^s \xi_u\,\mathrm{d}u} \max\left[E_Q\left[\mathrm{e}^{-\int_s^T \xi_u\,\mathrm{d}u}|\mathcal{F}_s\right] - \Lambda, 0\right]|\mathcal{F}_t\right].\end{aligned} \tag{33}$$

As one would expect intuitively, as $\Lambda \to 0$, $C_t(s, T, \Lambda) \to \Lambda_t(T)$, We now proceed to derive explicit expressions for $\Lambda_t(T)$ and $C_t(s, T, \Lambda)$, assuming a particular parameterisation of the process $\xi_t = r_t + h_t$.

3.2. *The CIR process for interest*

For completeness, we present the CIR model which is one of the most popular short-term models for interest rates. More specifically:

$$\mathrm{d}r_t = \kappa(\theta - r_t)\,\mathrm{d}t + \sigma_{\mathrm{r}}\sqrt{r_t}\,\mathrm{d}B_t^r. \tag{34}$$

The three interest rate parameters, κ, θ and σ_{r}, are risk-neutralised. Eq. (34) implies that $r_s|r_t$ is non-central chi-square distributed. This fact can be traced to original results by Feller, but is widely attributed to Cox et al. (1985). The expect value and variance of $r_s|r_t$ is

$$E[r_s|r_t] = r_t\,\mathrm{e}^{-\kappa(s-t)} + \theta(1-\mathrm{e}^{-\kappa(s-t)}), \tag{35}$$

$$\mathrm{var}[r_s|r_t] = r_t\left(\frac{\sigma_{\mathrm{r}}^2}{\kappa}\right)(\mathrm{e}^{-\kappa(s-t)} - \mathrm{e}^{-2\kappa(s-t)}) + \theta\left(\frac{\sigma_{\mathrm{r}}^2}{2\kappa}\right)(1-\mathrm{e}^{-\kappa(s-t)})^2. \tag{36}$$

The properties of the CIR short rate process are well known in the term structure literature, so we avoid deriving them here. Taking the Laplace transform in Eq. (22), the bond price is equal to

$$D_t(T) = C_1(t,T)\,\mathrm{e}^{-r_t C_2(t,T)}, \tag{37}$$

where

$$C_1(t,T) = \left[\frac{2\gamma\,\mathrm{e}^{(\kappa+\gamma)(T-t)/2}}{(\gamma+\kappa)(\mathrm{e}^{\gamma(T-t)}-1)+2\gamma}\right]^{2\kappa\theta/\sigma_{\mathrm{r}}^2}, \tag{38}$$

$$C_2(t,T) = \frac{2(\mathrm{e}^{\gamma(T-t)}-1)}{(\gamma+\kappa)(\mathrm{e}^{\gamma(T-t)}-1)+2\gamma}, \tag{39}$$

and $\gamma = \sqrt{\kappa^2 + 2\sigma_{\mathrm{r}}^2}$.

3.3. Mean reverting Brownian Gompertz

We are looking for a continuous-time diffusion process for the hazard rate, that can serve as the stochastic analogue of the Gompertz survival function. See Carriere (1994), Gutterman and Vanderhoof (1998) and Tennebein and Vanderhooof (1980) for references to the Gompertz and other continuous mortality models. Naturally, there are a variety of models that one can select as extension to Gompertz. After experimenting with a variety of approaches, the authors have selected a so-called mean reverting Brownian Gompertz (MRBG) specification. This process is expected to grow exponentially, the variance is proportional to the value of the hazard rate, the process will never hit zero, and exhibits mean reversion. All of these are desirable and perhaps even required conditions for a proper mortality hazard rate function. Technically, we have that:

$$h_t = h_0\,\mathrm{e}^{gt+\sigma Y_t}, \quad g, \sigma, h_0 > 0, \tag{40}$$

$$\mathrm{d}Y_t = -bY_t\,\mathrm{d}t + \mathrm{d}B_t^h, \quad Y_0 = 0, b \geq 0. \tag{41}$$

When $b = 0$, the process collapses to geometric Brownian motion, since Y_t becomes B_t. As b increases, Y_t displays stronger mean reversion. The stochastic differential equation for Y_t can be solved explicitly to yield:

$$Y_t = \int_0^t \mathrm{e}^{-b(t-u)}\,\mathrm{d}B_u^h \tag{42}$$

with a mean value (stochastic integral) of zero, and a variance (by Ito's isometry) of:

$$\sigma_t^2 = E[Y_t^2] = \int_0^t (\mathrm{e}^{-b(t-u)})^2\,\mathrm{d}B_u^h = \frac{1-\mathrm{e}^{-2bt}}{2b}, \tag{43}$$

which goes to t as $b \to 0$, and is always smaller than t, for $b > 0$. In other words, the process has a smaller variance compared to a standard Brownian motion. The higher moments of Y_t are given by

$$E[Y_t^{2n-1}] = 0, \quad n = 1, 2, \ldots, \tag{44}$$

$$E[Y_t^{2n}] = E\left[\left(\frac{Y_t}{\sigma_t}\right)^{2n}\right]\sigma_t^{2n}, \quad n = 1, 2, \dots, \tag{45}$$

$$E[Y_t^{2n}] = \frac{(2n)!}{2^n(n!)}\left[\frac{1-\mathrm{e}^{-2bt}}{2b}\right]^n, \quad n = 1, 2, \dots, \tag{46}$$

$$E[Y_t^{2n}] = \frac{(2n)!}{(n!)}\left[\frac{1-\mathrm{e}^{-2bt}}{4b}\right]^n, \quad n = 1, 2, \dots. \tag{47}$$

The log-hazard rate $\ln[h_t]$ is, therefore, *normally* distributed with a mean value of $\ln[gt] + \ln[h_0]$ and a variance of $\sigma^2(1-\exp(-2bt))/2b$. The (non-central) moments of h_t are given by

$$E[h_t^n] = E[(h_0\,\mathrm{e}^{gt+\sigma Y_t})^n], \tag{48}$$

$$E[h_t^n] = h_0^n\,\mathrm{e}^{ngt}E[\mathrm{e}^{n\sigma Y_t}] \tag{49}$$

so that

$$E[h_t^n] = h_0^n\,\mathrm{e}^{ngt}\exp\left[\frac{(n\sigma)^2}{2}\left(\frac{1-\mathrm{e}^{-2bt}}{2b}\right)\right], \quad n = 1, 2, \dots. \tag{50}$$

For example, when $n = 1$, we have the expected hazard rate of:

$$E[h_t] = h_0\exp\left(gt + \frac{\sigma^2}{2}\left(\frac{1-\mathrm{e}^{-2bt}}{2b}\right)\right). \tag{51}$$

The expected hazard rate will be the standard Gompertz function when $b = 0$ and therefore $E[h_t] = h_0\exp((g + (\sigma^2/2))t)$. To gain further insight into the structure of h_t, we let $F(t, y) = h_0\exp(gt + \sigma y)$, then by Ito's lemma, we have

$$\mathrm{d}F(t, y) = \frac{\partial F(t, y)}{\partial t}\,\mathrm{d}t + \frac{\partial F(t, y)}{\partial y}\,\mathrm{d}y + \frac{1}{2}\frac{\partial^2 F(t, y)}{\partial y^2}\,\mathrm{d}\langle Y\rangle_t. \tag{52}$$

This, in turn, implies that:

$$\mathrm{d}h_t = gh_t\,\mathrm{d}t + \sigma h_t\,\mathrm{d}Y_t + \tfrac{1}{2}\sigma^2 h_t\,\mathrm{d}t, \tag{53}$$

$$\mathrm{d}h_t = gh_t\,\mathrm{d}t + \sigma h_t(-bY_t\,\mathrm{d}t + \mathrm{d}B_t^h) + \tfrac{1}{2}\sigma^2 h_t\,\mathrm{d}t. \tag{54}$$

Finally, since $Y_t = (\ln[h_t/h_0] - gt)/\sigma$, as per the definition in Eq. (40), we arrive at

$$\mathrm{d}h_t = gh_t\,\mathrm{d}t + \sigma h_t\left(\frac{-b(\ln[h_t/h_0] - gt)}{\sigma}\,\mathrm{d}t + \mathrm{d}B_t^h\right) + \frac{\sigma^2}{2}h_t\,\mathrm{d}t, \tag{55}$$

$$\mathrm{d}h_t = ((g + \tfrac{1}{2}\sigma^2) - b(\ln[h_t] - \ln[h_0] - gt))h_t\,\mathrm{d}t + \sigma h_t\,\mathrm{d}B_t^h, \tag{56}$$

$$\mathrm{d}h_t = (g + \tfrac{1}{2}\sigma^2 + b\ln[h_0] + bgt + b\ln[h_t])h_t\,\mathrm{d}t + \sigma h_t\,\mathrm{d}B_t^h, \tag{57}$$

Table 2
Value of the option to annuitise[a]

Interest assumptions				
σ	15%			
κ	0.25			
ΔT	$\frac{1}{12}$			
r_0	5.50%			
r	7.50%			
Hazard assumptions				
σ	{5, 10, 15, 20}			
b	0.5			
ΔT	$\frac{1}{12}$			
h_0	{0.5, 1, 1.5}			
g	0.1			
	$\sigma = 5\%$	$\sigma = 10\%$	$\sigma = 15\%$	$\sigma = 20\%$
$h_0 = 0.5\%$	7.1% (0.006)	6.2% (0.005)	5.3% (0.004)	3.2% (0.003)
$h_0 = 1\%$	5.2% (0.004)	3.3% (0.004)	2.2% (0.003)	1.9% (0.002)
$h_0 = 1.5\%$	2.1% (0.0002)	1.7% (0.0002)	1.5% (0.002)	1.5% (0.0002)

[a] The enclosed numbers represent the outcome of 10,000 simulation runs for the underlying interest rate and hazard rate sample path. The option is 'struck' at the current forward rate, on a 20 year endowment, with a maturity of 10 years. Standard errors in brackets. The option price is decreasing in volatility b/c the option is struck at the forward price which will increase with volatility, de facto pushing the call option out-of-the-money.

which, interestingly enough, is quite similar to the Black–Derman–Toy model for the short-rate,

$$\mathrm{d}X_t = (A + Bt + \ln[X_t])X_t\,\mathrm{d}t + CX_t\,\mathrm{d}B_t, \tag{58}$$

see Black and Karasinski (1991), or Hull and White (1990) for more details.

3.4. Numerical examples

Although one can obtain closed-form expressions for the CIR bond and option prices, the same is not true for the MRBG process. This is mainly because the (integral) sum of lognormal variates is not lognormal. Now, although approximations are available in the literature, e.g. Hansen and Jorgensen (1999) or Milevsky and Posner (1998), we have decided to conduct a Monte Carlo simulation to obtain some bond and option values. To that end, we have discretised both diffusions, and simulated monthly interest and hazard rate changes to obtain the term structure of mortality-contingent claims. For example, when the interest rate parameters are $\sigma_\mathrm{r} = 0.15, \theta = 0.075, \kappa = 0.25, r_0 = 0.055$ and the hazard rate parameters are $\sigma_\mathrm{h} = 0.20, g = 0.10, b = 0.5, h_0 = 0.015$ we obtain the following values for the pure endowment contract and the options. With $n = 10,000$ simulations, $\Lambda_0(10) = 0.40867$, with a standard error of (0.0011), $\Lambda_0(20) = 0.10442$, with a standard error of (0.0005), and an option value of $C_0(10, 20, 0.25) = 0.01518$ with a standard error of (0.0002). These parameters correspond with a 70-year-old that is granted an option to annuitise at the current forward rates, in 10 years time. The option would be worth 1.5% of the face value of endowment policy. In general, it appears that both the option values and the bond prices are quite insensitive to the magnitude of the mean reversion parameter b, but highly sensitive (naturally) to the volatility $\sigma_\mathrm{r}, \sigma_\mathrm{h}$. Also, the higher the current hazard rate h_0, or the growth rate g, the lower the value of the option to annuitise. Table 2 provides some additional values for the option and bond prices under various parameter configurations.

4. Conclusion

In this paper, we have proposed a model for pricing options on future mortality (and interest) rates. These option currently exist in the market, but very little has been written on how they should be priced or reserved against. These options 'pay off' if life annuity (pure endowment) prices end up at a higher than some pre-specified strike price, at the time the contract was issued. We have presented both a discrete and continuous time model, and have demonstrated how to hedge these options using pure endowments, default free bonds and life insurance contracts. In the discrete model, the hedging details are shown explicitly. For the continuous model we rely on the theory of Duffie and Singleton (1996, 1997). Our main conceptual contribution is in treating the actuarial hazard rate itself as a stochastic variable, as opposed to a deterministic force of mortality. Indeed, in classical deterministic actuarial models, the option to annuitise — struck at the forward — should have zero value since actuaries do not allow for uncertainty in the mortality factors that will be applicable upon annuitisation. In contrast, we treat the probability of survival as an expectation over the path of the hazard rate over time. Pure endowment prices can then be viewed as generating a mortality term structure. We interpret the classical force of mortality as a forward curve, off which all claims are priced in a No Arbitrage framework. This mortality forward curve aggregates market expectations about futures mortality rates, but at the same time can change over time in response to new information. This approach is similar in spirit to recent work by Mullin and Philipson (1997), where they use market prices for term life insurance policies to imply expectations of future mortality rates. However, we take this process one step further by assuming the market has a cohort specific expectation that actually changes over time. The stochasticity in the future survival probabilities for any particular cohort is what gives value to the option. Future research will attempt to estimate the parameters of the hazard rate process by looking at annuity price data.

Acknowledgements

This research is partially supported by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC) and Mathematics of Information Technology and Complex System (MITACS) program. The authors would like to thank Scott Warlow for research assistance, and to thank a very helpful anonymous reviewer, participants at the IME2000 (Barcelona) conference, and Lowell Aronof, Glenn Daily, Mary Hardy, Harry Klaristenfeld, Mike Orszag, Mark Warshawsky for comments and discussions throughout the development of this research.

References

Black, F., Karasinski, P., 1991. Bond and option pricing when short rates are lognormal. Financial Analysts Journal 47, 52–59.

Carriere, J., 1994. An Investigation of the Gompertz Law of Mortality, Vol. 2. Actuarial Research Clearing House.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

Duffie, D., Singleton, K., 1996. Modeling the term structure of defaultable bonds. Working Paper. Graduate School of Business Administration, Stanford University.

Duffie, D., Singleton, K., 1997. An econometric model of the term structure of interest-rate swap yields. The Journal of Finance 52 (4).

Gutterman, S., Vanderhoof, I.T., 1998. Forecasting changes in mortality: a search for a law of causes and effects. The North American Actuarial Journal 2 (4), 135–138.

Hansen, A.T., Jorgensen, P.L., 1999. Fast and accurate analytical approximation of bond prices when short interest rates are lognormal. Journal of Computational Finance 3 (2), 27–45.

Hull, J., 2000. Options, Futures and Other Derivatives, 4th Edition. Prentice-Hall, Englewood Cliffs, NJ.

Hull, J., White, A., 1990. Pricing interest-rate derivative securities. The Review of Financial Studies 3 (4), 573–592.

Milevsky, M.A., Posner, S.E., 1998. Asian options, the sum of lognormals, and the reciprocal gamma distribution. Journal of Financial and Quantitative Analysis 33 (2), 409–422.

Milevsky, M.A., Posner, S.E., 2001. The titanic option: valuation of the guaranteed minimum death benefits in variable annuities and mutual funds. The Journal of Risk and Insurance 68 (1), 93–128.

Mullin, C., Philipson, T., 1997. The future of old-age longevity: competitive pricing of mortality contingent claims. Working Paper No. 6042. National Bureau of Economic Research.

Musiela, M., Rutkowski, M., 1997. Martingale Methods in Financial Modelling. Springer, Berlin.

Tennebein, A., Vanderhooof, I.T., 1980. New mathematical laws of select and ultimate mortality. Transactions of the Society of Actuaries 32, 119–158.

Woodbury, M.A., Manton, K.G., 1977. A random-walk model of human mortality and aging. Theoretical Population Biology 11, 37–48.

Yashin, A.I., Manton, K.G., Vaupel, J.W., 1985. Mortality and aging in a heterogeneous population: a stochastic process model with observed and unobserved variables. Theoretical Population Biology 27, 154–175.