Pricing Frameworks for Securitization of Mortality Risk[1]

Andrew J.G. Cairns Heriot-Watt University
David Blake Cass Business School
Kevin Dowd Nottingham University Business School

First version: June, 2004
This version: June 22, 2004

Abstract

It is now an accepted fact that stochastic mortality – the risk that future trends in mortality are different from those anticipated – is an important risk factor in both life insurance and pensions. As a risk factor it affects how we calculate fair values, premium rates, and risk reserves.

In this paper we discuss theoretical frameworks and models for pricing mortality derivatives and valuing liabilities which incorporate mortality guarantees. Models developed within one of these frameworks also facilitate the calculation of risk (or quantile) reserves and in a way that is consistent with an arbitrage-free pricing framework. The objective of the paper is to provide a foundation for further work which will look at the practical development and implementation of such models.

The different frameworks that we describe are all based on positive-interest-rate modelling frameworks since the force of mortality can be treated in a similar way to the short-term, risk-free rate of interest. The frameworks discussed are short-rate models, forward-mortality models, positive-mortality models and mortality market models.

Keywords: stochastic mortality, term structure of mortality, survivor index, spot survival probabilities, spot force of mortality, forward mortality surface, short-rate models, forward mortality models, positive mortality framework, mortality market models, annuity market model, SCOR market model.

[1]Corresponding author: Andrew Cairns E-mail: A.Cairns@ma.hw.ac.uk Tel: +44 131 451 3245, Fax: +44 131 451 3249. Full addresses are given at the end of the paper.

1 Introduction

A large number of products in life insurance and pensions by their very nature have mortality as a primary source of risk. By this we mean that products are exposed to *unanticipated* changes over time in the mortality rates of the appropriate reference populations. For example, annuity providers are exposed to the risk that the mortality rates of pensioners will fall at a faster rate than accounted for in their pricing and reserving calculations. A more specific example is that of the Equitable Life Assurance Society. The embedded options in a large number of their annuity contracts became very valuable in the 1990's due to a combination of falling interest rates and improvements in mortality. It is possible that the eventual downfall of this institution might have been avoided if they had been able to hedge their exposures to both interest-rate risk and mortality improvement risk. The theory and practice of interest-rate modelling within an arbitrage-free market is already well developed (see, for example, Cairns, 2004b, James and Webber, 2002, Rebonato, 2002, Brigo and Mercurio, 2001, Brace, Gatarek and Musiela, 1997, Jamshidian, 1997, Heath, Jarrow and Morton, 1992, Cox, Ingersoll and Ross, 1985, and Vasicek, 1977). Previous literature in insurance mathematics has tended to focus on interest-rate risks only, in combination with an assumption that mortality rates evolve in a deterministic way over time. This paper combines both stochastic interest and stochastic mortality and we develop a range of arbitrage-free frameworks for pricing and hedging mortality risk.

Early actuarial work treated mortality rates at different ages as being constant over time. Such an approach worked adequately in an environment dominated by with-profits contracts in combination with prudent mortality tables. As contract types evolved it was recognised that in certain cases calculations would require an accurate, rather than a prudent, assessment of future mortality rates. This meant that it would be necessary to build projected future improvements in mortality into pricing and reserving calculations. An early table which took account of mortality improvements was the $a(55)$ table used in the UK. This was based on graduated data for annuitants in 1947-48 with a projection to 1955. The next logical step was the use of mortality forecasts and double-entry tables. This was considered by the Joint Mortality Investigation Committee (JIMC, 1974, page 200) but they argued that such a table *"might lead those who will use the resulting tables to attribute to them an authority which they do not possess"*. Nevertheless, later tables in the UK produced by the Continuous Mortality Investigation Bureau (CMI) moved on to use deterministic double-entry tables (see, for example, the discussions in CMI, 1978, 1999).

Ultimately, the concern expressed in JMIC (1974) was well founded given that mortality improvements in the 1980's and 1990's turned out to be greater than forecast. This uncertainty was confirmed when Currie, Durban and Eilers (2004) (hereafter CDE) analysed historical trends in mortality using P-splines. The fitted surface of values for the force of mortality[2] $\hat{\mu}(t, x)$ is plotted on a log scale in Figure 1.1 while the development of the force of mortality for specific ages over time relative to values in 1947 is plotted in Figure 1.2.[3] Figure 1.2 reveals some detail that we cannot see in Figure 1.1: specifically that the rate of improvement has varied substantially over time, and that the improvements have varied substantially between different age groups. CDE constructed confidence bounds for the future development of mortality rates. Inevitably these confidence bounds get wider as the forecast horizon lengthens and CDE found that even 15-20 years ahead the bounds are very wide. Thus although the CMI's projections turned out to be inaccurate they were within the confidence bounds suggested by CDE. In general terms the analysis of CDE, as well as other analyses using stochastic mortality models discussed below, indicates that future mortality improvements cannot be forecast with any degree of precision. Short-term trends might be detected by keeping a close watch on recent medical advances, but even then the precise effect of such advances is difficult to judge. As we look further into the future it becomes even more difficult to predict what medical advances there might be, when they will happen, and what impact they will have on survival rates.

More recently it has also become apparent that deterministic mortality projections are inadequate for some applications, even where these are taken to be best estimates rather than prudent estimates. Cases where stochastic approaches are sometimes felt to be necessary include:

- Calculation of quantile (or Value-at-Risk) reserves. For example, a 99% quantile reserve (if correctly calculated) should be sufficient 99% of the time to meet all future contracted payments on a portfolio of liabilities. The *uncertain* future pattern of liability payments will depend, amongst other things, on the future evolution of the force of mortality $\mu(t, x)$.[4] Indeed, the UK financial services regulator is currently discussing the use of stochastic mortality in reserving calculations.

- Pricing and reserving for policies which incorporate certain types of guarantee. For example, a *guaranteed annuity option* is an investment-linked deferred-annuity contract which gives a policyholder the option to convert his accumulated fund at retirement at a guaranteed rate rather than at current market rates. The value of this option most obviously depends upon the level of interest rates at retirement but it also depends upon the mortality table being used by the life office at the time of retirement.

[2]The force of mortality $\mu(t, x)$ is described in more detail at the start of Section 2.

[3]Note that the limited amount of data at both low and high ages means that fitted values at these ages have relatively large confidence bounds.

[4]Here, t represents the current time, x the age at time t of a specified life and the probability that the individual will die between times t and $t + dt$ given he has survived until the current time is $\mu(t, x)dt + o(dt)$ as $dt \to 0$ (that is, approximately $\mu(t, x)dt$ for small dt).

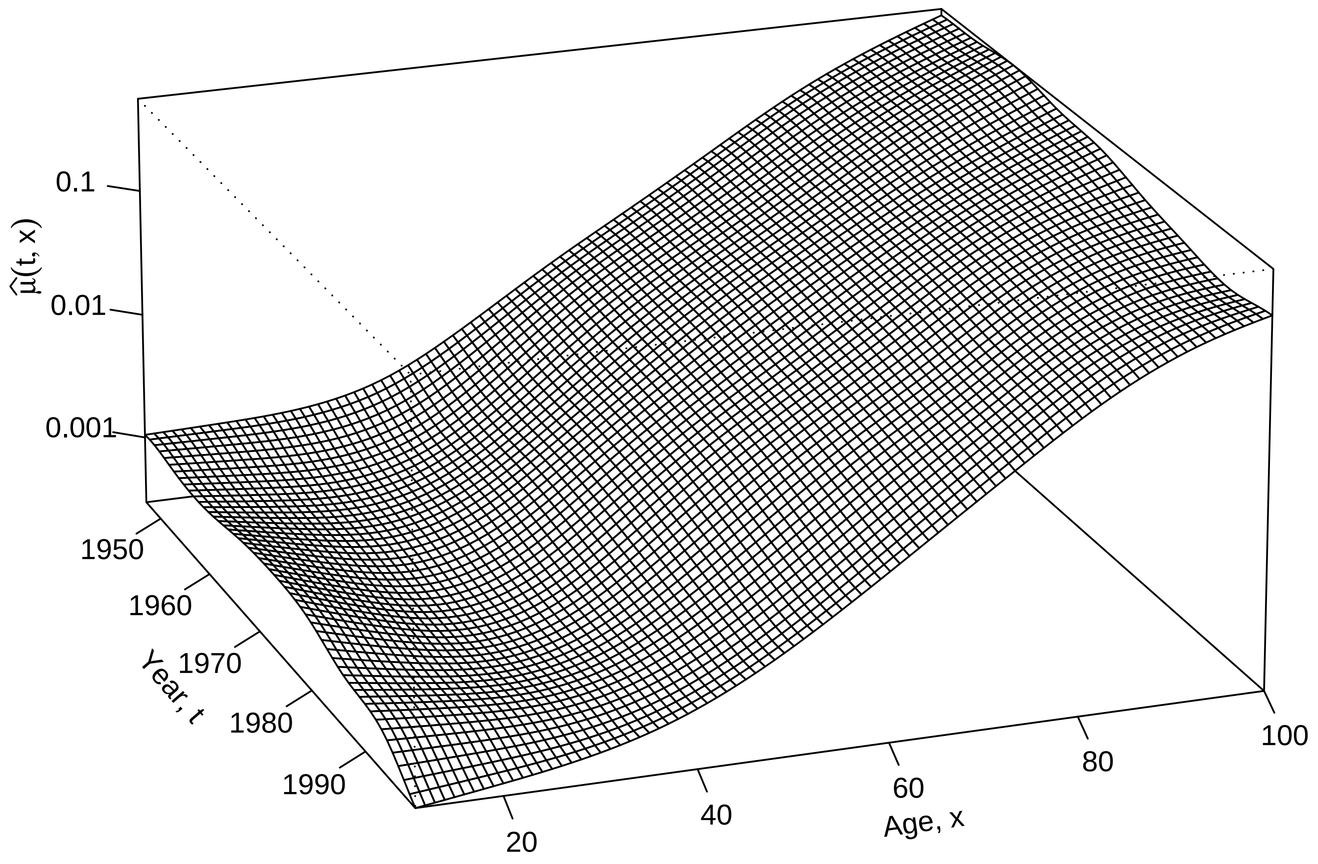

Figure 1.1: Fitted values using P-splines for the force of mortality $\hat{\mu}(t, x)$ for the years $t = 1947$ to 1999 and for ages $x = 11$ to 100 from Currie, Durban and Eilers, 2004.

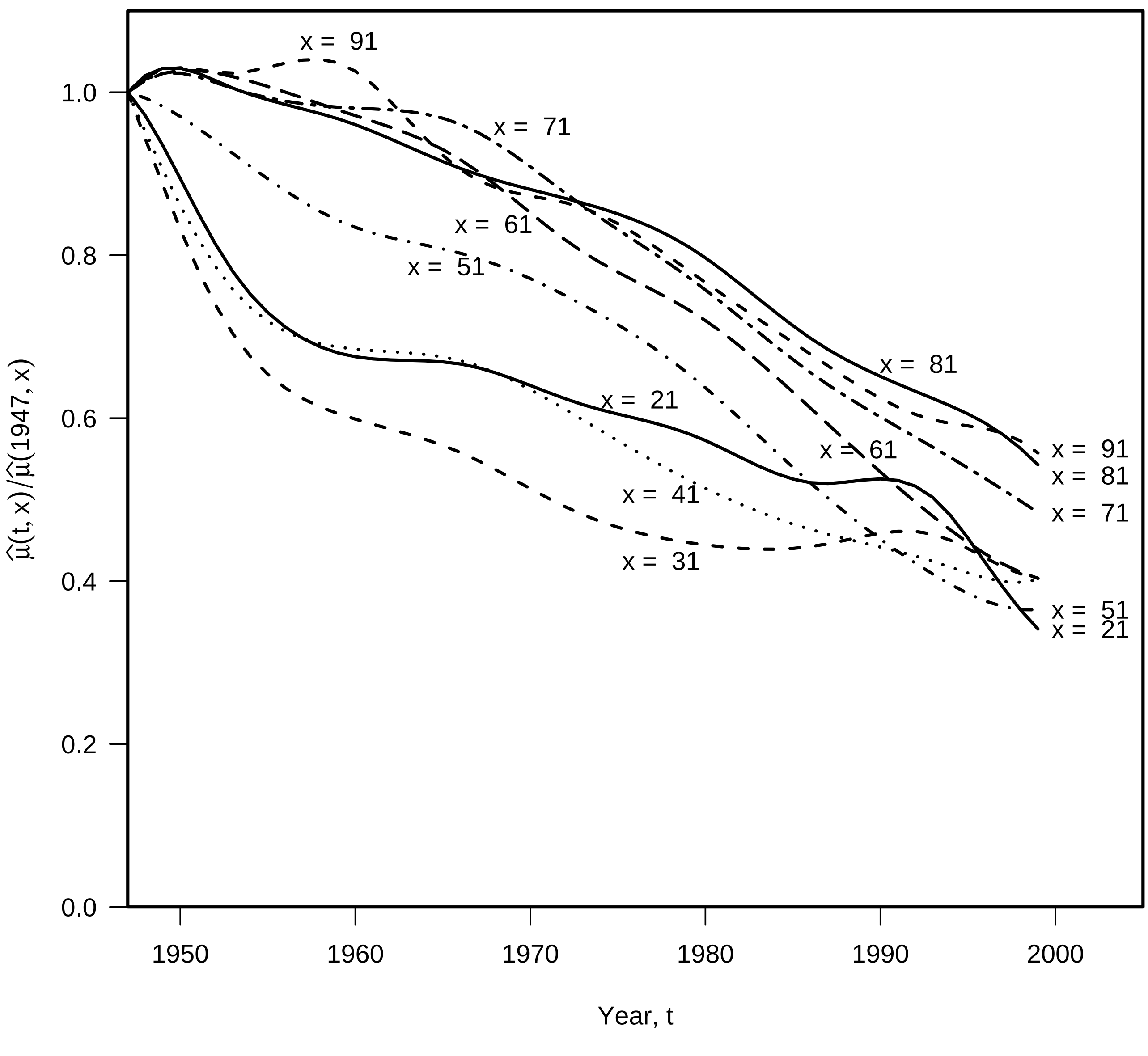

Figure 1.2: $\hat{\mu}(t,x)/\hat{\mu}(1947,x)$: Fitted values using P-splines for the force of mortality $\hat{\mu}(t,x)$, relative to the 1947 value for the years $t = 1947$ to 1999 and for ages $x =$ 21, 31, 41, 51, 61, 71 and 81 from Currie, Durban and Eilers, 2004. It can be seen that the pattern of improvements is different at different ages.

- Pricing of mortality derivatives. Examples of such contracts include:
  - survivor bonds (where coupon payments are linked to the number of survivors in a given cohort).[5] The concept of survivor bonds which deal with longevity risk[6] has recently been revived by Blake and Burrows (2001) and Lin and Cox (2004). Their origin dates back to Tontine bonds issued by a number of European governments in the 17^{th} and 18^{th} centuries.
  - survivor swaps (where counterparties swap a fixed series of payments in return for a series of payments linked to the number of survivors in a given cohort).[7] The case for survivor swaps is made by Dowd *et al* (2004).
  - annuity futures (where prices are linked to a specified future market annuity rate).[8]
  - mortality options (a range of contracts with option characteristics whose payoff depends on an underlying mortality table at the payment date). The guaranteed annuity contract mentioned above is an example a mortality option, although it is really a complex option involving interest rate risks as well. Contracts of this type are discussed further in Section 6.2.

For each of these derivatives, the reference population underlying the calculation of the mortality rates is central to both the viability and liquidity of the contracts. Specifically some investors will wish to use a contract to help hedge their mortality

[5]The first, widely-marketed, pure survivor bond was issued by Swiss Re in 2003. This was a relatively short-term contract which allowed the issuer to reduce its exposure to 1-in-25-year catastrophe risks (such as a severe outbreak of influenza, a major terrorist attack, or a natural catastrophe). In this regard the bond is addressing a different type of mortality risk from that being considered in this paper. The catastrophe risks being covered by the Swiss Re bond might be correlated with financial markets (past examples include the terrorist attack in New York and Washington on September 11, 2001, or the Kobe earthquake in 1995). In contrast, the types of systematic mortality risks we consider in this paper are assumed to be uncorrelated with the financial markets.

[6]We use in this context the term *longevity risk* to refer specifically to the risk that future survival rates are higher than anticipated. For most of the remainder of the paper we will use the more general term *mortality risk* to refer to all types of deviation from that anticipated in: (a) experienced mortality and survival rates; and (b) mortality tables in use on a given future date.

[7]A small number of survivor swaps have been arranged on an over-the-counter basis. They are not traded contracts and therefore only provide direct benefit to the counterparties in the transaction.

[8]As an example, suppose that $a(t,x)$ represents the market price at time t of a level annuity of £1 per annum payable monthly in arrears to a male aged x at time t. (This might, for example, be a weighted average of the top 5 prices in the market.) It is proposed that a traded futures market be set up with $a(t,x)$ as the underlying instrument for selected values of x and with a selection of maturity dates stretching out many years into the future. For a given maturity date, the market could be closed out some months or even a year before the maturity date itself, to reduce the impact, for example, of moral hazard, changes in expensing bases, or the movements of individual annuity providers in and out of the market.

risks. If the reference population is quite different in nature from the specific mortality risk facing the investor, then the investor (for example, a life office wishing to hedge its exposure) will be exposed to significant basis risk and may conclude that the mortality derivative is not worth holding. Other investors such as speculators and hedge funds may be less interested in using these derivatives for hedging purposes but will be interested in liquidity. Adequate liquidity will require a small number of reference populations, but these will need to be chosen carefully to ensure that the level of basis risk is relatively small for those investors hoping to use the contracts for hedging purposes.

The evolution of the prices of these derivative contracts will reflect as accurately as possible the stochastic evolution of $\mu(t,x)$. In addition, the pricing of these contracts needs to take account of uncertainty over the future values of $\mu(t,x)$. This happens in two ways. Most obviously, stochastic mortality has an impact on the value of mortality options: the greater the volatility in mortality rates, the greater is the value of a mortality option (as with equity options). However, the second effect is more subtle and relates to how financial contracts are priced. For example, the price of a contract based on expected cashflow may not be equal to the value of the contract assuming that mortality follows the median projection. Also (in line with the pricing of financial options) we may calculate expectations using a different probability measure (later in this paper we shall denote this by Q) from the real-world or true measure (which we shall denote by P).

We do not discuss in this paper the many practical issues related to the securitization of mortality risks. These issues are discussed elsewhere (see Cummins, 2004, Dowd *et al.*, 2004, and Lin and Cox, 2004).

It has been evident for many years that mortality rates have been evolving in an apparently stochastic fashion. This is most obvious when one looks at a sequence of mortality curves over a relatively long period of time (see, for example, Figures 1.1 and 1.2 as well as the papers by Forfar and Smith, 1987, Macdonald *et al*, 1998, Willetts, 1999, Macdonald *et al*, 2003, and Currie, Durban and Eilers, 2004). These sequences do exhibit a general trend, but the changes have an unpredictable element not only from one period to the next but also over the long run.

Some recent studies have explicitly modelled the development of the mortality curve over time as a stochastic process. Lee and Carter (1992) introduced a simple model (see Appendix A) for central mortality rates which involves age-dependent and time-dependent terms and applied their model to US population data. (See, also, Lee, 2000b.) The time-dependency is modelled using a univariate ARIMA time-series model implying that changes in the mortality curve at all ages are perfectly correlated. Statistical aspects of this work based on the same model for mortality rates were improved upon by Brouhns, Denuit and Vermunt (2002) who applied the model to Belgian data.

The possibility of imperfect correlation was investigated by Renshaw and Haberman

(2003) who extend the Lee and Carter approach by adding a second time-dependent set of changes. This means that changes in the mortality curve at different ages are no longer perfectly correlated. Another approach based on time series was taken by Felipe *et al.* (1998).

Some alternative approaches are proposed by Lee (2000a) and Yang (2001) (see Appendix B). They take a deterministic projection of the mortality curve, $\hat{q}(t,x)$, as given. They then apply an adjustment to this which evolves over time in a stochastic way.

Milevsky and Promislow (2001) (see Appendix C) take a more theoretical approach in continuous time which assumes that the force of mortality $\mu(t,x)$ has a Gompertz form $\xi_0(t)\exp(\xi_1 x)$ where the $\xi_0(t)$ term only varies over time and is modelled using a simple mean-reverting diffusion process. Dahl (2003) (see Appendix D) also takes a continuous-time approach and models for force of mortality using the affine class of processes. In contrast, this paper does not deal with specific models. Instead we provide a general formulation of the problem in continuous time. These frameworks nort only incorporate all of the models mentioned above but provide a basis for the development of other pricing models in the future.

Lin and Cox (2004) do not propose a specific model for stochastic mortality. Instead, they apply the Wang (1996, 2000, 2002) transform for adjusting the projected mortality rates into risk-neutral probabilities. This approach is gaining in popularity in non-life insurance applications where there is a lack of liquidity in the instruments subject to the underlying risks. However, it is not clear from Lin and Cox (2004) how different transforms for different cohorts and terms to maturity relate to one another to form a coherent whole.

In this paper we will take one step back from the use of specific models and investigate the different frameworks for stochastic mortality that can be used to develop arbitrage-free models to price mortality-linked derivatives. The paper is focused primarily on the pricing of new securities. However, the theory applies equally well to the fair valuation of insurance liabilities which incorporate mortality derivatives.

In Section 2 we introduce the fundamental processes for mortality (the force of mortality process $\mu(t,x)$) and for the risk-free rate of interest ($r(t)$). These processes feed into survivor indices $S(u,y)$ and a risk-free cash account $C(t)$ that play central roles in our analysis. We work with two fundamental types of financial contract:

- pure endowment contracts for a full range of ages *and* terms to maturity; and
- default-free zero-coupon bonds for a full range of terms to maturity.

We then describe, by noting parallels with interest-rate and credit-risk theory, how pure endowment contracts should be priced if they trade in a perfectly liquid, frictionless and arbitrage-free market.[9]

[9] We do not claim that real-world markets are perfectly liquid or frictionless. However, we can

In Sections 3 to 6 we describe the different frameworks that could be employed to build up models for stochastic mortality. Each of these frameworks is drawn from the field of interest-rate modelling but with the risk-free rate of interest $r(t)$ replaced by the force of mortality $\mu(t, x)$. These are all described in theoretical terms: no specific models are proposed or analysed. Rather, the aim of the paper is to leave readers with a choice of frameworks within which they can build their own continuous-time stochastic mortality models.

2 The term structure of mortality

In this section we will define the basic components of a model for stochastic mortality. We start by considering the force of mortality at time t for individuals aged x at time t, which we denote by $\mu(t, x)$. Traditional static mortality models implicitly assume that $\mu(t, x) \equiv \mu(x)$ for all t and x. Deterministic mortality projections imply that $\mu(t, x)$ is a determinstic function of t and x. The models we will consider here will treat $\mu(t, x)$ as a stochastic process.

There are two types of stochastic mortality:

- The first is specific (or unsystematic) mortality risk – the risk that the actual numbers of deaths deviate from anticipated numbers because of the finite number of lives in a given cohort. This type of risk can largely be diversified by investors if the usual assumption that future lifetimes for different individuals are independent random variables is valid.[10] Specific mortality risk therefore does not result in the incorporation of a significant risk premium in the price of mortality derivatives.

- systematic mortality risk – the risk that the force of mortality $\mu(t, x)$ evolves in a different way from that anticipated. This type of risk cannot be diversified away and therefore leads to the incorporation of a risk premium.

It follows that mortality derivatives might be priced using a risk-neutral probability measure, Q, which is different from the real-world probability measure, P.[11]

state that if prices are calculated in the way proposed then even an illiquid market with frictions will be arbitrage-free. Conversely, if we were to propose a pricing framework which violates the conditions in Section 2 then the possibility of arbitrage would emerge over time as the market becomes more liquid or trading costs begin to fall.

[10] Strictly speaking there may be some local dependencies such as thosed between husband and wife, or those between people who die in the same event: particularly the more-significant catastrophe risks of the type being covered by the Swiss Re mortality bond (including, for example, deaths cause by natural disasters or terrorist attacks).

[11] P is sometimes alternatively referred to as the true or objective or physical probability measure.

2.1 Basic building blocks: the survivor index

We have previously indicated that our aim is to develop a set of theoretical frameworks to price mortality derivatives. In order to do so, we will make the assumption that the force of mortality at time t, $\mu(t, y)$, is observable at time t for all y.[12] Furthermore this estimate is only calculated and published some months or years after the event. The length of this delay depends considerably on the reference population: for example, the UK industry-wide Continuous Mortality Investigation tables take longer to compile than tables relating to one specific life office. These are important practical issues but we will leave them for future work.

We will use as our basic building block a family of index-linked zero-coupon survivor bonds. The indexes we will employ are related to survival probabilities for different ages. Thus we define the survivor index

$$S(u, y) = \exp\left(-\int_0^u \mu(t, y+t)dt\right).$$

Looking forwards from time 0 this index is a random variable and not a probability. However, if $\mu(t, x)$ is deterministic then $S(u, y)$ is equal to the probability that an individual aged y at time 0 will survive to age $y + u$. Similarly, if $\mu(t, x)$ is deterministic, for two dates $t_1 < t_2$ the probability that an individual aged x at time t_1 will survive until time t_2 is $S(t_2, x - t_1)/S(t_1, x - t_1)$.

If $\mu(t, x)$ is stochastic then $S(u, y)$ can still be regarded as a survival probability, but one that can only be observed at time u rather than at time 0. However, it is straightforward to extract a survival probability by taking the expectation of a random variable $S(t, x)$ (equation (2.1) below). We prove this by using a combination of indicator random variables and conditional expectation. Thus, consider an individual aged x at time 0. Let $Y_x(u)$ be a Markov chain which is equal to 1 if the individual is still alive at time u. Also let $\mathcal{M}_t$ be the filtration generated by the evolution of the term-structure of mortality, $\mu(u, x)$, up to time t. The real-world or true survival probability, measured at time 0, that an individual aged x at time 0 survives until time u is

$$\begin{aligned} p_P(0, u, x) &= E_P[Y_x(u)] \\ &= E_P[E_P(Y_x(u)|\mathcal{M}_u)] \\ &= E_P[S(u, x)]. \end{aligned} \tag{2.1}$$

More generally we can define the survival probabilities at time t as follows. Let $p_P(t, u, x)$ be the probability under P that an individual aged x at time 0 and still alive at the current time t survives until time u: that is,

$$\begin{aligned} p_P(t, u, x) &= E_P[Y_x(u)|Y_x(t) = 1, \mathcal{M}_t] \\ &= E_P\left[\frac{S(u, x)}{S(t, x)} \,\middle|\, \mathcal{M}_t\right]. \end{aligned}$$

[12] In reality the force of mortality at time t can only be estimated rather than directly observed.

For the alternative risk-neutral probability measure Q, we can define the corresponding survival probabilities:

$$\begin{aligned} p_Q(t,u,x) &= E_Q[Y_x(u)|Y_x(t)=1,\mathcal{M}_t] \\ &= E_Q\left[\frac{S(u,x)}{S(t,x)}\;\middle|\;\mathcal{M}_t\right]. \end{aligned}$$

We are now in a position to consider the pricing of index-linked zero-coupon survivor bonds. There is (potentially) a different bond for each maturity date T and for each age x at time 0. We refer to a specific bond as the (T,x)-bond for compactness.

The (T,x)-bond pays the amount $S(T,x)$ at time T. This payment is well defined in the sense that $S(T,x)$ is an observable quantity at time T. The (T,x)-bond is an example of what financial mathematicians call a *tradeable asset*[13]: that is, an asset that pays no coupons or dividends and whose price at any time $t<T$ represents the total return on an investment in that asset.[14]

To price such bonds we also need to make reference to the term-structure of interest rates. Let $P(t,T)$ represent the price at time t of a zero-coupon bond that pays 1 at time T. The instantaneous forward rate curve at time t is given by

$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T)$$

and the instantaneous risk-free rate of interest is

$$r(t) = \lim_{T\to t} f(t,T)$$

(see, for example, Cairns, 2004b). The cash (or money-market) account invests at the risk-free rate of interest. Its value at time t is denoted by $C(t)$ with

$$\begin{aligned} dC(t) &= r(t)C(t)dt \\ \Rightarrow C(t) &= C(0)\exp\left(\int_0^t r(u)du\right). \end{aligned}$$

Let $\mathcal{F}_t$ be the filtration generated by the term-structure of interest rates up to time t, and $\mathcal{H}_t$ be the combined filtration for both the term-structure of interest rates and mortality. If there exists a measure Q (the risk-neutral measure) equivalent to the real-world measure P with

$$P(t,T) = E_Q\left[\frac{C(t)}{C(T)}\;\middle|\;\mathcal{F}_t\right]$$

[13] To financial economists this would be more commonly known as a pure discount asset.

[14] For an asset that does pay dividends or coupons a tradeable asset can be created by reinvesting the dividends in the underlying asset itself.

(which implies that $P(t,T)/C(t)$ is a Q-martingale) then the dynamics of the zero-coupon bond prices are arbitrage free.

Now let $\tilde{B}(t,T,x)$ represent the price at time t of the (T,x)-bond that pays $S(T,x)$ at time T. If there exists a measure Q equivalent to the real-world measure P with

$$\tilde{B}(t,T,x) = E_Q\left[\frac{C(t)}{C(T)}S(T,x)\;\middle|\;\mathcal{H}_t\right]$$

for all T and x then the dynamics of the index-linked zero-coupon bond prices are arbitrage free. This formula matches those of Milevsky and Promislow (2001) and Dahl (2004) but encompasses a much wider range of models.

Assumption 1

We now make the assumption that the dynamics of the term structure of mortality are independent of the dynamics of the term-structure of interest rates.

This assumption will allow us to separate pricing of mortality risk from pricing of interest-rate risk. It follows that

$$\begin{aligned}
\tilde{B}(t,T,x) &= E_Q\left[\frac{C(t)}{C(T)}\;\middle|\;\mathcal{F}_t\right]E_Q\left[S(T,x)\mid\mathcal{M}_t\right]\\
&= P(t,T)B(t,T,x)\\
\text{where}\quad B(t,T,x) &= E_Q\left[S(T,x)\mid\mathcal{M}_t\right].
\end{aligned}$$

Thus $B(t,T,x)$ is a martingale under Q. We can also assume that the $B(t,T,x)$ processes are strictly positive (barring the possibility of catastrophic events that wipe out the entire population).

This allows us to make three further observations.

- $B(t,T,x)/B(t,t,x) = p_Q(t,T,x)$. Since we can regard the $B(t,T,x)$ as spot prices we will refer to the $p_Q(t,T,x)$ as *spot survival probabilities.*

- We can use the $B(t,T,x)$ to define the *forward force of mortality surface* (we will sometimes shorten this to *forward mortality surface*)

  $$\bar{\mu}(t,T,x+T) = -\frac{\partial}{\partial T}\log B(t,T,x).$$

  Conversely, knowledge of the forward mortality surface allows us to price the bonds as follows:

  $$\frac{B(t,T,x)}{B(t,t,x)} = \exp\left[-\int_t^T \bar{\mu}(t,u,x+u)du\right].$$

  If we take $T=t$, we get the *spot force of mortality*

  $$\mu(t,x+t) = \bar{\mu}(t,t,x+t).$$

- Let us assume that the dynamics of the term structure of mortality are governed by an n-dimensional Brownian motion $\tilde{W}(t)$ under Q. The martingale property of $B(t,T,x)$ together with its positivity allows us to write down the stochastic differential equation for $B(t,T,x)$ in the following form

$$dB(t,T,x) = B(t,T,x)V(t,T,x)'d\tilde{W}(t)$$

where $V(t,T,x)$ is family of previsible vector processes which specify the volatility term structure of the bond prices.

We will now consider the possible frameworks which we can use to model the dynamics of the $B(t,T,x)$ processes. These correspond to a variety of frameworks used in modelling interest rates (see, for example, Cairns, 2004b):

- short-rate models for the dynamics of the $\mu(t,y)$ (which correspond to short-rate models for the risk-free rate of interest, $r(t)$, such as those of Vasicek, 1977, Cox, Ingersoll and Ross, 1985, and Black and Karasinski, 1991);
- forward-mortality models for the dynamics of the forward mortality surface, $\bar{\mu}(t,T,x+T)$ (corresponding to the framework of Heath, Jarrow and Morton, 1992);
- positive mortality models for the spot survival probabilities, $p_Q(t,T,x)$ (corresponding to the positive-interest models of Flesaker and Hughston, 1996, Rogers, 1997, and Rutkowski, 1997);
- market models for forward survival probabilities or forward annuity prices (corresponding to the LIBOR and swap market models of Brace, Gatarek and Musiela, 1997, and Jamshidian, 1997).

Towards the end of the paper we also discuss the parallels between pricing mortality derivatives and credit risk. We note that there are many similarities which allow the transfer to our context of some intensity-based models that have been developed for pricing credit risk.

3 Short-rate models

This type of framework explicitly models the dynamics of $\mu(t,y)$. Thus

$$d\mu(t,y) = a(t,y)dt + b(t,y)'d\tilde{W}(t)$$

where $a(t,y)$ and $b(t,y)$ (an $n \times 1$ vector) are previsible processes and $\tilde{W}(t)$ is a standard n-dimensional Brownian motion under the risk-neutral measure Q. We then have

$$\begin{aligned}\frac{B(t,T,x)}{B(t,t,x)} &= p_q(t,T,x) \\ &= E_Q\left[e^{-\int_t^T \mu(u,x+u)du} \;\middle|\; \mathcal{M}_t \right].\end{aligned}$$

We can make the following observations about this framework:

- We have specified that $\tilde{W}(t)$ and $b(t,y)$ are $n \times 1$ vectors. This means that we can allow for the possibility that short-term changes in the term-structure of mortality can be different at different ages. Different rates of change at different ages can also be achieved through the $a(t,y)$ drift function.
- $a(t,y)$ and $b(t,y)$ might depend on other diffusion processes which are themselves adapted to $\mathcal{M}_t$. Note that this dependence allows $b(t,y) \equiv 0$, in which case the force of mortality curve evolves in a smooth fashion over time. However, the evolution of the force of mortality curve is still stochastic because of its dependence on the stochastic drift rate $a(t,y)$. Other models might assume that $b(t,y) \neq 0$, in which case the force of mortality curve exhibits a degree of local volatility.
- The assumption that $b(t,y) \equiv 0$ is equivalent to the assumption that the volatility function $V(t,T,x)$ for the $B(t,T,x)$ processes tends to zero as $T \to t$. Thus, the shortest-dated bonds will have a very low volatility.
- This framework includes models that assume that $\mu(t,y)$ takes some parametric form (for example, the Gompertz-Makeham model $\mu(t,x) = \xi_0(t) + \xi_1(t)e^{\xi_2(t)x}$). We can model the parameters in this curve as diffusion processes. This class is a specific example of the type noted above where $a(t,y)$ and $b(t,y)$ depend on a number of other diffusion processes.

The framework includes the affine class of models for $\mu(t,x)$ considered by Dahl (2003), under which the spot survival probabilities have the closed form

$$p_Q(t,T,x) = \exp\left[A_0(t,T,x) - A_1(t,T,x)\mu(t,x+t)\right]$$

with $n = 1$ dimension. Dahl provides sufficient conditions on $a(t,y)$ and $b(t,y)$ that result in this affine representation for $p_Q(t,T,x)$. These conditions match those of Duffie and Kan (1996) for interest-rate models (see, also, Vasicek, 1977, and Cox, Ingersoll, and Ross, 1985). An important criterion of mortality models is that the spot survival probability function $p_Q(t,T,x)$ is decreasing in T – otherwise this would imply the potential for future negative mortality rates. One potential drawback of this affine class is that the only models which ensure that $p_Q(t,T,x)$ is decreasing in T requires the use of a mean-reverting process for $\mu(t,y)$. This mean reversion might be towards a time-dependent, but deterministic, local mean-reversion level, in which case mortality improvements can be systematically built into the model. However, if mortality improvements have been faster than anticipated in the past then the mean reversion assumption implies that the potential for further mortality improvements will be significantly reduced in the future. In extreme cases significant past mortality improvements may be reversed if the level of mean reversion is too strong. This is clearly a very strong assumption which is difficult to justify on the basis of previous observed mortality changes and with reference to our perception of the timing and impact of, for example, future medical advances. It may be that one component of $\mu(t, y)$ is mean reverting. However, we believe that another more dominant component of $\mu(t, y)$ should not be subject to mean reversion, since it is impossible to predict the range and pace of future medical advances.

4 Forward mortality models

The next set of models are forward mortality models.

Suppose that we have the two stochastic differential equations

$$\begin{aligned} dB(t,T,x) &= B(t,T,x)V(t,T,x)'d\tilde{W}(t) & (4.1)\\ \text{and}\quad d\bar{\mu}(t,T,x+T) &= \alpha(t,T,x+T)dt + \beta(t,T,x+T)'d\tilde{W}(t) & (4.2) \end{aligned}$$

where $V(t,T,x)$, $\alpha(t,T,x+T)$ and $\beta(t,T,x+T)$ are previsible processes. In the general arbitrage-free modelling context, what is the relationship between $V(t,T,x)$, $\alpha(t,T,x+T)$ and $\beta(t,T,x+T)$? We provide an answer to this question which is similar to that of Heath, Jarrow and Morton (1992) (HJM) approach. However, the present context presents us with a richer and more complex modelling environment with an additional dimension to consider compared with that of the classical HJM framework.

From equation (4.2) we have

$$\begin{aligned} \bar{\mu}(t,T,x+T) &= \bar{\mu}(0,T,x+T) + \int_0^t \alpha(s,T,x+T)ds + \int_0^t \beta(s,T,x+T)'d\tilde{W}(s)\\ \Rightarrow \mu(t,x+t) &= \bar{\mu}(0,t,x+t) + \int_0^t \alpha(s,t,x+t)ds + \int_0^t \beta(s,t,x+t)'d\tilde{W}(s)\\ \text{and}\quad S(t,x) &= \exp\left[-\int_0^t \mu(s,x+s)ds\right]\\ &= \exp\Big[-\int_0^t \bar{\mu}(0,u,x+u)du - \int_0^t\int_u^t \alpha(u,s,x+s)ds\ du\\ &\qquad - \int_0^t\int_u^t \beta(u,s,x+s)'ds\ d\tilde{W}(u)\Big]. \end{aligned}$$

Next note that

$$\begin{aligned} B(t,T,x) &= S(t,x)\exp\left[-\int_t^T \bar{\mu}(t,s,x+s)ds\right] \\ &= \exp\Big[-\int_0^T \bar{\mu}(0,u,x+u)du - \int_0^t\int_u^T \alpha(u,s,x+s)ds\ du \\ &\qquad -\int_0^t\int_u^T \beta(u,s,x+s)'ds\ d\tilde{W}(u)\Big]. \end{aligned} \tag{4.3}$$

Now define $V(u,T,x) = -\int_u^T \beta(u,s,x+s)'ds$. We can then apply Ito's formula to $B(t,T,x)$ in equation (4.3) to get the SDE

$$\begin{aligned} dB(t,T,x) &= B(t,T,x)\Big[\left(\frac{1}{2}|V(t,T,x)|^2 - \int_t^T \alpha(t,s,x+s)ds\right)dt \\ &\qquad + V(t,T,x)'d\tilde{W}(t)\Big]. \end{aligned}$$

(This confirms our earlier claim that $V(t,T,x) = -\int_u^T \beta(u,s,x+s)'ds$.)

Now we require the drift under Q to be zero. Therefore

$$\frac{1}{2}|V(t,T,x)|^2 = \int_t^T \alpha(t,s,x+s)ds$$

and by taking the partial derivative with respect to T we get

$$\alpha(t,T,x+T) = -V(t,T,x)'\beta(t,T,x+T).$$

As with the other frameworks, the challenge is to specify an appropriate form for $\beta(t,T,x+T)$ or $V(t,T,x)$. The chosen formulation needs to ensure that the forward mortality surface remains strictly positive. This is most easily achieved by making $\beta(t,T,x+T)$ explicitly dependent on the current forward mortality surface. In addition, the chosen form needs to ensure that the spot force of mortality curve, $\mu(t,y)$, retains an appropriate shape (for example, generally increasing with age). How these criteria can be met we leave for further work!

5 The positive mortality framework

We now turn to our third class of models, the positive mortality framework.

Let $\tilde{P}$ be some measure equivalent to Q, and let $A(t,x)$ be some family of $\mathcal{M}_t$ adapted, strictly-positive supermartingales.

Define

$$p_Q(t,T,x) = \frac{B(t,T,x)}{B(t,t,x)} = \frac{E_{\tilde{P}}[A(T,x)|\mathcal{M}_t]}{A(t,x)}. \tag{5.1}$$

The strict positivity of $A(t,x)$ means that $p_Q(t,T,x)$ is positive. The supermartingale property of $A(t,x)$ ensures that the $p_Q(t,T,x)$ are less than or equal to 1 and decreasing in $T > t$. It is straightforward to demonstrate (for example, through the application of the Radon-Nikodym derivative $dQ/d\tilde{P}$) that the resulting dynamics of $B(t,T,x)$ are appropriate for an arbitrage-free pricing model (see, also, Rogers, 1997, and Rutkowski, 1997). Within this pricing framework, the drift of $A(t,x)$ under $\tilde{P}$ is equal to $-\mu(t,x+t) \times A(t,x)$. (In the corresponding positive-interest model the drift of $A(t)$ is equal to $-r(t) \times A(t)$ – see, for example, Cairns, 2004b.)

Equation (5.1) is deceptively simple as a pricing formula. However, the effort comes in specifying a model for the processes $A(t,x)$ and in calculating the expectations. (For examples in interest-rate modelling see Flesaker and Hughston, 1996, Rogers, 1997, and Cairns, 2004a.)

A special case of this framework is an adaptation of Flesaker and Hughston (1996) (FH). Let $N(t,s,x)$ for $0 < t < s$ be a family of strictly-positive martingales under $\tilde{P}$. Define

$$A(t,x) = \int_t^\infty N(t,s,x)ds.$$

The martingale property of $N(t,s,x)$ means that

$$\begin{aligned} E_{\tilde{P}}[A(T,x)|\mathcal{M}_t] &= \int_T^\infty N(t,s,x)ds &\quad (5.2)\\ &< A(t,x). &\quad (5.3)\end{aligned}$$

It follows from (5.3) that $A(t,x)$ satisfies the Rogers/Rutkowski requirements for a strictly-positive supermartingale.

Combining equations (5.2) and (5.1) we now get

$$p_Q(t,T,x) = \frac{\int_T^\infty N(t,s,x)ds}{\int_t^\infty N(t,s,x)ds}.$$

This turns the problem into one of devising an appropriate model for the family $N(t,s,x)$. From a computational point of view this involves, at worst, the numerical evaluation of a one-dimensional integral. However, the challenge remains to specify a suitable model of martingales for $N(t,s,x)$.

It is common in interest-rate-derivatives markets to calibrate the initial term structure of the model to the observed interest-rate term structure. We can also apply this approach to the mortality term structure. Suppose then that we take as given at time 0 the market prices of the zero-coupon bonds, $P(0,T)$, and the (T,x)-bonds, $\tilde{B}(0,T,x)$, for all x and $T > 0$. From this we can derive the implied spot survival probabilities $p_Q(0,T,x) = \tilde{B}(0,T,x)/P(0,T)$. The initial values for the family $N(t,T,s)$ can then be calibrated as follows

$$N(0,T,x) = -\frac{\partial}{\partial T} p_Q(0,T,x) = \bar{\mu}(0,T,x+T)p_Q(0,T,x).$$

This initial calibration is unique up to a strictly-positive, constant scaling factor.

By analogy with interest-rate modelling, this framework might contain natural model formulations that are difficult to identify in other frameworks. Again this is left for further work.

6 Mortality market models

6.1 Introduction: change of numeraire

We come now to the mortality market models, and begin with some preliminaries about the type of model.

Recall that the processes $B(t, T, x)$ in a zero-interest environment are martingales under Q with SDE's

$$dB(t, T, x) = B(t, T, x)V(t, T, x)'d\tilde{W}(t)$$

for appropriate previsible volatility functions $V(t, T, x)$.[15]

Now consider some, strictly-positive, tradeable assets as numeraires. As a specific first example consider $B(t, \tau, y)$ as the numeraire. We then consider processes of the type

$$Z(t, T, x) = \frac{B(t, T, x)}{B(t, \tau, y)}.$$

For most problems we are likely to consider it is likely that the most productive choice of y will be x itself (since then $Z(\tau, \tau, x) = p_Q(\tau, T, x)$). If we then apply Ito's formula and the Product Rule we find that

$$dZ(t, T, x) = Z(t, T, x)\Big(V(t, T, x) - V(t, \tau, x)\Big)'\big(d\tilde{W}(t) - V(t, \tau, x)dt\big).$$

Now define a new process $W^{\tau,x}(t) = \tilde{W}(t) - \int_0^t V(s, \tau, x)ds$. Provided that $V(t, \tau, x)$ satisfies the Novikov condition we can use the Girsanov theorem (see, for example, Karatzas and Shreve, 1998) to infer that there exists a measure $P_{\tau,x}$ equivalent to Q under which $W^{\tau,x}(t)$ is a standard Brownian motion. We then have

$$dZ(t, T, x) = Z(t, T, x)\Big(V(t, T, x) - V(t, \tau, x)\Big)'dW^{\tau,x}(t)$$

and we see that $Z(t, T, x)$ is a martingale under $P_{\tau,x}$.

[15]Readers who are familiar with interest-rate market models can consider *unity* as being the numeraire.

6.2 The annuity market model

6.2.1 Zero interest

For simplicity we will restrict ourselves initially to a market where interest rates are set to zero. Non-zero stochastic interest rates will be added later.

Let

$$F(t,x) = \frac{B(t,T,x)}{\sum_{s=T+1}^{\infty} B(t,s,x)}$$

be a forward annuity rate under which survivors at T pay £1 at T and receive back $F(t,x)$ at times $T+1, T+2, \ldots$ so long as they are still alive at each of those times.[16] In the assumed zero-interest environment this contract has zero value at time t. Note specifically that $F(T,x) = 1/\sum_{s=T+1}^{\infty} p_Q(T,s,x)$ is the spot (market) annuity rate at T.

This suggests the use of a different numeraire $X(t) = \sum_{s=T+1}^{\infty} B(t,s,x)$. Since $X(t)$ is a stricly-positive martingale we can write its SDE as

$$dX(t) = X(t)V_X(t)'d\tilde{W}(t)$$

for an appropriate previsible volatility function $V_X(t)$. Then

$$\begin{aligned} dF(t,x) &= F(t,x)\Big(V(t,T,x) - V_X(t)\Big)'\Big(d\tilde{W}(t) - V_X(t)dt\Big) \\ &= F(t,x)\gamma(t,x)'dW^X(t) \end{aligned}$$

where $\gamma(t,x) = \Big(V(t,T,x) - V_X(t)\Big)$ and $W^X(t) = \tilde{W}(t) - \int_0^t V_X(s)ds$ is a standard Brownian motion under an appropriate measure P_X equivalent to Q.

The standard modelling assumption for market models is to specify that $\gamma(t,x)$ is a deterministic function. It follows in this case that $F(s,x)$ for $t < s \leq T$ is log-normally distributed under P_X with

$$\begin{aligned} E_{P_X}[F(s,x)|\mathcal{M}_t] &= F(t,s) \\ \text{and} \quad Var_{P_X}[\log F(s,x)|\mathcal{M}_t] &= \int_t^s |\gamma(u,x)|^2 du. \end{aligned}$$

Now consider an annuity contract which includes a guaranteed annuity rate. In the open market £1 at time T will secure a pension of $F(T,x)$ per annum from time T payable annually in arrears (assuming no expenses and a fair price). The contract

[16] Readers may be more familiar with a deferred annuity contract. Under this contract survivors at t pay £1 at t in return for a defined series of payments at times $T+1, T+2, \ldots$ payable only to those who are still alive at each of those times. In contrast, with the forward contract the purchase price is not paid until time T, and then only by those who are still alive at that time.

also includes a gaurantee that the amount of the pension will be K per annum (the guaranteed annuity rate) if this rate is higher than the open market rate.

When we wish to value the guarantee we need to consider carefully the nominal amount being converted into an annuity at T. We claim that the appropriate amount is $S(T,x)$. To see why, suppose that we have a group of $N(t,x)$ lives at time t aged $x+t$. At time T, $N(T,x)$ of these individuals will still be alive. Suppose that each of these survivors will have available a nominal amount of £1 for conversion into an annuity at T. Then, given $\mathcal{M}_T$, $N(T,x)$ will have a binomial distribution with parameters $N(t,x)$ and $S(T,x)/S(t,x)$ and expected value $\kappa S(T,x)$ where $\kappa = N(t,x)/S(t,x)$. Justification for our claim is concluded with the developments leading up to equation (6.1) below where we see that $N(T,x)$ is conditionally independent of the mortality table in use at time T. This allows us to replace $N(T,x)$ by $\kappa S(T,s)$.

The total value of the contract at T is

$$N(T,x)\max\{F(T,x),K\}\sum_{s=T+1}^{\infty} p_Q(T,s,x).$$

The value of the option itself at T is therefore

$$G(T) = \frac{N(T,x)(K-F(T,x))_+}{F(T,x)}.$$

Now the option itself is a tradeable asset with price $G(t)$ at time t, so $G(t)/X(t)$ is a P_X martingale. Hence

$$\begin{aligned}
\frac{G(t)}{X(t)} &= E_{P_X}\left[\frac{G(T)}{X(T)}\;\middle|\;\mathcal{M}_t\right] \\
&= E_{P_X}\left[\frac{N(T,x)(K-F(T,x))_+}{F(T,x)X(T)}\;\middle|\;\mathcal{M}_t\right] \\
&= E_{P_X}\left[E_{P_X}\left(\frac{N(T,x)(K-F(T,x))_+}{F(T,x)X(T)}\;\middle|\;\mathcal{M}_T\right)\;\middle|\;\mathcal{M}_t\right] \\
&= E_{P_X}\left[E_{P_X}\left(N(T,x)\mid\mathcal{M}_T\right)\frac{(K-F(T,x))_+}{F(T,x)X(T)}\;\middle|\;\mathcal{M}_t\right] \\
&= E_{P_X}\left[\kappa S(T,x)\frac{(K-F(T,x))_+}{F(T,x)X(T)}\;\middle|\;\mathcal{M}_t\right] \qquad (6.1) \\
&= E_{P_X}\left[\kappa(K-F(T,x))_+\mid\mathcal{M}_t\right] \\
\Rightarrow G(t) &= \kappa X(t)\left(K\Phi(-d_2)-F(t,x)\Phi(-d_1)\right) \\
\text{where }\; d_1 &= \frac{\log F(t,x)/K+\frac{1}{2}\sigma_F^2}{\sigma_F} \\
d_2 &= d_1-\sigma_F \\
\sigma_F^2 &= \int_t^T|\gamma(u,x)|^2du
\end{aligned}$$

and $\Phi(z)$ is a cumulative distribution function for the standard Normal distribution.

We have explained here how the swap market model can be adapted to mortality modelling. However, it remains for an empirical study to determine whether the assumption of a deterministic $\gamma(u,x)$ is reasonable or not.

6.2.2 Stochastic interest rates

Now consider the case with stochastic interest. In this case we have

$$\begin{aligned} dP(t,s) &= P(t,s)(r(t)dt + V_P(t,s)'d\tilde{Z}(t)) \\ dB(t,s,x) &= B(t,s,x)V_B(t,s,x)'d\tilde{W}(t) \end{aligned}$$

where $\tilde{Z}(t)$ and $\tilde{W}(t)$ are independent Brownian motions. Application of the product rule gives us

$$d(P(t,s)B(t,s,x)) = P(t,s)B(t,s,x)(r(t)dt + V_P(t,s)d\tilde{Z}(t) + V_B(t,s,x)d\tilde{W}(t)).$$

Now consider the annuity contract described above with a guaranteed annuity rate of K. The actual annuity rate at time T per £1 at T is $F(T,x)$ where

$$F(t,x) = \frac{P(t,T)B(t,T,x)}{\sum_{s=T+1}^{\infty} P(T,s)B(T,s,x)}.$$

(With $t=T$ this equates to $F(T,x) = 1/\sum_s P(T,s)p_Q(T,s,x)$.) This suggests the use of the numeraire

$$X(t) = \sum_{s=T+1}^{\infty} P(T,s)B(T,s,x)$$

with

$$\begin{aligned} dX(t) &= X(t)\left[r(t)dt + V_{PX}(t)d\tilde{Z}(t) + V_{BX}(t)d\tilde{W}(t)\right] \\ \text{where } V_{PX}(t) &= X(t)^{-1}\sum_{s=T+1}^{\infty} V_P(t,s)P(t,s)B(t,s,x) \\ \text{and } V_{BX}(t) &= X(t)^{-1}\sum_{s=T+1}^{\infty} V_B(t,s,x)P(t,s)B(t,s,x). \end{aligned}$$

Under the measure P_X, the prices of all tradeable assets discounted by $X(t)$ are martingales. Specifically this implies that $F(t,x)$ is a P_X-martingale with SDE under P_X

$$dF(t,x) = F(t,x)[\gamma_P(t,x)dZ^X(t) + \gamma_B(t,x)dW^X(t)]$$

for suitable previsible processes $\gamma_P(t,x)$ and $\gamma_B(t,x)$. In the annuity market model we assume that $\gamma_P(t,x)$ and $\gamma_B(t,x)$ are deterministic functions. As before we assume that the nominal amount to be converted into an annuity at T is $S(T,x)$. (The earlier argument converting actual numbers of lives surviving to T into $S(T,x)$ applies equally well here.) The value of the option component is denoted by $G(t)$ with

$$G(T) = \frac{S(T,x)(K - F(T,x))_+}{F(T,x)}.$$

The martingale property implies that

$$\begin{aligned}
\frac{G(t)}{X(t)} &= E_{P_X}\left[\frac{G(T)}{X(T)} \,\middle|\, \mathcal{H}_t\right] \\
&= E_{P_X}\left[\frac{S(T,x)(K - F(T,x))_+}{P(T,T)B(T,T,x)} \,\middle|\, \mathcal{H}_t\right] \\
&= E_{P_X}\left[(K - F(T,x))_+ | \mathcal{H}_t\right].
\end{aligned}$$

With the assumption that the volatility functions $\gamma_P(t,x)$ and $\gamma_B(t,x)$ are determinstic this gives us the pricing formula

$$\begin{aligned}
G(t) &= X(t)\left(K\Phi(-d_2) - F(t,x)\Phi(-d_1)\right) \\
\text{where } d_1 &= \frac{\log F(t,x)/K + \frac{1}{2}\sigma_F^2}{\sigma_F} \\
d_2 &= d_1 - \sigma_F \\
\text{and } \sigma_F^2 &= \int_t^T (|\gamma_P(u,x)|^2 + |\gamma_B(u,x)|^2)du.
\end{aligned}$$

It can be seen, therefore, that the annuity-market model offers a simple but powerful tool which can allow us to tackle some important questions involving annuity guarantees. Again, though, the underlying assumptions need to be tested against historical data to see if the framework is a reasonable one or not.

6.3 The SCOR market model

We will now consider a market model which looks directly at annualised forward mortality rates. This type of model is less tractable than the annuity market model (Section 6.2) if we wish to use it to value annuity guarantees. On the other hand, this type of model can be applied much more easily to a wider class of product.

As before we will start by considering the situation where interest rates are equal to zero. We now define the concept of survival credits (see, also, Blake, Cairns and Dowd, 2003). These are, in effect, bonuses payable to survivors within a pool of life office policyholders in a way which ensures that no systematic profits or losses accrue to the life office. The survival credit payable to survivors at $t+1$ is calculated at time t by the life office based on the latest mortality tables available at time t.

In the event, actual survivorship from t to $t+1$ may differ from that anticipated, so the variation or risk over that year is borne by the life office.

The risk-neutral survival probability from t to $t+1$ measured at time t is $p_Q(t,t+1,x)$ and this implies that the actuarially and financial-economically fair survival credit payable at $t+1$ is

$$\frac{1-p_Q(t,t+1,x)}{p_Q(t,t+1,x)}.$$

This represents a fair subdivision (as far as it can be anticipated at time t) of the amount invested at t by those who die before $t+1$ amongst those who survive to $t+1$.[17]

This survivor credit is reminiscent of the τ-LIBOR (the London Interbank Offer Rate with duration or tenor τ) in the money markets which is equal (in a world with non-zero interest rates) to $L=(1-P(t,t+\tau))/\tau P(t,t+\tau)$. The τ-LIBOR contract states that for each £1 deposited at t, £$1+\tau L$ will be returned at $t+\tau$. For this reason we will refer to

$$L(T_{k-1},T_{k-1},T_k,x)=\frac{1-p_Q(T_{k-1},T_k,x)}{(T_k-T_{k-1})p_Q(T_{k-1},T_k,x)} \tag{6.2}$$

as the *Survivor Credit Offer Rate* (or *SCOR*). In general we will assume that $T_k-T_{k-1}=1$ for all k.[18]

Note also that we can rewrite (6.2) as

$$L(T_{k-1},T_{k-1},T_k,x)=\frac{B(T_{k-1},T_{k-1},x)-B(T_{k-1},T_k,x)}{(T_k-T_{k-1})B(T_{k-1},T_k,x)}.$$

This allows us to define the forward SCOR as follows

$$L(t,T_{k-1},T_k,x)=\frac{B(t,T_{k-1},x)-B(t,T_k,x)}{(T_k-T_{k-1})B(t,T_k,x)}. \tag{6.3}$$

Under a forward SCOR contract arranged at t we are fixing in advance the survivor credit that will be payable at T_k to survivors at T_k. That is for each £1 payable by survivors at T_{k-1}, those still alive at T_k will receive £$1+(T_k-T_{k-1})L(t,T_{k-1},T_k,x)$ at T_k. This will have zero value at t using the risk-neutral pricing approach discussed in Section 2.

[17] This is rather like a pool of annuitants as considered by Blake, Cairns and Dowd (2003), but here we are not making any assumptions about how much income is paid out of the fund to the survivors. However, Blake, Cairns and Dowd do not consider in detail the possibility of stochastic mortality.

[18] Note that, while $p_Q(T_{k-1},T_k,x)$ must lie between 0 and 1, $L(T_{k-1},T_{k-1},T_k,x)$ can lie between 0 and ∞. This means that we can model $L(T_{k-1},T_{k-1},T_k,x)$, if we so choose, as a log-normal random variable.

For simplicity of notation in what follows, let us assume that $T_k - T_{k-1} = 1$ for all k and denote

$$L_k(t) \equiv L(t, T_{k-1}, T_k, x).$$

From equation (6.3) we see that $L_k(t)$ is equal to the value of a tradeable asset or portfolio $(B(t, T_{k-1}, x) - B(t, T_k, x))$ with the tradeable asset $B(t, T_k, x)$ as the numeraire. As noted at the start of this section on market models, this implies that there exists a measure P_{T_k} equivalent to Q under which the prices of all tradeable assets divided by $B(t, T_k, x)$ are martingales and under which $W^{T_k}(t) = \tilde{W}(t) - \int_0^t V_B(u, T_k, x)du$ is a standard Brownian motion.

Application of the Product Rule to $L_k(t)$ (following a similar argument in Cairns, 2004b, Section 9.1) gives us

$$\begin{aligned} dL_k(t) &= \frac{B(t, T_{k-1}, x)}{B(t, T_k, x)} \left(V_B(t, T_{k-1}, x) - V_B(t, T_k, x)\right)' \left\{d\tilde{W}(t) - V_B(t, T_k, x)dt\right\} \\ &= L_k(t) V_L(t)' dW^{T_k}(t) \end{aligned} \tag{6.4}$$

where

$$\begin{aligned} W^{T_k}(t) &= \tilde{W}(t) - \int_0^t V_B(u, T_k, x)du \\ \text{and} \quad V_{Lk}(t) &\equiv V_L(t, T_{k-1}, T_k, x) \\ &= \left(V_B(t, T_{k-1}, x) - V_B(t, T_k, x)\right) \frac{(1 + L_k(t))}{L_k(t)}. \end{aligned} \tag{6.5}$$

With reference to equation (6.4), first we note that the martingale property implies that $E_{P_{T_k}}[L_k(u)|\mathcal{M}_t] = L_k(t)$ for $t < u < T_{k-1}$. Second, if we make the usual market model assumption that $V_{Lk}(t)$ is a deterministic function then $L_k(u)$, given $\mathcal{M}_t$ for $t < u < T_{k-1}$, is log-normal under P_{T_k} with $Var_{P_{T_k}}[\log L_k(u)|\mathcal{M}_t] = \int_t^u |V_{Lk}(s)|^2 ds$.

Equation (6.5) can be rearranged to give

$$V_B(t, T_{k-1}, x) - V_B(t, T_k, x) = \frac{L_k(t)}{1 + L_k(t)} V_{Lk}(t). \tag{6.6}$$

Bearing in mind the relationship between the $W^{T_k}(t)$ and $\tilde{W}(t)$, for $l > k$ we can use (6.6) to show that

$$dW^{T_l}(t) = dW^{T_k}(t) + \sum_{j=k+1}^{l} \frac{L_j(t)}{1 + L_j(t)} V_{Lj}(t)dt.$$

Thus simulating under P_{T_1} we have

$$dL_1(t) = L_1(t) V_{L1}(t) dW^{T_1}(t)$$

and for $k > 1$

$$dL_k(t) \quad = \quad L_k(t)V_{Lk}(t)\left(dW^{T_1}(t) + \sum_{j=2}^{k} \frac{L_j(t)}{1+L_j(t)} V_{Lj}(t)dt\right).$$

In addition we can simulate under the real world measure P by replacing $dW^{T_1}(t)$ by $dW(t) + \lambda(t)dt$ for a suitable process $\lambda(t)$.

Once again, this class of model offers us a powerful toolkit. Again, though, we need to test potential models against historical data. Equally, there are challenges in choosing a suitable market-price-of-risk process $\lambda(t)$ which is statistically justifiable and which leaves the model reasonably tractable.

7 Credit risk models

Finally, we consider the last class of models, the credit-risk models.

To start, we note that the zero-coupon survivor bond with price $\tilde{B}(t, T, x)$ at time t is similar to a zero-coupon corporate bond which pays 1 at T if there has been no default and 0 if the bond has defaulted. There are many models which address the problem of how to price such bonds (see, for example, the textbooks by Schönbucher, 2003, or Lando, 2004). In the present context, the most useful models for default risk which could be translated into a stochastic mortality model are *intensity-based* models (see, for example, Schönbucher, 2003, Chapter 7). In these models the default intensity, $\lambda(t)$ corresponds to the force of mortality $\mu(t, x+t)$. Thus from the theoretical point of view pricing can be approached in the same way.

However, there are some differences between mortality risk and credit risk which means that the types of model employed might be different:

- In a credit risk context different companies are equivalent to different cohorts in the mortality model. However, there is no reason why, from the structural point of view, that individual companies should be linked to each other in the way that adjacent cohorts are.
- The default intensity is likely to be modelled as a mean-reverting process that is also possibly time-homogeneous. In contrast, mortality models are certainly time inhomogeneous and need to incorporate non-mean-reverting elements. This has the important implication that Cox, Ingersoll and Ross (1985)-type models can be used for credit-risk models, but not for mortality-based models.
- The default intensity is likely to be correlated with the interest-rate term structure.

We can conclude that credit-risk modelling does have something to offer us in the mortality context. However, we need to proceed with caution instead of blindly applying some credit-risk models that have unsuitable charateristics.

8 Conclusions

We have presented here a variety of theoretical frameworks that could be used for pricing many different types of mortality derivatives. These frameworks can provide other researchers with the basis for the development of specific models for stochastic mortality.

We take as granted that many of the assumptions underpinning these frameworks (such as liquid, frictionless markets) do not hold in practice. Nevertheless we can still state that, in such imperfect markets, *if* prices evolve in the way suggested by these pricing frameworks, then the model will be arbitrage free. We are not assuming that the market must be complete, or that transactions costs must be zero or that assets are infinitely divisible, and so on.

The challenges for future research are as follows. The first two challenges relate to the models themselves:

1. We need to investigate a range of specific stochastic mortality models. Do they give an adequate statistical description of the past? Do they satisfy certain "reasonableness" criteria in terms of their potential future dynamics and mortality curve shapes? Which models are straightforward to implement numerically?

2. There is also the related issue of the number of risk factors. Is one risk-factor adequate or do we need to have two or more factors so that we can have imperfectly correlated mortality improvements at different ages? From the discrete- and continuous-time models described in the introduction only that of Renshaw and Haberman (2003) considers the possibility of a second factor. One can easily argue that two or more factors might be desirable since different medical advances, for example, are likely to have different impacts on different parts of the mortality curve.

Then we have practical issues relating to the time lag:

3. How do we allow for the time lag between the measurement date and the date when mortality rates for that date have been graduated and made public? Is there something that can be learned here from catastrophe derivatives, where information gradually emerges after a catastrophe event? One can certainly argue that information about mortality rates at some time T_0 will emerge

between T_0 and the publication date T_1 of the official results through other *ad hoc* sources.[19]

Finally, there are contract design issues:

4. What contract designs are likely to prove successful? Success can be measured in different ways. In one sense this can be measured in terms of the amount of risk transferred. In another sense it can be measured in terms of market liquidity and volumes of business.

   Both of these measures are related to basis risk. Some securities are issued by individual insurers (using a Special Purpose Vehicle) and are designed to minimise basis risk for the issuer (for example, the Swiss Re survivor bond). Investors in the security are assumed to be buying the security as a means of diversifying their portfolios and will, therefore, be less concerned about basis risk. If the aim is a create a highly-liquid market then many investors will be aiming to use the security to hedge their mortality risks. The contract will need to make reference to some standard reference population. This reference population will need to be close enough to those of the majority of individual insurers to keep basis risk down. If basis risk is substantial for all but a small minority of insurers then it is unlikely that a liquid market will develop.

   Connected to this issue is the need to specify the contract in a way which minimises the moral hazard associated with time lags in the release of information (that is, the possibility of insider trading). This is a critical issue. For a healthy and liquid investors need to feel that they have adequate protection against insiders or that moral hazard is negligible. Some early attempts to introduce catastrophe derivatives failed to excite the market precisely because of this issue. We must learn from these past experiences to increase our chances of getting the design mortality derivatives right first time round.

In summary, the challenges ahead of us are substantial. We need to develop good stochastic models, we need reliable and timely mortality indices, and we need good contract design. However, the time is more right than ever for the introduction of mortality derivatives and this offers exciting times ahead for those of us who choose to take up the challenge!

Acknowledgements

AC would like to thank David Wilkie for some invaluable discussions on the topic of mortality improvements and Sam Cox for some useful insights. We are also grateful

[19]Note that the adequately-subscribed Swiss Re bond discussed earlier in this paper makes reference to national population statistics. We can infer that the linkage in the contract between payments and time-lagged indices was satisfactory for investors.

to Iain Currie for providing the data in Figure 1.1.

References

Black, F., and Karasinski, P. (1991) "Bond and option pricing when short rates are log-normal," Financial Analysts Journal July-Aug, 52-59.

Blake, D., Cairns, A.J.G., and Dowd, K. (2003) "Pensionmetrics II: Stochastic pension plan design during the distribution phase," Insurance: Mathematics and Economics 33, 29-47.

Blake, D., and Burrows, W. (2001) "Survivor bonds: helping to hedge mortality risk," Journal of Risk and Insurance 68, 339-348.

Brace, A., Gatarek, D., and Musiela, M. (1997) "The market model of interest-rate dynamics," Mathematical Finance 7, 127-155.

Brigo, D., and Mercurio, F. (2001) *Interest Rate Models: Theory and Practice.* Springer: Berlin.

Brouhns, N., Denuit, M., and Vermunt J.K. (2002) "A Poisson log-bilinear regression approach to the construction of projected life tables," Insurance: Mathematics and Economics 31, 373-393.

Cairns, A.J.G. (2004a) "A family of term-structure models for long-term risk management and derivative pricing," Mathematical Finance To appear.

Cairns, A.J.G. (2004b) *Interest-Rate Models: An Introduction.* Princeton University Press: Princeton.

CMI (1978) "Proposed standard tables for life office pensioners and annuitants," Continuous Mortality Investigation Reports 3, 1-30.

CMI (1999) "Projection factors for mortality improvement," Continuous Mortality Investigation Reports 17, 89-108.

Cox, J., Ingersoll, J., and Ross, S. (1985) "A theory of the term-structure of interest rates," Econometrica 53, 385-408.

Cummins, J.D. (2004) "Securitization of life insurance assets and liabilities," Working Paper, The Wharton School.

Currie I.D., Durban, M. and Eilers, P.H.C. (2004) "Smoothing and forecasting mortality rates," Working Paper, Heriot-Watt University.

Dahl, M. (2003) "Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts," Working paper, Laboratory of Actuarial Mathematics, University of Copenhagen..

Dowd, K., Blake, D., Cairns, A.J.G., and Dawson, P. (2004) "Survivor swaps," In preparation.

Duffie, D., and Kan, R. (1996) "A yield-factor model of interest rates," Mathematical Finance 6, 379-406.

Felipe, A., Guilén, M., and Artis, M. (1998) "Recent mortality trends in the Spanish population," Transactions of the 26th International Congress of Actuaries, Birmingham 9, 55-74.

Flesaker, B., and Hughston, L.P. (1996) "Positive interest," Risk 9(1), 46-49.

Forfar, D.O., and Smith, D.M. (1987) "The changing shape of English life Tables," Transactions of the Faculty of Actuaries 40, 98-134.

Heath, D., Jarrow, R. and Morton, A. (1992) "Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation," Econometrica 60, 77-105.

James, J., and Webber, N. (2000) *Interest Rate Modelling.* Wiley: Chichester.

Jamshidian, F. (1997) "LIBOR and swap market models and measures," Finance and Stochastics 1, 293-330.

JMIC (1974) "Considerations affecting the preparation of standard tables of mortality," Journal of the Institute of Actuaries 101, 135-201.

Karatzas, I., and Shreve, S.E. (1998) *Methods of Mathematical Finance.* Springer-Verlag: New York.

Lando, D. (2004) *Credit Risk Modeling: Theory and Applications.* Princeton University Press: Princeton.

Lee, P.J. (2000a) "A general framework for stochastic investigations of mortality and investment risks," Presented at the Wilkiefest, Heriot-Watt University, March 2000.

Lee, R. (2000b) "The Lee-Carter method for forecasting mortality, with various extensions and applications," North American Actuarial Journal 4, 80-93.

Lee, R.D., and Carter, L.R. (1992) "Modeling and forecasting U.S. mortality," Journal of the American Statistical Association 87, 659-675.

Lin, Y., and Cox, S.H. (2004) "Securitization of mortality risks in life annuities," Working paper, Georgia State University.

Macdonald, A.S., Cairns, A.J.G., Gwilt, P.L., and Miller, K.A. (1998) "An international comparison of recent trends in population mortality," British Actuarial Journal 4, 3-141.

Macdonald, A.S., Bartlett, D., Berman, C., Daykin, C., Grimshaw, D., Savill, P. (2003) "Mortality improvements and the cohort effect," CMI Working Papers 1 and 2; Presented to the Staple Inn Actuarial Society on 11 March 2003; 46 pages. (Available online at www.sias.org.uk.)

Milevsky, M.A., and Promislow, S.D. (2001) "Mortality derivatives and the option to annuitise," Insurance: Mathematics and Economics 29, 299-318.

Rebonato, R. (2002) *Modern Pricing of Interest-Rate Derivatives.* Princeton University Press: Princeton.

Renshaw, A.E., and Haberman, S. (2003) "Lee-Carter mortality forecasting with age-specific enhancement," Insurance: Mathematics and Economics 33, 255-272.

Rogers, L.C.G. (1997) "The potential approach to the term-structure of interest rates and foreign exchange rates," Mathematical Finance 7, 157-164.

Rutkowski, M. (1997) "A note on the Flesaker & Hughston model of the term structure of interest rates," Applied Mathematical Finance 4, 151-163.

Schönbucher, P. (2003) *Credit Derivatives Pricing Models: Models Pricing and Implementation.* Wiley: Chichester.

Vasicek, O. (1977) "An equilibrium characterisation of the term structure," Journal of Financial Economics 5, 177-188.

Wang, S. (1996) "Premium calculation by transforming the layer premium density," ASTIN Bulletin 26, 71-92.

Wang, S. (2000) "A class of distortion operations for pricing financial and insurance risks," Journal of Risk and Insurance 67, 15-36.

Wang, S. (2002) "A universal framework for pricing financial and insurance risks," ASTIN Bulletin 32, 213-234.

Willetts, R.C. (1999) "Mortality in the next millenium," Presented to the Staple Inn Actuarial Society on 7 December 1999. (Available online at www.sias.org.uk.)

Yang, S. (2001) "Reserving, pricing and hedging for guaranteed annuity options," PhD Thesis, Heriot-Watt University, Edinburgh.

Appendices

A The Lee and Carter model for stochastic mortality

Lee and Carter (1992) investigate the dynamics of the observed central mortality rates $m(t,x)$ for integer t and x. Their model breaks $m(t,x)$ down into a log-bilinear model

$$\log m(t,x) = a(x) + b(x)k(t)$$

with the translation and scaling constraints that $\sum_x b(x) = 1$ and $\sum_{t=T_0}^{T_1} k(t) = 0$. $a(x)$ and $b(x)$ are non-parametric functions without any smoothing applied or functional form. $k(t)$ is estimated directly from the data without any assumption about its dynamic form

B The Lee and Yang model for stochastic mortality

Lee (2000a) and Yang (2001) proposed the following model for stochastic mortality.

Suppose that a deterministic forecast of annual mortality rates is made at time 0. Thus $\hat{q}(x,t)$ represents the probability (as estimated at time 0) that an individual aged x at time t will die before time $t+1$ for each integer x and t. The actual mortality experience is modelled as

$$\begin{aligned} q(x,t) &= \hat{q}(x,t) \exp\left[X(t) - \frac{1}{2}\sigma_Y^2 + \sigma_Y Z_Y(t)\right] \\ \text{where } X(t) &= X(t-1) - \frac{1}{2}\sigma_X^2 + \sigma_X Z_X(t) \end{aligned}$$

and $Z_X(t)$ and $Z_Y(t)$ are mutually independent sequences of i.i.d. standard normal random variables.

It follows that the $X(t)$ models the stochastic trend in the development of the mortality curve while the $-\frac{1}{2}\sigma_Y^2 + \sigma_Y Z_Y(t)$ models one-off environmental variations in mortality (such as a major flu epidemic). From the limited data available Yang found that σ_Y was not significantly different from 0.

C The Milevsky and Promislow model for stochastic mortality

Milevsky and Promislow (2001) model the force of mortality in the form $\mu(t,x) = \xi_0 \exp(\xi_1 x + Y_t)$ where Y_t is an Ornstein-Uhlenbeck process with SDE $dY_t = -\alpha Y_t dt + \sigma dW_t$. Essentially this translates into a Gompertz model with a time-varying scaling factor.

D The Dahl model for stochastic mortality

Dahl (2003) models the process for $\mu(t, x+t)$ as follows

$$d\mu(t,x+t) = \alpha^\mu(t,x,\mu(t,x+t))dt + \sigma^\mu(t,x,\mu(t,x+t))d\tilde{W}(t).$$

He finds that if the drift and volatility are of the form

$$\begin{aligned} \alpha^\mu(t,x,\mu(t,x+t)) &= \delta^\alpha(t,x)\mu(t,x+t) + \zeta^\alpha(t,x) \\ \text{and} \quad \sigma^\mu(t,x,\mu(t,x+t)) &= \sqrt{\delta^\sigma(t,x)\mu(t,x+t) + \zeta^\sigma(t,x)} \end{aligned}$$

for some deterministic functions $\delta^\alpha(t,x)$, $\delta^\sigma(t,x)$, $\zeta^\alpha(t,x)$ and $\zeta^\sigma(t,x)$ then

$$p_Q(t,T,x) = e^{A(t,T,x) - B(t,T,x)\mu(t,x+t)}$$

where the deterministic functions $A(t,T,x)$ and $B(t,T,x)$ are derived from differential equations involving $\delta^\alpha(t,x)$, $\delta^\sigma(t,x)$, $\zeta^\alpha(t,x)$ and $\zeta^\sigma(t,x)$.

Andrew J.G. Cairns
Department of Actuarial Mathematics and Statistics
School of Mathematical and Computer Sciences
Heriot-Watt University
Edinburgh, EH14 4AS
United Kingdom

David Blake
Cass Business School
106 Bunhill Row
London, EC1Y 8TZ
United Kingdom

Kevin Dowd
Centre for Risk & Insurance Studies
Nottingham University Business School
Jubilee Campus
Nottingham, NG8 1BB
United Kingdom