THE FAIR VALUE OF GUARANTEED ANNUITY OPTIONS

Enrico Biffis°, Pietro Millossovich•
Department of Applied Mathematics, University of Trieste
Piazzale Europa, 1 - 34127 Trieste - Italy
Tel.: +39 040 5587117; Fax: +39 040 54209
°enricob@econ.units.it •pietrom@econ.units.it

Abstract. We discuss the fair valuation of Guaranteed Annuity Options, i.e. options providing the right to convert deferred survival benefits into annuities at fixed conversion rates. The use of doubly stochastic stopping times and of affine processes provides great computational and analytical tractability, while enabling to set up a very general valuation framework. For example, the valuation of options on traditional, unit-linked or indexed annuities is encompassed. Moreover, security and reference fund prices may feature stochastic volatility or discontinuous dynamics. The longevity risk is also taken into account, by letting the evolution of mortality present stochastic dynamics subject not only to random fluctuations but also to systematic deviations.

Keywords: fair value, options to annuitize, stochastic mortality, longevity risk, financial risk, doubly stochastic stopping times, affine processes.

1. Introduction

Recent mortality and interest rate trends have proved to be particularly dangerous for the pricing of insurance contracts providing long term living benefits (pensions and annuities), as well as for their reserving. On the mortality side, the increase of life expectancy and the decline in mortality rates at adult-old ages have made clear the importance of dealing explicitly with the so-called longevity risk, i.e. the risk of systematic departures from expected levels of mortality. Due to its non-pooling character, the possibility of benefiting from offsetting effects by holding a large enough portfolio of policies is precluded (see Olivieri (2001)). This problem has been amplified by a general decline in interest rates over the last years, affecting in particular those contracts providing joint financial and demographic guarantees. Among these, we focus on Guaranteed Annuity Options (GAOs), i.e. contracts providing the holder the right to convert deferred survival benefits, possibly unit-linked or indexed, into an annuity at a fixed conversion rate. The underpricing of such guarantees has caused several solvency problems to insurers, for example in the UK, where Equitable Life (the world's oldest life insurer) had to close to new business in 2000.

The pricing of GAOs has been tackled by several authors, e.g. by Van Bezooyen, Exley and Mehta (1998), Milevsky and Promislow (2001), Ballotta and Haberman (2003b,a), Boyle and Hardy (2003), Olivieri and Pitacco (2003), Pelsser (2003) and Wilkie, Waters and Yang (2003). Here, we provide a fairly general framework that includes and extends most of the models appeared. The aim is to deal effectively with several sources of risk, including asset, interest rate and mortality (systematic and unsystematic) risk, all this by allowing for (insurance) security price dynamics featuring stochastic volatility or jumps.

Date: April 20, 2004.

We exploit a parallel between defaultable securities and insurance contracts first suggested in Artzner and Delbaen (1995) and exploited in Milevsky and Promislow (2001) and Biffis (2003, 2004). We may regard annuities, for example, as defaultable coupon bonds (or consols) with zero recovery. Of course, the interpretation of the default event is reverted: an annuity stops being paid because an insured (the receiver of the benefits) dies, while coupons from a bond cease to be paid because the issuer (the one who pays coupons and principal) defaults. The valuation framework of intensity based models (relying on the exogenous specification of the conditional probability of default, given that default has not yet occurred: see Duffie (2001, Ch. 11)) can be extended to insurance securities in a sense to be made precise in the following.

We treat explicitly the issue of the exercise boundary involved by GAOs. Indeed, GAOs are exercised in a primary market environment, while triggering cash flows that one may want to value from an internal or secondary market perspective. In particular, the exercise decision made by the policyholder may not be rational from the insurer's point of view. We allow for such issue in the model, putting particular emphasis on the market-oriented accounting standards proposed by the International Accounting Standards Board (IASB).

In Sec. 2, we provide a formal description of GAOs and of the fair valuation setup proposed by the IASB. In Sec. 3, we introduce affine processes and show why they are appealing from the analytical and computational point of view. Sec. 4 and Sec. 5 describe the financial market and the mortality model. The latter exploits doubly stochastic stopping times driven by affine jump-diffusions in order to obtain closed form expressions (up to ordinary differential equations solutions) for survival probabilities. For other proposals regarding the issue of stochastic mortality, see for example Milevsky and Promislow (2001), Ballotta and Haberman (2003a) or Dahl (2003). In Sec. 6, the stochastic valuation framework is described with indications regarding the extent to which no-arbitrage type valuations can be used in the insurance market of concern. Valuation expressions for GAOs are provided in Sec. 6.1. Sec. 6.2 extends the setup to death benefits and indexed annuities. Concluding remarks are offered in Sec. 7. In what follows, we refer the reader to: Bowers *et al.* (1997) for actuarial terminology and notation; Duffie (2001) for no-arbitrage pricing; Brémaud (1981) for point processes; Protter (1990) and Jacod and Shiryaev (2003) for background on stochastic processes and integration.

2. Guaranteed Annuity Options

As introduced in the previous section, GAOs are contracts providing the policyholder with the right to convert a deferred survival benefit into an annuity at a fixed conversion rate. The option is exercised at maturity, conditional on survival, if the cash benefit then available is greater than the *value to the policyholder* of the annuity payable throughout his/her remaining life time, the annuity amounts being determined by the conversion rate specified at inception. Usually, the moneyness of such option at maturity happens to depend not only on the price of annuities available in the primary market at that time (determined by interest rates, mortality levels and charges then prevailing), but also on policyholders' preferences and expectations not captured by market prices.

Formally, let us consider a policyholder entering the contract at time 0, then aged x years, and denote its residual life time by τ. Moreover, suppose the amount of the cash benefit available at maturity, i.e. at time $T > 0$, is given by the value of a reference fund with price process S and that a guaranteed conversion rate $0 < g < 1$ is fixed at inception. Then, if we denote by $a_t \doteq \ddot{a}_{x+T}(t)$ the time-t 'value' process of an annuity payable from time T to a policyholder then aged $x + T$, the 'value' of the contract at time T, V_T, is given by the following expression:

$$\begin{aligned} V_T &= \mathbb{I}_{\{\tau>T\}} \left[S_T \mathbb{I}_{A^c} + g S_T a_T \mathbb{I}_A\right] \\ &= \mathbb{I}_{\{\tau>T\}} S_T + \mathbb{I}_{\{\tau>T\}\cap A}\ a_T S_T \left[g - 1/a_T\right] \\ &= \mathbb{I}_{\{\tau>T\}} S_T + \mathbb{I}_{\{\tau>T\}\cap A}\ g S_T \left[a_T - 1/g\right] \end{aligned} \tag{1}$$

where A denotes the exercise set of the option, A^c its complement and $\mathbb{I}_B$ the indicator function of a set B. Note that in the second line the quantity $1/a_T$ is the conversion rate corresponding to the annuity 'value' a_T, while the quantity $1/g$ is the 'value' of the guaranteed unitary annuity implied by g. Thus, the analysis of V_T can be tackled equivalently in terms of conversion rates or annuity values.

The setup outlined in expression (1) is quite general. First, we have not specified what kind of 'value' V_T represents, whether an internal (to the insurance company) or external (in terms of financial statement) quantification of the payoff, or the result of a primary or secondary market valuation. Second, in the valuation of the cash flows of concern, we explicitly allow for the previously outlined mismatch between the viewpoint of the issuer of the contract (short position) and that of the buyer (long position), as is made clear by the reference to an exercise set not necessarily coinciding with $\{a_T \geq 1/g\}$.

For example, when no distinction is made between primary and secondary markets, when the exercise is assumed to be rational and to depend on interest rates and mortality levels as seen equivalently by the insurer or the insureds, then $A = \{a_T \geq 1/g\}$. This is the case of all papers related to GAOs of which we are aware, where the maturity value of the contract is given by:

$$\begin{aligned} V_T &= \mathbb{I}_{\{\tau>T\}} \max\{g S_T a_T, S_T\} \\ &= \mathbb{I}_{\{\tau>T\}} S_T + \mathbb{I}_{\{\tau>T\}}\ \ddot{a}_T S_T \max\{g - 1/a_T, 0\} \\ &= \mathbb{I}_{\{\tau>T\}} S_T + \mathbb{I}_{\{\tau>T\}}\ g S_T \max\{a_T - 1/g, 0\} \end{aligned} \tag{2}$$

The expressions appearing in (2) make clear the interpretation of GAOs in terms of standard (vanilla) options: in the second line we have a put option written on the conversion rate $1/a_T$ with strike g, while in the third line we have a call option on the annuity value a_T with strike the guaranteed annuity value $1/g$. In our framework, for example, the valuation of the last term in (2) parallels that of a vulnerable option on a defaultable coupon bearing bond with zero recovery. The results provided in Sec. 6 include and generalize this particular case: specifically we will see GAOs as vulnerable digital options providing random payoffs, as the expressions in (1) show.

In order to provide a concrete example of application of the setting described, we make reference to the IASB's proposals for market-value oriented accounting of insurance liabilities, as described for example in IASB (2001). The IASB defines the *fair value* of a book of contracts as the exchange price in a (hypothetical) secondary market transaction. Since a deep wholesale market for books of contracts does not exist at the moment, exchange prices cannot be easily observed. The IASB favours the use of an expected discounted cash flow approach consistent with risk-neutral valuation. Indeed, 'no-arbitrage type' arguments imply that the fair value of an insurance liability should not be different from the market value of a portfolio of traded assets matching the liability cash flows with a sufficient degree of certainty. According to IASB (2001), all sources of risk should be taken into account, including non-financial risks, both diversifiable and non fully diversifiable (e.g. the longevity risk); therefore, the computation of suitable margins for such risks, the so-called market value margins, is needed. The guidance on the form of market-value margins is fairly broad: our valuation framework expresses them in terms of adjustments to the riskless short rate and/or to best-estimate assumptions.

In computing the 'fair value' of the payoff V_T described by (1), a_T would represent the 'fair value' of an annuity payable from time T to a policyholder then aged $x+T$, a value different in general from the price charged in the primary market. Moreover, exercise probabilities, which are of course crucial in fair valuation (see Sec. 6), may be based on company/market statistical data and possibly incorporate suitable adjustments for systematic risk. The setting allows explicitly for the fact that the insurer's risk-neutral probabilities are in general different from those of policyholders, in particular when concerning the evolution of mortality. Moreover, adjustments for 'irrational' exercise of the GAOs can be made at a very detailed level, allowing for more accurate sensitivity analysis for example.

3. Affine Processes

Affine processes are essentially Markov processes with conditional characteristic function of the exponential affine form. They are thoroughly treated in Duffie, Filipovič and Schachermayer (2003) and Filipovič (2001). In this section, we adopt a narrower but more intuitive perspective that is usually adopted in financial applications. Following closely (also for terminology and notation) Duffie and Kan (1996) and Duffie, Pan and Singleton (2000), we define affine processes in terms of strong solutions to specific stochastic differential equations (SDE) in a given filtered probability space.

Given a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, an $\mathbb{R}^n$-valued *affine jump-diffusion process* X is an $\mathbb{F}$-Markov process specified as the unique strong solution of the following Stochastic Differential Equation (SDE):

$$dX_t = \delta(t, X_t)dt + \sigma(t, X_t)dW_t + \sum_{h=1}^{m} dJ_t^h, \tag{3}$$

where W is an $\mathbb{F}$-standard Brownian motion in $\mathbb{R}^n$ and each J^h is a pure-jump process in $\mathbb{R}^n$ with jump-arrival intensity $\{\lambda^h(t, X_t) : t \geq 0\}$ and jump distribution ν_t^h on $\mathbb{R}^n$. We require the drift δ, the instantaneous covariance matrix $\sigma\sigma^\intercal$ and the jump-arrival intensities (λ^h) to have all affine dependence on X. In more explicit terms, the affine dependence requires that the coefficients appearing in the SDE (3) have the following form:

$$\delta(t, x) = d_0(t) + d_1(t)x \tag{4}$$

$$\left(\sigma(t, x)\sigma(t, x)^\intercal\right)_{i,j} = (V_0(t))_{i,j} + (V_1(t))_{i,j} \cdot x \qquad i, j = 1, \ldots, n \tag{5}$$

$$\lambda^h(t, x) = l_0^h(t) + l_1^h(t) \cdot x \qquad h = 1, \ldots, m \tag{6}$$

where $c \cdot d = \sum_{j=1}^n c_j d_j$ for all $c, d \in \mathbb{C}^n$. The functions $d \doteq (d_0, d_1)$, $V \doteq (V_0, V_1)$ and $l^h \doteq (l_0^h, l_1^h)$ are defined on $[0, \infty)$, take values respectively in $\mathbb{R}^n \times \mathbb{R}^{n\times n}$, $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$ and $\mathbb{R} \times \mathbb{R}^n$, and are assumed to be bounded and continuous. Moreover, the jump-size distribution of the h-th jump process is determined by its Laplace transform (called *jump transform* henceforth):

$$\theta^h(t,c) = \int_{\mathbb{R}^n} e^{c\cdot z} d\nu_t^h(z),$$

defined for $t \in [0,\infty)$, $c \in \mathbb{C}^n$ and such that the integral is finite. The jump transform $\theta \doteq (\theta^h)$ and the functions d, V and $l \doteq (l^h)$ completely determine the distribution of X, once an initial condition X_0 is given.

An important result holds for analytical approaches based on the affine structure described. For any $a, b \in \mathbb{C}^n$, $c \in \mathbb{C}$ and for given $T \geq t$ and affine function R defined by $R(t,x) = \rho_0(t) + \rho_1(t) \cdot x$ (for some bounded continuous $\mathbb{R} \times \mathbb{R}^n$-valued function $\rho \doteq (\rho_0, \rho_1)$), the transform ϕ, defined by:

$$\phi(a,b,c,X_t,t,T) = E\left[e^{-\int_t^T R(s,X_s)ds}\, e^{a\cdot X_T}\,(b\cdot X_T + c)\,\Big|\mathcal{F}_t\right], \tag{7}$$

admits the following representation, under technical conditions provided in Duffie, Pan and Singleton (2000, Prop. 3):

$$\phi(a,b,c,X_t,t,T) = e^{\alpha(t)+\beta(t)\cdot X_t}\left(\hat{\alpha}(t) + \hat{\beta}(t)\cdot X_t\right), \tag{8}$$

where $\alpha(\cdot) \doteq \alpha(\cdot; a, T)$ and $\beta(\cdot) \doteq \beta(\cdot; a, T)$, are functions solving uniquely the following ordinary differential equations (ODEs):

$$\dot{\beta}(t) = \rho_1(t) - d_1(t)^\mathsf{T}\beta(t) - \frac{1}{2}\beta(t)^\mathsf{T}V_1(t)\beta(t) - \sum_{h=1}^m l_1^h(t)\left[\theta^h(t,\beta(t)) - 1\right] \tag{9}$$

$$\dot{\alpha}(t) = \rho_0(t) - d_0(t)\cdot\beta(t) - \frac{1}{2}\beta(t)^\mathsf{T}V_0(t)\beta(t) - \sum_{h=1}^m l_0^h(t)\left[\theta^h(t,\beta(t)) - 1\right] \tag{10}$$

with boundary conditions $\alpha(T) = 0$ and $\beta(T) = a$, while the functions $\hat{\alpha}(\cdot) \doteq \hat{\alpha}(\cdot; a, b, c, T)$ and $\hat{\beta}(\cdot) \doteq \hat{\beta}(\cdot; a, b, T)$ solve the ODEs:

$$\dot{\hat{\beta}}(t) = -d_1(t)^\mathsf{T}\hat{\beta}(t) - \beta(t)^\mathsf{T}V_1(t)\hat{\beta}(t) - \sum_{h=1}^m l_1^h(t)\left[\Theta^h(t,\beta(t))\cdot\hat{\beta}(t)\right] \tag{11}$$

$$\dot{\hat{\alpha}}(t) = -d_0(t)\cdot\hat{\beta}(t) - \beta(t)^\mathsf{T}V_0(t)\hat{\beta}(t) - \sum_{h=1}^m l_0^h(t)\left[\Theta^h(t,\beta(t))\cdot\hat{\beta}(t)\right] \tag{12}$$

with boundary conditions $\hat{\alpha}(T) = c$ and $\hat{\beta}(T) = b$, where $\Theta^h(t,c)$ denotes the gradient of $\theta^h(t,c)$ with respect to $c \in \mathbb{C}^n$, i.e. $\Theta^h(t,c) = \int_{\mathbb{R}^n} c\exp(c\cdot z)\nu_t^h(dz)$. We remind that for all $c, d \in \mathbb{C}^n$ the vector in $\mathbb{C}^n$ with k-th element $\sum_{i,j} c_i(V_1(t))_{ijk}d_j$ is denoted by $c^\mathsf{T}V_1(t)d$. Note that for $b = 0$ and $c = 1$ the transform ϕ has exponential affine form.

The analytical tractability of affine processes is essentially linked to the ODEs associated with the transform (7). These ODEs are generalized Riccati equations that can be solved by using standard numerical methods. For some choice of (d, V, l, θ, ρ), explicit solutions are available. They are derived, for example, in the simple case of the Vasicek (1977) and Cox, Ingersoll and Ross (1985) models without jumps. When including jumps, explicit solutions may be available when the jump-arrival process is Poisson, depending on the specification of ν. In the Vasicek case (Poisson-Gaussian jump-diffusion), the choice of the jump distribution can be very general. In the Cox, Ingersoll and Ross case (square-root jump-diffusion), the choice is more restricted, due to the non-negativity requirement for X: closed-form solutions are available with degenerate (fixed jump size), uniform, exponential and binomial jump-size distributions (see Duffie and Kan (1996)).

In many applications, including empirical estimation, we are interested in the valuation of a modified version of transform (7). For fixed $T \in [0, \infty)$ and for $a, b, d \in \mathbb{R}^n$ and $c \in \mathbb{R}$, let $G \doteq G_{a,b,c,d,\rho}(y; X_0, T)$ be the function defined by:

$$G(y) = E\left[e^{-\int_0^T R(s,X_s)ds} e^{a \cdot X_T} \left(b \cdot X_T + c\right) \mathbb{I}_{\{d \cdot X_T \leq y\}}\right] \tag{13}$$

The function $G(\cdot)$ is of finite variation on compacts (on the whole real line, provided integrability conditions hold) and can be treated as the distribution function of a signed measure on $\mathbb{R}$. The Fourier-Stieltjes transform $\mathcal{G}(\cdot)$ of $G(\cdot)$ is a special case of (7), in that:

$$\mathcal{G}(u) \doteq \int_{-\infty}^{+\infty} e^{iuy} dG(y) = \phi(a + iud, b, c, X_0, 0, T),$$

and can be inverted by using the Lévy Inversion Formula. For example, in the case $b = 0$ and $c = 1$, the following holds, under technical conditions provided in Duffie, Pan and Singleton (2000, Prop. 2):

$$G_{a,0,1,d,\rho}(y; X_0, T) = \frac{\phi(a, 0, 1, X_0, 0, T)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\mathrm{Im}[\phi(a + iud, 0, 1, X_0, 0, T) e^{-iuy}]}{u} du \tag{14}$$

where $\mathrm{Im}(c)$ denotes the imaginary part of $c \in \mathbb{C}^n$.

We now move on to the issue of changes of measure. In particular, we are interested in whether a process X that is affine under a probability measure $\mathbb{P}$ remains affine after a change to an equivalent probability measure $\widetilde{\mathbb{P}}$ and, if that is the case, in the form of the new dynamics. For fixed $T > 0$, under technical conditions guaranteeing that (8) holds for $b = 0$ and $c = 1$ (see Duffie, Pan and Singleton (2000, Prop. 1)), the process $\xi_t \doteq \xi_t^{(a,T)}$ defined by:

$$\xi_t = e^{-\int_0^t R(s,X_s)ds} e^{\alpha(t;a,T) + \beta(t;a,T) \cdot X_t} \tag{15}$$

is a positive martingale, and a probability measure $\widetilde{\mathbb{P}} \doteq \mathbb{P}_T^{(a)}$ equivalent to $\mathbb{P}$ can be defined on $(\Omega, \mathcal{F})$ by setting its density equal to ξ_T / ξ_0. The notation used underlines the dependence of the measure change not only on a given time horizon $T > 0$, but also on the final condition $\beta(T) = a$ of the ODEs (9) and (10). (Note, however, that ξ depends also on (d, V, l, θ, ρ).) This will be important for valuation purposes, since a convenient choice of the measure $\widetilde{\mathbb{P}}$ will enable to simplifies the payoffs under conditional expectations in order to exploit (8).

Duffie, Pan and Singleton (2000, Prop. 5) show that X remains affine under the new measure $\widetilde{\mathbb{P}}$ with new dynamics $(\widetilde{d}, \widetilde{V}, \widetilde{l}, \widetilde{\theta})$ defined by:

$$\widetilde{d}_0(t) = d_0(t) + V_0(t)\beta(t) \qquad \widetilde{d}_1(t) = d_1(t) + V_1(t)\beta(t) \tag{16}$$

$$\widetilde{l}_0^h(t) = l_0^h(t)\ \theta^h(t, \beta(t)) \qquad \widetilde{l}_1^h(t) = l_1^h(t)\ \theta^h(t, \beta(t)) \tag{17}$$

$$\widetilde{V}(t) = V(t) \qquad \widetilde{\theta}^h(t, c) = \frac{\theta^h(t, c + \beta(t))}{\theta^h(t, \beta(t))} \tag{18}$$

where $h = 1, \ldots, m$, $c \in \mathbb{C}^n$, $t \in [0, T]$, $\beta(t) \doteq \beta(t; a, T)$ and $\widetilde{V}_1(t)w$ denotes the $n \times n$ matrix with k-th column $\widetilde{V}_1(t)^k w$ for $w \in \mathbb{R}^n$.

4. Financial Market

We fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and take as given an adapted short-rate process r (such that $\int_0^t |r_s|ds < \infty$ for every $t \geq 0$, $\mathbb{P}$-a.s.) representing the continuously compounded rate of interest on riskless securities. This can be formalized by assuming the presence in the market of a money-market account, a security with price process B defined by $B_t = \exp(\int_0^t r_s ds)$ and representing the amount of money available at time t from investing one unit at time 0 in risk-free deposits and 'rolling over' the proceeds until t.

Moreover, we assume that at least a security is traded continuously in the market, with a nonnegative semimartingale S representing its (ex-dividend) price. The absence of arbitrage is essentially equivalent to the existence of an equivalent martingale measure $\mathbb{Q}$ (see Harrison and Kreps (1979) and Delbaen and Schachermayer (1994)) under which the gain (from holding the security) process is a martingale after deflation by the money-market account. Specifically, let D be a nonnegative semimartingale representing the security cumulated dividend process. Then, the discounted gain process is given by $(B_t^{-1} S_t + \int_0^t B_s^{-1} dD_s)$ and the following convenient formula applies:

$$S_t = E^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} S_T + \int_t^T e^{-\int_t^u r_s ds} dD_u \middle| \mathcal{F}_t \right]. \tag{19}$$

In what follows, we will assume that the price of any security is zero after a given time $t > 0$ if the securities pays no dividends thereafter. If the security has dividend yield process ζ, i.e. the instantaneous yield from holding the security is $\zeta_t S_t dt$, then $D_t = \int_0^t \zeta_u S_u du$ and the martingale property implies that S has drift $r - \zeta$ under $\mathbb{Q}$, justifying the appellation 'risk-neutral' for this measure. When considering several securities, including zero-coupon bonds, the no-arbitrage restriction imposed by (19) must apply simultaneously to each security price process. From now on, we assume that the dynamics of all security processes are specified under $\mathbb{Q}$ unless otherwise stated.

We then postulate that all security prices are driven by a Markov state vector X following an affine (jump-)diffusion in $\mathbb{R}^k$. Specifically, we assume that r is expressed as $r_t \doteq r(t, X_t)$, where X_t satisfies (3) and the conditions there given and r is an affine function defined by $r(t, x) = p_0(t) + p_1(t) \cdot x$, with $p \doteq (p_0, p_1)$ an $\mathbb{R} \times \mathbb{R}^k$-valued bounded continuous function on $[0, \infty)$. For fixed $T \geq t > 0$, the time t-price $B(t, T)$ of a zero coupon bond with maturity T (i.e. a security paying a single dividend equal to 1 at time T), has the following expression by (19) and (8):

$$B(t, T) = E^{\mathbb{Q}} \left[e^{-\int_t^T r(s, X_s) ds} \middle| \mathcal{F}_t \right] = e^{\alpha(t;0,T) + \beta(t;0,T) \cdot X_t} \tag{20}$$

Moreover, we consider a risky security and assume that its log-price is also affine, in the sense that $\log(S) = X^i$, where X^i denotes the i-th component of X. Moreover, let S have an affine dividend-yield process $\zeta(t, X_t) = q_0(t) + q_1(t) \cdot X_t$, with $q \doteq (q_0, q_1)$ an $\mathbb{R} \times \mathbb{R}^k$-valued bounded continuous function on $[0, \infty)$. In the absence of arbitrage, the dynamics of S must obey the restrictions implied by (19). By applying Itô's formula to S and forcing the drift to be equal to $r - \zeta$ under $\mathbb{Q}$, we get the following conditions, on the lines of Duffie, Pan and Singleton (2000, Sec. 3.1):

$$(d_0(t))_i = p_0(t) - q_0(t) - \frac{1}{2}(V_0(t))_{i,i} - \sum_{h=1}^{m} l_0^h(t)\left[\theta^h(t,\epsilon(i)) - 1\right] \tag{21}$$

$$(d_1(t))_i = p_1(t) - q_1(t) - \frac{1}{2}(V_1(t))_{i,i} - \sum_{h=1}^{m} l_1^h(t)\left[\theta^h(t,\epsilon(i)) - 1\right] \tag{22}$$

where $\epsilon(i)$ indicates the vector in $\mathbb{R}^k$ with all null components but the i-th, which is equal to 1. Similar restrictions must be imposed for any additional risky security of this type considered in the market.

The setup outlined is fairly general: security price processes featuring stochastic volatility or discontinuous dynamics are naturally included. For example, the financial market models used by Heston (1993), Bates (1997), Scott (1997), Bakshi, Cao and Chen (1997) and Bakshi and Madan (2000) are all encompassed by the framework described.

5. Stochastic Mortality Modeling

We fix a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and focus on a generic individual aged x at time 0, whose random residual lifetime is modeled as an $\mathbb{F}$-stopping time τ_x. If the individual belongs to a homogeneous group of persons (in particular, of the same age and with the same health status) whose random residual life times can be considered independent and identically distributed, the results obtained extend to all individuals belonging to the homogeneous population.

We look at τ_x as at the first jump-time of a non-explosive $\mathbb{F}$-counting process N with random intensity μ_x, and say that τ_x has intensity μ_x. One may regard N as recording at each time $t \geq 0$ whether the individual has died ($N_t \neq 0$) or not ($N_t = 0$). The idea behind the specification of a random intensity of mortality μ_x is that, at any time $t \geq 0$ and state $\omega \in \Omega$ such that $\tau_x(\omega) > t$, we have:

$$\mathbb{P}(\tau_x \leq t + \Delta | \mathcal{F}_t)(\omega) \cong \mu_x(t,\omega)\Delta \tag{23}$$

This is the stochastic analogous of the expression for the 'instantaneous death probability', which is familiar to actuaries and usually emerges when defining the deterministic intensity itself. However, here the setup is both dynamic and stochastic: at time $t \geq 0$, the random intensity is expressed as $\mu_x(t,\omega)$, where $\omega \in \Omega$ is the 'state of the world' determining the particular trajectory of μ_x, while t is the continuous-time counterpart of the calendar year of reference used in longitudinal tables. As a consequence, we remark that we are naturally adopting a 'diagonal' (or cohort-based) approach.

From now on, we drop reference to the age x and set $\tau = \tau_x$ and $\mu_t(\omega) = \mu_x(t,\omega)$. For analytical tractability, we assume that N is a doubly stochastic process driven by a subfiltration $\mathbb{G}$ of $\mathbb{F}$ and with $\mathbb{G}$-predictable intensity μ. The intuitive meaning of such assumption is that, conditional on any given trajectory $\mu(\cdot,\omega)$ of μ (for fixed $\omega \in \Omega$), the counting process N associated with τ becomes Poisson-inhomogeneous with parameter $\int_0^{\cdot} \mu(s,\omega)ds$. In other words, for all $T \geq t \geq 0$ and nonnegative integer k, we have:

$$\mathbb{P}\left(N_T - N_t = k | \mathcal{F}_t \vee \mathcal{G}_T\right) = \frac{\left(\int_t^T \mu_s ds\right)^k}{k!} e^{-\int_t^T \mu_s ds}. \tag{24}$$

We note that the time of death τ is an $\mathbb{F}$-stopping time, but not necessarily a $\mathbb{G}$-stopping time: the idea behind the specification of a $\mathbb{G}$-predictable intensity μ is that $\mathbb{G}$ carries enough information about the likelihood of death happening, but not about the actual occurrence of death, which in turn is carried by the larger filtration $\mathbb{F}$. Note that assuming the time of death has an intensity implies that τ is a 'totally inaccessible stopping time'. Intuitively, that means the individual's death comes as a total surprise, in an informational sense.

Within this framework, we can use the law of iterated expectations and (24) to express the time-t survival probabilities over the time horizon $(t, T]$ (for fixed $T \geq t \geq 0$), and on the event $\{\tau > t\}$, as follows:

$$\mathbb{P}(\tau > T|\mathcal{F}_t) = E\left[e^{-\int_t^T \mu_s ds}\middle|\mathcal{F}_t\right]. \tag{25}$$

This can be compared with its deterministic analogous, ${}_{T-t}p_{x+t}$, that can be found in standard actuarial texts, e.g. in Bowers *et al.* (1997). When computing expectations of some functionals of τ, its density may be needed. In the doubly stochastic framework, under technical conditions provided in Grandell (1976, pp. 105-107), the $\mathcal{F}_t$-conditional density $f_t(\cdot)$ is given, on the set $\{\tau > t\}$, by the expression:

$$f_t(s) = \frac{\partial}{\partial s}\mathbb{P}\left(\tau \leq s|\mathcal{F}_t\right) = E\left[\mu_s\ e^{-\int_t^s \mu_u du}\middle|\mathcal{F}_t\right] \tag{26}$$

It is clear that both (25) and (26) are particular cases covered by the transform (7), so that working in an affine framework would be very convenient from the computational point of view. We therefore take an affine jump-diffusion Y in $\mathbb{R}^d$, i.e. a process solving the SDE (3) and satisfying the conditions there given. We then set $\mathbb{G} = \mathbb{G}^Y$ (with $\mathbb{G}^Y$ denoting the natural filtration of Y) and consider $\mu_t = \mu(t, Y_{t-})$ for some function $\mu(t, y) = \eta_0(t) + \eta_1(t) \cdot y$, where $\eta \doteq (\eta_0, \eta_1)$ is an $\mathbb{R} \times \mathbb{R}^d$-valued bounded continuous function on $[0, \infty)$ such that the intensity μ_t is nonnegative.

The computational advantages offered by this setup are remarkable. For example, survival probabilities as those defined by (25) have the following convenient expression, on the event $\{\tau > t\}$:

$$\mathbb{P}(\tau > T|\mathcal{F}_t) = e^{\alpha(t)+\beta(t)\cdot Y_t} \tag{27}$$

where the functions $\alpha(\cdot) \doteq \alpha(\cdot; 0, T)$ and $\beta(\cdot) \doteq \beta(\cdot; 0, T)$, respectively $\mathbb{R}$ and $\mathbb{R}^d$-valued, solve the ODEs (9) and (10) with boundary conditions $\alpha(T) = 0$ and $\beta(T) = 0$.

We note that the process Y may include observable as well as unobservable variables driving the evolution of the intensity. In the former case, we could include suitable 'markers', such as life expectancy, death rates referring to particular ages or the 'entropy' measure (see Keyfitz (1982, 1985), for example). In the latter case, we may think of generic unobservable factors, whose dynamics may be estimated by calibration to 'target' marker values. We refer the reader to Piazzesi (2003, Sec. 6) and references therein for an overview of the most common methods used for the estimation of affine jump-diffusions.

In actuarial analysis and computations, it appears very convenient to make explicit reference to suitable deterministic 'demographical bases', such as available mortality tables, or 'best-estimate' and 'prudential' assumptions about the evolution of mortality. Best-estimate assumptions are realistic (sometimes called 'true') hypotheses representing unbiased expectations about the future based on the best available company/industry/market information. Prudential assumptions are conservative expectations, often used for reserving or pricing purposes. The choice of which demographic basis to use depends on the purpose of analysis. To this regard, we note that we have been very general so far about the nature of the probability measure $\mathbb{P}$, which will be meant as a 'risk-neutral' measure when performing market valuations, as a prudential or 'on the safe-side' demographical basis when carrying out conservative reserving or pricing. In any case, the doubly stochastic property of τ and the affine dynamics of μ are assumed to be specified under the probability measure of concern.

We can implement what just described in different ways. We provide two examples of affine intensity, one including suitable deterministic assumptions directly into the parameterization of μ, the other focusing on random departures from an initially chosen demographic basis or mortality table. As a first example, we consider a two-dimensional square-root diffusion $Y = (\mu, \bar{\mu})$ with the first component representing the random intensity of mortality, the second component entering its stochastic drift. The dynamics of Y are described by the following SDE:

$$\begin{cases} d\mu_t = k_1(\bar{\mu}_t - \mu_t)dt + \sigma_1(t)\sqrt{\mu_t}\, dW_t^{(1)} \\ \\ d\bar{\mu}_t = k_2(m(t) - \bar{\mu}_t)dt + \sigma_2(t)\sqrt{\bar{\mu}_t - m^*(t)}\, dW_t^{(2)} \end{cases} \tag{28}$$

where: $W = (W^{(1)}, W^{(2)})$ is a standard brownian motion in $\mathbb{R}^2$; $k_1, k_2 > 0$ are parameters representing the 'speed of mean reversion' of μ to $\bar{\mu}$ and of $\bar{\mu}$ to m; σ_1, σ_2, m and m^* are bounded continuous nonnegative functions. The function m may for example represent an intensity of mortality derived from an available life table (from which we want to project mortality improvements) or a best-estimate assumption. The function m^* may instead represent an optimistic enough level of mortality bounding the intensity $\bar{\mu}$ from below. The process $\bar{\mu}$ is well-defined (i.e. $\bar{\mu} > m^*$ a.s.) provided $m(t) > m^*(t) + \sigma_2(t)^2/2k_2$ for all $t \geq 0$. By taking $m^*(t) = \sigma_1(t)^2/2k_1$ for all $t \geq 0$ we ensure that $\bar{\mu}_t > \sigma_1(t)^2/2k_1$ a.s. and thus that Y is well-defined: see Duffie and Kan (1996, p. 387). We note that the model takes into account the risk of random fluctuations around μ and around the drift target m. This is important when dealing with the risk of deviations from expected mortality levels, as explained in Sec. 1 and Sec. 2. See Biffis (2004) for more details and some numerical examples relative to pensionable ages.

A second interpretative example of affine intensity of mortality is provided by the specification:

$$\mu_t = m(t) + \Lambda(t, Y_{t-}), \tag{29}$$

where m is a real-valued bounded continuous function on $[0, \infty)$ and where $\Lambda(t, x) = \eta_1(t) \cdot x$ (η_1 being bounded continuous and $\mathbb{R}^d$-valued) is specified so as to ensure that μ is non-negative. Note that μ is of the affine form (in Y_-) with time-dependent 'intersect' m. The deterministic component m represents a suitable assumption about the intensity of mortality, as discussed above. The random component Λ represents in turn departures from the initially chosen basis, capturing random fluctuations as well as systematic deviations, depending on the specification of the drift, volatility and jump parameters of the affine jump-diffusion. In particular, the extent to which m is 'binding' for the evolution of the random intensity μ will depend on the degree of mean reversion (if any) presented by the diffusive element of Y and on the frequency and amplitude of any jumps occurring. In Biffis (2003, 2004) some numerical examples concerning a one-dimensional Poisson-Gaussian process are provided.

The intensity specification (29) presents some simple features that can be helpful in a number of respects. For example, we can see that survival probabilities are expressed as multiplicative adjustments to the survival probabilities derived from m, the chosen demographic basis. Since $\Lambda(t,x)$ is affine, on the set $\{\tau > t\}$ we have:

$$\begin{aligned}\mathbb{P}(\tau > T|\mathcal{F}_t) &= e^{-\int_t^T m(s)ds}\ E\left[e^{-\int_t^T \Lambda(s,Y_s)ds}\middle|\mathcal{F}_t\right] \\ &= {}_{T-t}p_{x+t}\ e^{\alpha(t)+\beta(t)\cdot Y_t}\end{aligned} \tag{30}$$

where ${}_{T-t}p_{x+t}$ is the survival probability implied by the demographic basis m, while the functions $\alpha(\cdot)$ and $\beta(\cdot)$ solve the ODEs (9) and (10) (associated with the transform of Λ appearing in (30)) with boundary conditions $\alpha(T) = 0$ and $\beta(T) = 0$. We see therefore that, as is customary in actuarial practice, the computation of survival probabilities can be based on adjustments to a reference mortality table. In our framework, however, adjustment factors derive from a stochastic model for the evolution of mortality. The adjustment mechanism provided by (30) is not too different from the UK projection model based on 'reduction factors' for death probabilities (e.g. CMI (1999)). Here, we would interpret the adjustments as 'increase factors' for survival probabilities at old ages. Expression (30) can be very useful in the context of market valuations as well, when the underlying probability measure is risk-neutral and one calibrates the fair-value of insurance contracts to proxies for secondary market prices (e.g. assets backing a book of contracts minus the value of in-force business, provided prudential reserves are considered and no allowance for cost of capital or deferred tax liabilities is made; see Abbink and Saker (2002), Biffis (2003, 2004)).

6. Valuation Results

We are now ready to introduce the insurance market in which the valuation of GAOs takes place. We take a probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting a process X in $\mathbb{R}^k$, representing the evolution of financial variables, and a process Y in $\mathbb{R}^d$, representing the evolution of mortality. We assume quite naturally that X and Y are independent under any probability measure of concern. Moreover, we focus on a policyholder aged x at time 0, with random residual lifetime described as in Sec. 5 by an $\mathbb{F}$-stopping time τ. The flow of information available at time $t \geq 0$ is assumed to be represented by the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ including knowledge of the evolution of all state variables up to time t and of whether the policyholder has died by time t. Formally, we set $\mathbb{F} = \mathbb{G} \vee \mathbb{H}$, where $\mathbb{G} = \mathbb{G}^X \vee \mathbb{G}^Y$ and $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$, with $\mathcal{H}_t = \sigma(\mathbb{I}_{\{\tau\leq s\}}, s \leq t)$. Thus $\mathbb{H}$ is the minimal filtration with respect to which τ is a stopping time.

In the absence of arbitrage, an equivalent martingale measure $\mathbb{Q}$ exists, under which all financial security prices are martingales after deflation by the money market account (see Sec. 4). Under $\mathbb{Q}$, the $\mathbb{F}$-stopping time τ is assumed to be doubly stochastic driven by $\mathbb{G}^Y \subset \mathbb{F}$ with random intensity μ. It is worth emphasizing that the doubly stochastic property does not need to preserve under measure changes. However, an intensity is guaranteed to exist under $\mathbb{Q}$ if τ admits an intensity under, say, the 'physical' measure $\mathbb{P}$ (Artzner and Delbaen (1995)). In the setup specified, the standard risk-neutral machinery can be used, as shown for example in Lando (1998) and Duffie, Schröder and Skiadas (1996), provided we use a fictitious short rate process $r+\mu$ accounting for the risk of mortality. This will become clearer below, when a mortality-risk-adjusted money-market account will show up.

The financial market described in Sec. 4 may not be complete, that is, some contingent claims may not be spanned by the securities available in the market. Even if we assume it to be so, the 'insurance market' considered is not, unless *ad hoc* assumptions are made. For internal or fair (in the sense proposed by the IASB: see Sec. 2) valuation purposes, we will refer to secondary markets where (re)insurers can exchange books of policies, so that both long and short positions can be taken on insurance contracts. Depending on the type of contracts under valuation, suitable basic insurance contracts will be assumed to be continuously traded in the market and represent the primitive securities used for arbitrage pricing. For example, when valuing annuities, pure endowments of (possibly) every maturity will be implicitly taken as primitive securities. Prices obtained in this context will be called *fair values*, consistently with the terminology proposed by the IASB. When valuations are aimed at reserving or (primary market) pricing, the situation is more delicate, since one typically makes reference to a market where insurers take short positions in insurance contracts, while insureds take only long positions. The trading constraints on the insurer side can be weakened by assuming that unlimited reinsurance is available, corresponding to a long position on the contracts sold.

However, a delicate caveat must be kept in mind whatever the market (primary or secondary) of reference is. Namely, each single policy refers to a specific policyholder, so that arbitrage pricing results referring to single contracts can only approximately be scaled back to portfolio levels. Indeed, standard arguments involving perfect hedging and replicating strategies only apply to policies considered in their own. Extensions of results to policy portfolios must be meant as approximations, their precision level depending on the degree of homogeneity (from the risk selection point of view) of the policyholders considered and on the dimension of the portfolio in case of pooling risks (e.g. the risk of mortality random fluctuations around expected values).

6.1. **GAOs Results.** In this section we provide the valuation results relative to a single premium unit-linked pure endowment contract embedding a GAO. The inclusion of a benefit payable in case of death during the deferment period and of indexed annuity amounts is discussed in Sec. 6.2. We start by assuming that the value of the benefits is linked to the market price of a reference fund during the deferment period. On the lines of Sec. 2 and Sec. 4, we let S and $(a_t) \doteq (\ddot{a}_{x+T}(t))$ denote the $\mathbb{F}$-adapted processes respectively of the reference fund price and of the time-t fair value of the annuity payable from time $T \geq 0$ to an individual then aged $x+T$. Such annuity pays unitary amounts (conditional on survival) at times $T_0 \leq T_1 \leq \ldots \leq T_h \leq \ldots$, with $T_0 \geq T$. We assume that under $\mathbb{Q}$ the state variables processes X and Y are affine with respect to $\mathbb{F}$ and set $Z \doteq (X,Y)$. (The process Z in $\mathbb{R}^{k+d}$ is also affine.) We then denote by r the riskless short rate, by B the money-market account price process and assume that r and B are defined as in Sec. 4 and satisfy the conditions stated therein. Similarly, we assume that the reference fund price process is exponential affine. In particular, we let $S_t = e^{X_t^i}$ and $r_t = r(t, X_t) = p_0(t) + p_1(t) \cdot X_t$. The intensity of mortality is assumed to be affine as in the framework described in Sec. 5, that is $\mu_t = \mu(t, Y_{t-}) = \eta_0(t) + \eta_1(t) \cdot Y_{t-}$.

Consistently with the fair valuation framework described above, the time-t fair value a_t is given by:

$$a_t = B_t \sum_{h=0}^{\infty} E^{\mathbb{Q}}\left[\mathbb{I}_{\{\tau > T_h\}} B_{T_h}^{-1} | \mathcal{F}_t\right], \tag{31}$$

where the sum above may be stopped at $h = h^* - 1$, provided $x + T_{h^*}$ is the minimum age to which no one is assumed to survive. The following proposition provides the time-t fair value expression for the payoff V_T described by (1), in which we make reference to the third line.

Proposition 6.1. *For fixed $T \geq t \geq 0$, the time-t fair value V_t of the payoff V_T described by expression* (1)*, on the event $\{\tau > t\}$, is given by:*

$$\begin{aligned} V_t =& e^{\alpha(t;\epsilon(i),T)+\beta(t;\epsilon(i),T)\cdot Z_t} + g \sum_{h=0}^{\infty} e^{\alpha(T;0,T_h)+\alpha(t;\gamma_h,T)+\beta(t;\gamma_h,T)\cdot Z_t} \mathbb{Q}_T^{(\gamma_h)}(A|\mathcal{F}_t) \\ & - e^{\alpha(t;\epsilon(i),T)+\beta(t;\epsilon(i),T)\cdot Z_t} \mathbb{Q}_T^{(\epsilon(i))}(A|\mathcal{F}_t) \end{aligned} \tag{32}$$

where: $\epsilon(i)$ denotes the vector in $\mathbb{R}^{k+d}$ with all null components except the i-th, equal to 1; γ_h is the vector in $\mathbb{R}^{k+d}$ defined by $\gamma_h \doteq \epsilon(i) + \beta(T;0,T_h)$; $\mathbb{Q}_T^{(v)}$ is the probability defined by means of (15) *for $v \in \mathbb{R}^{k+d}$; the functions $\alpha(\cdot;b,u)$ and $\beta(\cdot;b,u)$ solve uniquely the ODEs* (9) *and* (10)*, relative to the affine function $R(s,z) = r(s,x) + \mu(s,y)$ (where $z = (x,y) \in \mathbb{R}^{k+d}$), with boundary conditions $\alpha(u;b,u) = 0$ and $\beta(u;b,u) = b$ (for $b \in \mathbb{R}^{k+d}$ and $u \geq 0$).*

Proof. First, we note that, by using the tower property, the doubly stochastic assumption and the transform formula (8), the time-t value of the survival benefit is given by:

$$\begin{aligned} B_t E^{\mathbb{Q}}\left[B_T^{-1}\mathbb{I}_{\{\tau > T\}} S_T \middle| \mathcal{F}_t\right] &= E^{\mathbb{Q}}\left[e^{-\int_t^T (r(s,X_s)+\mu(s,Y_s))ds} e^{\epsilon(i)\cdot Z_T} \middle| \mathcal{F}_t\right] \\ &= e^{\alpha(t;\epsilon(i),T)+\beta(t;\epsilon(i),T)\cdot Z_t}. \end{aligned}$$

Similar arguments applied to (31) yield immediately that the following holds:

$$a_t = \sum_{h=0}^{\infty} e^{\alpha(t;0,T_h)+\beta(t;0,T_h)\cdot Z_t} \tag{33}$$

This can be used into the (digital) option component of (1) to get the following expression for its time-t fair value, on the event $\{\tau > t\}$:

$$\begin{aligned} & gB_t E^{\mathbb{Q}}\left[\mathbb{I}_A B_T^{-1} S_T (a_T - 1/g) \mathbb{I}_{\{\tau > T\}} \middle| \mathcal{F}_t\right] = \\ & = g \sum_{h=0}^{\infty} e^{\alpha(T;0,T_h)} E^{\mathbb{Q}}\left[e^{-\int_t^{T_h} (r(s,X_s)+\mu(s,Y_s))ds} e^{\gamma_h \cdot Z_T} \mathbb{I}_A \middle| \mathcal{F}_t\right] \\ & \quad - E^{\mathbb{Q}}\left[e^{-\int_t^T (r(s,X_s)+\mu(s,Y_s))ds} e^{\epsilon(i)\cdot Z_T} \mathbb{I}_A \middle| \mathcal{F}_t\right], \end{aligned} \tag{34}$$

Finally, repeated application of the change of measure described at the end of Sec. 3 completes the proof. ∎

The result provided by Prop. 6.1 can be employed in several ways. First, it can be used directly, provided the insurer is able or willing to specify the exercise probabilities $\mathbb{Q}_T^{(\cdot)}$, possibly relying on available internal/market data regarding historical rates of exercise. However, we must stress that the exercise probabilities appearing in (32) are 'generalized forward risk-adjusted' probabilities and thus their specification/estimation is quite difficult, although the dynamics of Z under $\mathbb{Q}_T^{(\cdot)}$ are easily recovered by formulae (16) to (18). In order to overcome such problem, we offer some other approaches.

We can for example specify the dynamics of the (primary market) annuity price $\tilde{a}$ (alternatively, of the conversion rate $1/\tilde{a}$) in terms of the affine state process Z. We keep on reasoning under $\mathbb{Q}$, so that the dynamics of $\widetilde{a}$ should be adjusted for the risk of adverse (to the insurer) exercise. We assume that $\tilde{a}$ is specified as $\tilde{a} = f(b \cdot Z)$ for some $b \in \mathbb{R}^{k+d}$ and some positive strictly increasing function f on $\mathbb{R}$ (for example, $\tilde{a} = k + \exp(b \cdot Z)$). Both f and b can allow for the correlation of $\tilde{a}$ with a, interest rates and mortality, as well as with any additional factors of concern (charges, insurance market trends), provided they are included in the state variables. The exercise probabilities appearing in (32) are then expressed as follows:

$$\begin{aligned} \mathbb{Q}_T^{(v)}(A|\mathcal{F}_t) &= \mathbb{Q}_T^{(v)}(-b \cdot Z_T \le f^{-1}(1/g)|\mathcal{F}_t) \\ &= G^{(v,T)}_{0,0,1,-b,(0,0)}(f^{-1}(1/g), Z_t, T) \\ &= G_{v,0,1,-b,(p,\eta)}(f^{-1}(1/g), Z_t, T)\ e^{-\alpha(t;v,T)-\beta(t;v,T)\cdot Z_t}, \end{aligned} \tag{35}$$

where $G_{\cdot}^{(v,T)}$ and $G_{\cdot}$ are functions defined by (13), respectively under $\mathbb{Q}_T^{(v)}$ and $\mathbb{Q}$, and can be computed as in (14).

In case no distinction is made between primary and secondary market, as exemplified by (2), we can consider $A = \{a_T \ge 1/g\}$ and the valuation of GAOs is similar to the pricing of vulnerable options on defaultable coupon-bonds (consols) with zero recovery. The exercise boundary is in this case a concave surface and the transform inversion methods examined in Sec. 3 cannot be used directly. However, an approximating hyperplane can be found, as explained in Munk (1999) and Singleton and Umantsev (2002), and the usual formula (14) used. The particular case in which the underlying of the option is a pure endowment instead of an annuity poses no problem, since we are then back to a linear exercise boundary.

6.2. **Death Benefits and Indexed Annuities.** We now briefly discuss the inclusion of a death benefit process C payable in case death occurs before maturity. We let C be a bounded $\mathbb{G}$-adapted $\mathbb{F}$-predictable process. Intuitively, that means the death benefit is known an instant before death happens. The same arguments used in previous section yield that the time-t fair value V_t^d of the amount $C_\tau \mathbb{I}_{\{t<\tau\le T\}}$, for $T \ge t \ge 0$, is given by:

$$\begin{aligned} V_t^d &= E^{\mathbb{Q}}\left[e^{-\int_t^\tau r(s,X_s)ds} C_\tau \mathbb{I}_{\{t<\tau\le T\}} \middle| \mathcal{F}_t \right] \\ &= E^{\mathbb{Q}}\left[E^{\mathbb{Q}}\left[e^{-\int_t^\tau r(s,X_s)ds} \mathbb{I}_{\{t<\tau\le T\}} C_\tau \middle| \mathcal{F}_t \vee \mathcal{G}_T \right] \middle| \mathcal{F}_t \right] \\ &= \mathbb{I}_{\{\tau>t\}} \int_t^T E^{\mathbb{Q}}\left[e^{-\int_t^s (r(u,X_u)+\mu(u,Y_u))du} \mu(s,Y_s) C_s \middle| \mathcal{F}_t \right] ds, \end{aligned} \tag{36}$$

where we have used the fact that on the set $\{\tau > t\}$ the $\mathcal{F}_t \vee \mathcal{G}_T$-conditional density of τ is given by $\mu_s \exp\left(-\int_t^s \mu_u du\right)$ for $t \le s \le T$. We can let $C = k\ \exp(b \cdot Z)$ for $b \in \mathbb{R}^{k+d}$ and $k > 0$. (In particular, the death benefit could be linked to the reference fund by setting $b = \epsilon(i)$.) Then (36) can be developed further to obtain:

$$V_t^d = \int_t^T E^{\mathbb{Q}} \left[e^{-\int_t^s (r(u,X_u)+\mu(u,Y_u))du} \left(\eta_0(s) + \eta_1(s) \cdot Y_s\right) k e^{b \cdot Z_s} \middle| \mathcal{F}_t \right] ds$$
$$= k \int_t^T e^{\alpha(t;b,s)+\beta(t;b,s)\cdot Z_t} \left(\hat{\alpha}(t; b, \hat{\eta}_1(s), \eta_0(s), s) + \hat{\beta}(t; b, \hat{\eta}_1(s), s) \cdot Z_t \right) ds$$

on the set $\{\tau > t\}$, where α, β, $\hat{\alpha}$ and $\hat{\beta}$ solve, for fixed $s \in [t, T]$, the ODEs (9) to (12) with boundary conditions $\alpha(s) = 0$, $\beta(s) = b$, $\hat{\alpha}(s) = \eta_0(s)$ and $\hat{\beta}(s) = \hat{\eta}_1(s)$, where we set $\hat{\eta}_1(s) = (0, \eta_1(s))$.

The framework proposed can easily handle indexed annuities. Specifically, suppose that the annuity underlying the GAO provides an annuity amount $w_0(u) + w_1(u) \cdot Z_u$ at the generic payment date $u \geq T$. For example, the indexation could relate to the only state variables process X and be defined through the policy $w_0(u) = 1 + s\, p_0(u)$, $w_1(u) = s\, (p_1(u), 0)$, for some $s > 0$, yielding an annuity amount of the form $1 + s\, r_u$. Similar reasoning can be used to link the benefits to a reference fund price or to index them on other quantities of concern (e.g. inflation), provided the latter are included in the state variables. The results of Prop. 6.1 can be adjusted to take indexation into account by replacing expressions (31) and (33) with:

$$a_t = B_t \sum_{h=0}^{\infty} E^{\mathbb{Q}} \left[\mathbb{I}_{\{\tau > T_h\}} B_{T_h}^{-1} \left(w_0(T_h) + w_1(T_h) \cdot Z_{T_h} \right) | \mathcal{F}_t \right]$$
$$= \sum_{h=0}^{\infty} e^{\alpha(t;0,T_h)+\beta(t;0,T_h)\cdot Z_t} \left(\hat{\alpha}(t; 0, w_1(T_h), w_0(T_h), T_h) + \hat{\beta}(t; 0, w_1(T_h), T_h) \cdot Z_t \right)$$

where, for all h, α, β, $\hat{\alpha}$ and $\hat{\beta}$ solve the ODEs (9) to (12) with boundary conditions $\alpha(T_h) = 0$, $\beta(T_h) = a$, $\hat{\alpha}(T_h) = w_0(T_h)$ and $\hat{\beta}(T_h) = w_1(T_h)$. The expectations appearing in (34) would then be solved by using transform inversion methods, on the lines of expression (14).

7. Conclusion

In this work we have demonstrated how the jointly affine and doubly stochastic setup enables to deal effectively with several sources of risk that simultaneously affect insurance contracts embedding GAOs. In particular, the risk of mortality (systematic and unsystematic) has been taken into account, by exploiting a parallel with the pricing of credit-risky securities in intensity-based models. Moreover, we have explicitly allowed for the mismatch between the insurer's viewpoint in valuing the cash flows triggered by the exercise of such options and the policyholder's exercise decision. Special emphasis has been put on the fair valuation framework proposed by the IASB and aimed at market-oriented accounting principles for insurance liabilities.

References

Abbink, M. and M. Saker (2002). Getting to grips with fair value. Staple Inn Actuarial Society, London.

Artzner, P. and F. Delbaen (1995). Default risk insurance and incomplete markets. *Mathematical Finance*, vol. 5(3):187–195.

Bakshi, G. and D. Madan (2000). Spanning and derivative security valuation. *Journal of Financial Economics*, (55):205–238.

Bakshi, G., C. Cao and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, (52):2003–2049.

Ballotta, L. and S. Haberman (2003a). The fair valuation problem of guaranteed annuity options: the stochastic mortality environment case. Working paper n. 145, Cass Business School.

Ballotta, L. and S. Haberman (2003b). Valuation of guaranteed annuity conversion options. *Insurance: Mathematics & Economics*, vol. 33(1):87–108.

Bates, D. (1997). Post-87' crash fears in S&P500 futures options. *Journal of Econometrics*, (94):181–283.

Biffis, E. (2003). Affine processes in mortality modelling: an actuarial application. Quaderno n. 3/2003, Department of Applied Mathematics, University of Trieste.

Biffis, E. (2004). Stochastic models for dynamic mortality: Actuarial applications. PhD thesis, Department of Applied Mathematics, University of Trieste.

Bowers, N., H. Gerber, J. Hickman, D. Jones and C. Nesbitt (1997). *Actuarial Mathematics*. The Society of Actuaries, Schaumburg, Illinois, second edn.

Boyle, P. and M. Hardy (2003). Guaranteed annuity options. *ASTIN Bulletin*, vol. 33(2):125–152.

Brémaud, P. (1981). *Point Processes and Queues - Martingale Dynamics*. Springer-Verlag, New York.

CMI (1999). *Continuous Mortality Investigation Committee*. The Institute of Actuaries and The Faculty of Actuaries, London. Report n. 17.

Cox, J., J. Ingersoll and S. Ross (1985). A theory of the term structure of interest rates. *Econometrica*, vol. 53:385–408.

Dahl, M. (2003). Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts. Working paper n.185, Laboratory of Actuarial Mathematics, University of Copenhagen.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, (463):463–520.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, third edn.

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance*, vol. 6(4):379–406.

Duffie, D., M. Schröder and C. Skiadas (1996). Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *The Annals of Applied Probability*, vol. 6(4):1075–1090.

Duffie, D., J. Pan and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, vol. 68(6):1343–1376.

Duffie, D., D. Filipovič and W. Schachermayer (2003). Affine processes and applications in finance. *Annals of Applied Probability*, vol. 13:984–1053.

Filipovič, D. (2001). Time-inhomogeneous affine processes. Forthcoming in *Stochastic Processes and Their Applications*.

Grandell, J. (1976). *Doubly Stochastic Poisson Processes*. Springer-Verlag, New York.

Harrison, M. and D. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, (11):418–430.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, (6):327–343.

IASB (2001). *Draft Statement of Principles*. International Accounting Standards Board. Available at: www.iasb.org.uk.

Jacod, J. and A. Shiryaev (2003). *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, second edn.

Keyfitz, N. (1982). Choice of function for mortality analysis: Effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology*, vol. 21:329–352.

Keyfitz, N. (1985). *Applied Mathematical Demography*. Springer-Verlag, New York, second edn.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research*, (2):99–120.

Milevsky, M. and S. Promislow (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics & Economics*, vol. 29(3):299–318.

Munk, C. (1999). Stochastic duration and fast coupon bond option pricing in multi-factor models. *Review of Derivatives Research*, (3):157–181.

Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics & Economics*, vol. 29:231–245.

Olivieri, A. and E. Pitacco (2003). Annuitisation guarantee and uncertainty in mortality trends. Working paper n. 30, CERAP - Università Bocconi.

Pelsser, A. (2003). Pricing and hedging guaranteed annuity options via static option replication. Working paper, Nationale-Nederlanden and Erasmus University, Rotterdam.

Piazzesi, M. (2003). Affine term structure models. Preprint: The Handbook of Financial Econometrics.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer-Verlag, New York.

Scott, L. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Application of Fourier inversion methods. *Mathematical Finance*, (7):345–358.

Singleton, K. and L. Umantsev (2002). Pricing coupon-bond options and swaptions in affine term structure models. Working paper, Graduate School of Business, Stanford University.

Van Bezooyen, J., C. Exley and S. Mehta (1998). Valuing and hedging guaranteed annuity options. Working paper, GEMS. Presented to the Joint Investment and Faculty Investment Conference, Cambridge, 14th Sep 1998.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, vol. 5(2):177–188.

Wilkie, A., H. Waters and S. Yang (2003). Reserving, pricing and hedging for policies with guaranteed annuity options. *British Actuarial Journal*, vol. 9(2):263–391.