AFFINE PROCESSES FOR DYNAMIC MORTALITY AND ACTUARIAL VALUATIONS

ENRICO BIFFIS*

Istituto di Metodi Quantitativi, Università Bocconi

Abstract. We address the risk analysis and market valuation of life insurance contracts in a jump-diffusion setup. We exploit the analytical tractability of affine processes to deal simultaneously with financial and demographic risks affecting a wide range of insurance covers. We then focus on mortality at pensionable ages and show how the risk of longevity can be taken into account. A parallel with the pricing of certain credit risky securities is drawn, in order to employ important results derived in that field.

Keywords: affine jump-diffusion, stochastic mortality, doubly stochastic processes, longevity risk, fair value.

JEL: G13, G22, J11, M40.
Subj. Class.: IM40, IB10, IB13.

* Istituto di Metodi Quantitativi, Università Bocconi - Viale Isonzo 25, 20135 Milano, Italy. Tel.: +390258365556. Fax: +390258365630. E-mail: ebiffis@unibocconi.it.

First version: 1st June 2004. This version: 4th October 2004. This paper revises and extends Biffis (2003). Earlier versions have been presented at the 8^{th} IME Congress in Lyon, at the 6^{th} Spanish-Italian Meeting on Financial Mathematics in Trieste, at the University of Parma and at the ETH in Zürich. I would like to thank the participants in those conferences and seminars for helpful comments and suggestions. All errors are mine. I acknowledge financial support from the Italian MIUR.

1. Introduction

Actuarial valuations involve the consideration of at least two types of uncertainty in life insurance: uncertainty related to financial factors (financial risk) and uncertainty related to the random residual lifetime of insureds (mortality risk). Mortality is traditionally modeled deterministically, either assuming suitably parameterized analytical models, either adopting adjusted/projected mortality tables. Recent mortality trends have proved particularly challenging for the pricing and reserving related to contracts providing long-term living benefits (e.g. Olivieri, 2001; Pitacco, 2003a) and have called for more effective approaches.

Valuation models cannot fully capture the dynamics of insurance business unless both risks are modeled adequately. In this regard, a new impulse to a sound and articulated valuation of insurance contracts has been recently prompted by the release of IASB (2001, 2004) by the International Accounting Standards Board (IASB) and by the support shown by several institutions to the principles there proposed. The IASB's proposals acknowledge the need to deal explicitly with all sources of risk, including non-financial risks, both diversifiable (e.g. mortality random fluctuations around expected levels) and non fully diversifiable (e.g. systematic departures of mortality from expected levels), and advocate the accounting of insurance liabilities at market value. Specifically, the IASB defines *fair value* of a book of insurance liabilities its exchange price in a (hypothetical) secondary market transaction. Since a deep wholesale market for such liabilities does not exist, exchange prices are not easily observable. However, 'no-arbitrage type' arguments imply that the fair value of an insurance liability should not be too different from the market value of a portfolio of traded assets matching broadly the liability cashflows. Indeed, the IASB favours the use of an expected discounted cashflow approach consistent with risk-neutral valuation.

Any valuation model confronts actuaries and insurance practitioners with a trade-off between complexity and computational tractability of pricing and estimation. In this respect, affine jump-diffusions are stochastic processes that have shown themselves to be useful in developing tractable and flexible dynamic asset pricing models. They have been extensively used in modeling the term structure of interest rates (see Dai and Singleton, 2000, and references therein), stochastic volatility for currency and equity prices (e.g. Bakshi, Cao and Chen, 1997) and the risk of default of corporate bonds and other credit-risky securities (e.g. Lando, 1998).

In this work, we model asset prices and mortality dynamics by using affine jump-diffusions. In this way, we are able to fully exploit the analytical tractability of affine processes in the context of both financial and mortality risk. Moreover, by drawing a parallel between insurance contracts and certain credit-sensitive securities (first suggested in Artzner and Delbaen, 1995, and recently used by Milevsky and Promislow, 2001; see also Dahl, 2004) we exploit some results of the so called *intensity-based* approach to credit risk modelling to derive closed-form expressions for the pricing and reserving of the most common life insurance contracts. We also outline a link with a popular actuarial valuation methodology, the Embedded Value Method, and provide valuation formulae whose implementation is straightforward and consistent with the framework proposed in IASB (2001, 2004).

The paper is organized as follows. In Sec. 2, we draw a parallel between insurance contracts and some defaultable securities. In Sec. 3, we sketch the main characteristics of affine processes and present the most useful results for financial applications. Sec. 4 describes the financial and insurance market of concern, presenting the framework to be employed for risk-analysis and market valuations or just for stand-alone mortality modeling. Sec. 5 offers valuation expressions for a large class of (possibly unit-linked or indexed) life insurance contracts where mortality, the term structure of interest rates and asset prices are all modeled stochastically. In Sec. 6, we show how affine jump-diffusions can be used to describe the dynamics of human mortality and to deal with the so called 'longevity risk', providing in turn some numerical examples. Sec. 7 offers concluding remarks. Additional results and details are provided in the Appendix.

2. Insurance Contracts and Credit-risky Securities

There are two main approaches to valuing credit-sensitive securities (see Jeanblanc and Rutkowski, 2000, for a detailed overview): the *structural approach*, which aims at determining the default event by modeling the dynamics of the firm's assets, and the *reduced-form approach*, under which the default time is modeled as a stopping time that occurs as a total surprise. The latter approach relies on the *exogenous* specification of the conditional probability of default, given that default has not yet occurred. This is usually done by modeling the intensity of default, which is allowed to depend on observables or unobservables variables. Among the first papers to advocate the use of such approach was Artzner and Delbaen (1995), where a parallel with the pricing of annuities was drawn. It seems thus logical to go back to that insight and employ it for actuarial purposes.

A key result of intensity-based modeling is that the pricing of credit-sensitive securities can be carried out in the same way as for risk-free financial instruments, provided discounting (under a 'risk-neutral' probability) is performed by using an *adjusted* (or fictitious) short rate accounting for the probability and timing of default. This means that all of the arsenal of term-structure models can be applied by suitably parameterizing the adjusted short rate instead of the risk-free rate. In particular, let R be the adjusted short rate process and $B^d(0,T)$ the price of a defaultable zero coupon bond with maturity T and face value 1. Then, from results in (e.g.) Lando (1998), under technical conditions, we have:

$$B^d(0,T) = E\left[e^{-\int_0^T R_s \mathrm{d}s}\right] + \int_0^T w_s q_s \mathrm{d}s, \tag{1}$$

where the first term on the right hand side is the current price of a claim paying one unit at maturity in case of no default, zero otherwise, while the second term involves w_s, representing the proceeds from default (recovery) if it occurs at time s, and q_s, the price density of a claim paying one unit if default occurs in $(s, s+\mathrm{d}s)$. See Sec. 5.1 for a detailed analysis of the result.

The form of expression (1) is remarkably useful from the point of view of interpretation and extremely appealing from that of computation, as will be shown in Sec. 3, where the analytical tractability of the affine setting is described. If we interpret default as the event 'insured's death' and the intensity of default as the intensity of mortality, we can see that expression (1) covers quite a large class of traditional insurance contracts. For example, endowments of arbitrary maturity, as well as of different benefit amounts, are encompassed by (1). Moreover, annuities are covered by combinations[1] of terms like the expectation appearing in (1) for different maturities (annuity payment dates), while assurances are covered by the only second term in (1). Of course, the interpretation of the default event is reverted: for example, 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) has defaulted. We point out that we will not treat the issue of a possible default by the insurer.

Viewing an endowment as a defaultable zero-coupon bond, or an annuity as a defaultable consol (with zero recovery) yields several advantages. For example, it is easier to take into account the joint effect of mortality and financial risks, as will be shown in Sec. 5. This is particularly important for those products featuring embedded options depending on both sources of risk (see Biffis and Millossovich, 2004, for an application to guaranteed annuity options). Furthermore, adjusted discount rates, which are crucial to any profit testing or embedded value computation, can be given a sound theoretical background for the mortality risk component. The framework proposed can therefore be beneficial in many respects, from pricing and reserving to preparing financial statements consistent with the IASB's accounting standards.

3. Affine Processes

Affine processes are a class of Markov processes with conditional characteristic function of the exponential affine form. A thorough treatment of such processes is provided in Duffie, Filipovič and Schachermayer (2003) and Filipovič (2001). In this section, we adopt the narrower but more usual (in financial applications) perspective based on the definition of affine processes in terms of strong solutions to specific stochastic differential equations (SDEs) in a given filtered probability space (e.g. Duffie and Kan, 1996; Duffie, Pan and Singleton, 2000).

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions[2] and representing the information available up to time t. An $\mathbb{R}^n$-valued *affine jump-diffusion* X is an $\mathbb{F}$-Markov process specified as the strong solution to the following SDE:

$$dX_t = \delta(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t + \mathrm{d}J_t, \tag{2}$$

where W is an $\mathbb{F}$-standard Brownian motion in $\mathbb{R}^n$ and J is a pure-jump process in $\mathbb{R}^n$ with jump-arrival intensity $\{\kappa(t, X_t) : t \geq 0\}$ and time-dependent jump distribution ν_t on $\mathbb{R}^n$. We require the drift δ, the instantaneous covariance matrix $\sigma\sigma^{\mathsf{T}}$ and the jump-arrival intensity κ to have all affine dependence on X, as explained in App. A.

An important result holds for analytical approaches based on the affine structure described. For any $c \in \mathbb{C}$ and $a, b \in \mathbb{C}^n$, for given $T \geq t$ and affine function Λ

[1] We can always see an annuity as a set of pure endowments, just as coupon bearing bonds can be stripped.

[2] That is, $(\mathbb{P}, \mathcal{F})$-completeness and right continuity: see Protter (2004, p. 3). We refer the reader to the same source for background and terminology. In what follows, all filtrations are assumed to satisfy the usual conditions.

defined by $\Lambda(t,x) = \lambda_0(t) + \lambda_1(t) \cdot x$ (for some bounded continuous $\mathbb{R} \times \mathbb{R}^n$-valued function $\lambda \doteq (\lambda_0, \lambda_1)$), under technical conditions the following expression holds:

$$E\left[e^{-\int_t^T \Lambda(s,X_s)\mathrm{d}s}\ e^{a \cdot X_T}\ (b \cdot X_T + c)\Big|\mathcal{F}_t\right] = e^{\alpha(t)+\beta(t)\cdot X_t}\ \left[\hat{\alpha}(t) + \hat{\beta}(t) \cdot X_t\right], \quad (3)$$

where $\alpha(\cdot) \doteq \alpha(\cdot; a, T)$, $\beta(\cdot) \doteq \beta(\cdot; a, T)$, $\hat{\alpha}(\cdot) \doteq \hat{\alpha}(\cdot; a, b, c, T)$ and $\hat{\beta}(\cdot) \doteq \hat{\beta}(\cdot; a, b, T)$ are functions solving uniquely a set of ordinary differential equations (ODEs) specified in App. A, with boundary conditions $\alpha(T) = 0$, $\beta(T) = a$, $\hat{\alpha}(T) = c$ and $\hat{\beta}(T) = b$. One can immediately see the importance of these results in deriving analytical solutions for a variety of financial applications. For example, we will be interested in the following:

First, the entire class of affine term structure models is obtained as a special case of (3) when we take expectation under a risk-neutral probability measure and set $r_s \doteq \Lambda(s, X_s)$ (r_s being the short rate process), $a, b = 0$ and $c = 1$. This class includes: the model proposed by Vasicek (1977), where X is a Gaussian Ohrnstein-Uhlenbeck process; the CIR model (Cox, Ingersoll and Ross, 1985), where X is a Feller difffusion; the general multivariate models introduced by Duffie and Kan (1996) and several other extensions (see Dai and Singleton, 2000, and references therein).

Second, expression (3) can be used in a similar fashion for risk-neutral valuation, when asset prices have an exponential affine form, as is assumed in many financial applications (see Sec. 4.1 and references therein). More generally, the *expected discounted cash flow approach* (with discount rate $\Lambda(s, X_s)$) can be implemented for all discounted cash flows whose form is encompassed by the term under expectation in (3).

Third, the transform (3) can be used in credit risk modeling, when the time of default of a financial counter-party is modeled as the first jump time of a doubly stochastic process N driven by an affine process X. In such framework, the intensity of N has the form $\{\Lambda(s, X_{s-}) : s \geq 0\}$ and (3) can be used, for example, with $a, b = 0$ and $c = 1$ to compute the probability of no default by T, conditional on X_t and survival to t (see Sec. 4.2).

The analytical tractability of affine processes is essentially linked to the ODEs associated with the transform (3). They are generalized Riccati equations that can be solved by using standard numerical methods. For some specifications of the parameters δ, σ, κ and ν, explicit solutions are available. They are derived, for example, in the simple case of the Vasicek and CIR model 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, the choice of the jump distribution can be very general. In the CIR case, 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, for example).

4. Modeling Framework

In this section we introduce the modeling framework. We shall first introduce the financial market (Sec. 4.1) and the mortality model (Sec. 4.2) *separately*, and then successively *combine* them to describe the insurance market model (Sec. 4.3).

4.1. **Financial Market.** We fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We take as given an adapted short-rate process r (such that $\int_0^t |r_s| \mathrm{d}s < \infty$ for all $t \geq 0$) 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 $B_t = \exp(\int_0^t r_s \mathrm{d}s)$ 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.

We assume that at least a security is traded continuously in the market, with a nonnegative semimartingale S representing its ex-dividend price. In the absence of arbitrage, an equivalent martingale measure $\mathbb{Q}$ exists, under which which the gain (from holding the security) process is a martingale after deflation by the money-market account. Specifically, let D be a process of locally integrable variation representing the security's cumulated dividend. Then, the discounted gain process is given by $(B_t^{-1} S_t + \int_0^t B_s^{-1} \mathrm{d}D_s)$ and the following convenient formula applies:

$$S_t = E^{\mathbb{Q}} \left[e^{-\int_t^T r_s \mathrm{d}s} S_T + \int_t^T e^{-\int_t^u r_s \mathrm{d}s} \mathrm{d}D_u \Big| \mathcal{F}_t \right]. \tag{4}$$

In what follows, we will assume that the price of any security is zero after a given time $t > 0$ if the security pays no dividends thereafter. If the security has dividend yield process ζ, i.e. the instantaneous yield from holding the security is $\zeta_t S_t \mathrm{d}t$, then $D_t = \int_0^t \zeta_u S_u \mathrm{d}u$ 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 (4) 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 refer the reader to Duffie (2001) for more details on no-arbitrage pricing.

We then postulate that all security prices are driven by an affine process X in $\mathbb{R}^k$. Specifically, we assume that r is expressed as $r_t \doteq R(t, X_t)$, where X_t satisfies (2) and the conditions there given and R is an affine function defined by $R(t, x) = \rho_0(t) + \rho_1(t) \cdot x$, with $\rho \doteq (\rho_0, \rho_1)$ an $\mathbb{R} \times \mathbb{R}^k$-valued bounded continuous function on $[0, \infty)$. An immediate consequence (use (4) and (3)) is that 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 exponential affine form.

When considering risky securities, we take their log-price to be an affine process. Specifically, let us consider the price process S of a risky security having 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 $[0, \infty)$. Without loss of generality (see Duffie, Pan and Singleton, 2000, Sec. 3.1) we can set $\log(S)$ equal to (say) X^i, the i-th component of X. In the absence of arbitrage, the dynamics of S must obey the restrictions implied by (4) and outlined in App. B. 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 Bakshi, Cao and Chen (1997), Bakshi and Madan

(2000), Bates (2000), Heston (1993) and Scott (1997) are all encompassed by the framework described. See also Eraker, Johannes and Pohlson (2003).

4.2. **Mortality Model.** We fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and focus on an insured aged x at time 0. We model his/her random residual lifetime as an $\mathbb{F}$-stopping time τ_x admitting a random intensity μ_x. Specifically, we regard τ_x as the first jump-time of a nonexplosive $\mathbb{F}$-counting process N recording at each time $t \geq 0$ whether the individual has died ($N_t \neq 0$) or not ($N_t = 0$). The stopping time τ_x is said to admit an intensity μ_x if N does, i.e. if μ_x is a nonnegative predictable process such that $\int_0^t \mu_x(s)\mathrm{d}s < \infty$ for all $t \geq 0$ and such that the compensated process $M = \{N_t - \int_0^t \mu_x(s)\mathrm{d}s : t \geq 0\}$ is a local $\mathbb{F}$-martingale. If, in addition, $E[\int_0^t \mu_x(s)\mathrm{d}s] < \infty$ for all $t \geq 0$, then M is an $\mathbb{F}$-martingale. (Alternatively, we could look at the single jump process $\widetilde{N}_t = \mathbb{I}_{\{\tau_x \leq t\}}$ and take as intensity of $\widetilde{N}$ a predictable process $\widetilde{\mu}_x$ that is equal to μ_x up to time τ_x, the intensity of $\widetilde{N}$ loosing meaning thereafter.) We refer the reader to Brémaud (1981), Jeanblanc and Rutkowski (2000) and Duffie (2001) for details.

The idea behind a stochastic intensity is that, at any time $t \geq 0$ and state $\omega \in \Omega$ such that $\tau_x(\omega) > t$, we have:

$$\mathbb{P}(\tau_x \leq t + \Delta | \mathcal{F}_t)(\omega) \cong \mu_x(t, \omega)\Delta \tag{5}$$

for small $\Delta > 0$. This is the stochastic analogue of the expression for the 'instantaneous death probability':

$${}_{\Delta}q_{x+t} \cong m(x+t)\Delta, \tag{6}$$

which is familiar to actuaries and usually emerges when defining the deterministic intensity m itself (see Bowers, Gerber, Hickman, Jones and Nesbitt, 1997, for actuarial background and notation). Actually, expressions (5) and (6) coincide when $\mathbb{F}$ is the smallest filtration making τ_x a stopping time, i.e. when $\mathcal{F}_t = \sigma(\mathbb{I}_{\{\tau_x \leq s\}}; 0 \leq s \leq t)$ for all $t \geq 0$. We remark, however, that our setup is both stochastic and dynamic, since the distinguishing features of $\mu_t(\omega)$ are: the 'state of the world' $\omega \in \Omega$ determining the particular trajectory of μ; the date $t \geq 0$, i.e. the continuous-time counterpart of the calendar year of reference used in longitudinal tables. Thus, we are naturally adopting a diagonal, or cohort-based, approach (see Pitacco, 2003b).

From an actuarial viewpoint, we may see expression (5) as emphasizing the stochastic variation of the intensity over time, as new information about the insureds'mortality becomes available. This corresponds to a change of demographic basis, represented by the choice/availability of a new intensity (or of a new life table from which the intensity is derived) at each time t. We do not specify what kind of basis, realistic or prudential for example, consistently with the choice of remaining very general in this section about the probability measure $\mathbb{P}$ under which expectations are taken (see Sec. 5.4 for details).

If the insured is *representative* of a group of policyholders (e.g. same age and health status), then the results obtained in the sequel extend to all individuals belonging to the homogeneous population, provided the individual random residual lifetimes can be thought as independent and identically distributed. Keeping this in mind, from now on we drop reference to the age x and set $\tau = \tau_x$, $\mu_t(\omega) = \mu_x(t, \omega)$ and $m(x+t) = m(t)$.

To improve analytical tractability, we further assume that N is a *doubly stochastic* (or Cox) process driven by a subfiltration $\mathbb{G}$ of $\mathbb{F}$, with $\mathbb{G}$-predictable intensity μ. Intuitively, this means that, conditionally on any particular trajectory $t \mapsto \mu_t(\omega)$ of μ, the counting process N becomes Poisson-inhomogeneous with parameter $\int_0^{\cdot} \mu_s(\omega)\mathrm{d}s$. Formally, we have that for all $T \geq t \geq 0$ and nonnegative integer k, the following holds:

$$\mathbb{P}(N_T - N_t = k | \mathcal{F}_t \vee \mathcal{G}_T) = \frac{\left(\int_t^T \mu_s \mathrm{d}s\right)^k}{k!} e^{-\int_t^T \mu_s \mathrm{d}s} \tag{7}$$

The idea behind the specification of $\mathbb{G}$ is that it provides enough information about the evolution of the intensity of mortality, i.e. about the likelihood of death happening, but not enough information about the actual occurrence of death. Such information is carried by the larger filtration $\mathbb{F}$, with respect to which τ is a stopping time.

The application of the law of iterated expectations and the use of (7) with $k = 0$ yield that the 'probability of survival' up to time $T \geq t$, on the set $\{\tau > t\}$, is given by:

$$\mathbb{P}(\tau > T | \mathcal{F}_t) = E\left[e^{-\int_t^T \mu_s \mathrm{d}s} \middle| \mathcal{F}_t \right] \tag{8}$$

This can be compared with its deterministic analogue, which, in usual actuarial notation, reads:

$${}_{T-t}p_{x+t} = e^{-\int_t^T m(s)\mathrm{d}s}, \tag{9}$$

where ${}_{T-t}p_{x+t}$ indicates the probability of surviving $T - t$ years for an individual aged $x + t$ years at time t.

Furthermore, under technical conditions reviewed in Grandell (1976, pp. 105-107), the $\mathcal{F}_t$-conditional density $f_t(\cdot)$ of τ is given, on the set $\{\tau > t\}$, by the expression:

$$f_t(s) = \frac{\partial}{\partial s} \mathbb{P}(\tau \leq s | \mathcal{F}_t) = E\left[\mu_s \, e^{-\int_t^s \mu_u \mathrm{d}u} \middle| \mathcal{F}_t \right] \tag{10}$$

It is clear that both (8) and (10) are particular cases covered by the transform (3), so that working in an affine framework would be very convenient from the computational point of view. We therefore take an $\mathbb{R}^d$-valued affine jump-diffusion Y and its natural filtration $\mathbb{G}^Y = (\mathcal{G}_t^Y)_{t \geq 0}$, with $\mathcal{G}_t^Y = \sigma(Y_s : 0 \leq s \leq t)$, as subfiltration of $\mathbb{F}$. We then consider $\mu_s \doteq M(s, Y_{s-})$, for some function $M(t, y) = \eta_0(t) + \eta_1(t) \cdot y$, with $\eta \doteq (\eta_0, \eta_1)$ an $\mathbb{R} \times \mathbb{R}^d$-valued bounded continuous function such that μ is nonnegative. We take the left limits of Y, because the state variables process needs not be predictable. That makes no difference for expressions (8) and (10), however, since the pure jump process of (2) has at most a countable number of jumps and thus $\int_t^T M(s, Y_{s-})\mathrm{d}s = \int_t^T M(s, Y_s)\mathrm{d}s$ a.s. for all $T \geq t$. In Sec. 6 we show how suitable choices of η and Y lead to an effective description of the dynamics of human mortality.

4.3. **Combined Model.** We parallel one of the standard setups employed in the valuation of credit-risky securities. We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ large enough to support 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. Moreover, we focus on a representative insured aged x at time 0, with random residual lifetime described as in Sec. 4.2 by an $\mathbb{F}$-stopping time τ.

The filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ represents the flow of information available as time goes by: this includes knowledge of the evolution of all state variables up to each time t and of whether the policyholder has died by then. Formally, we write:

$$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$$

where $\mathcal{G}_t \vee \mathcal{H}_t$ is the σ-algebra generated by $\mathcal{G}_t \cup \mathcal{H}_t$, with

$$\mathcal{G}_t = \sigma(Z_s : 0 \leq s \leq t)$$
$$\mathcal{H}_t = \sigma(\mathbb{I}_{\{\tau \leq s\}} : 0 \leq s \leq t),$$

and where $Z = (X, Y)$ is the joint state variables process in $\mathbb{R}^{k+d}$. Thus, we have $\mathbb{F} = \mathbb{G} \vee \mathbb{H}$, with $\mathbb{G} = \mathbb{G}^X \vee \mathbb{G}^Y$ and with $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ being the smallest 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. 4.1). Under $\mathbb{Q}$, we assume that X and Y are independent and, on the lines of Sec. 4.2, that τ is doubly stochastic driven by $\mathbb{G}^Y \subset \mathbb{F}$. It is worth emphasizing that the latter property does not need to preserve under measure changes and that any such changes determine a modification in the intensity process μ. However, an intensity is guaranteed to exist under $\mathbb{Q}$ if τ admits an intensity under, say, the 'physical' measure $\mathbb{P}$ (see Artzner and Delbaen, 1995).

5. Valuation of Life Insurance Contracts

In this section, we start with the valuation results derived for the basic insurance contracts paying a single (random) benefit upon survival at a given date or upon death within a certain period of time (Sec. 5.1). These results are extended in Sec. 5.2 to several types of insurance contracts, including traditional, unit-linked and indexed policies. We adopt a 'no-arbitrage type' valuation framework, consistent with the methodology proposed by IASB (2001, 2004) for the measurement of insurance liabilities at market value and refer to the prices obtained as *fair values.* The caveats under which such approach can be applied are discussed in Sec. 5.3, while Sec. 5.4 discusses the meaning of the risk-neutral probability measure $\mathbb{Q}$ introduced in the previous section. More generally, the expressions provided in the sequel can be looked to in terms of *expected discounted cash-flow approach* (see Sec. 3): the process r can be replaced by a suitable discount rate, the measure $\mathbb{Q}$ by a measure other than risk-neutral; the results obtained can then be used for risk-management purposes (e.g. sensitivity analysis, stress testing, etc.). In any case, all expectations appearing in the sequel are to be understood as taken under a given probability measure $\mathbb{Q}$.

5.1. **Building-blocks.** We provide the fair values of two basic payoffs involved by standard insurance contracts. These are benefits, of amount possibly linked to other security prices, contingent on survival or death over a given time period. The modeling framework is as described in Sec. 4.3, but for the moment no explicit reference is made to the affine dynamics of the process $Z = (X, Y)$ in $\mathbb{R}^{k+d}$. We only require the short rate process r and the intensity of mortality μ to satisfy the technical conditions stated in Sec. 4.1 and Sec. 4.2.

Proposition 5.1 (Survival Benefit). *Let C be a bounded $\mathbb{G}$-adapted process. Then, the time-t fair value $\mathrm{SB}_t(C_T;T)$ of the time-T survival benefit of amount C_T, with $0 \le t \le T$, is given by:*

$$\begin{aligned}\mathrm{SB}_t(C_T;T) &= E\left[e^{-\int_t^T r_s \mathrm{d}s}\mathbb{I}_{\{\tau>T\}}C_T \middle| \mathcal{F}_t\right] \\ &= 1_{\{\tau>t\}} E\left[e^{-\int_t^T (r_s+\mu_s)\mathrm{d}s}C_T \middle| \mathcal{G}_t\right]\end{aligned} \tag{11}$$

In particular, if C is $\mathbb{G}^X$-adapted, the following holds:

$$\mathrm{SB}_t(C_T;T) = 1_{\{\tau>t\}} E\left[e^{-\int_t^T r_s \mathrm{d}s}C_T \middle| \mathcal{G}_t^X\right] E\left[e^{-\int_t^T \mu_s \mathrm{d}s} \middle| \mathcal{G}_t^Y\right] \tag{12}$$

Proof. We can use the law of iterated expectations, the fact that r and C are $\mathbb{G}$-adapted, to write:

$$\mathrm{SB}_t(C_T;T) = E\left[e^{-\int_t^T r_s \mathrm{d}s}C_T E\left[\mathbb{I}_{\{\tau>T\}} \middle| \mathcal{F}_t \vee \mathcal{G}_T\right] \middle| \mathcal{F}_t\right]$$

From this, we note that $\mathrm{SB}_t(C_T;T)$ is zero on the set $\{\tau \le t\}$. Focusing on the set $\{\tau > t\}$, we exploit the doubly stochastic property, obtaining:

$$SB_t(C_T;T) = 1_{\{\tau>t\}} E\left[e^{-\int_t^T (r_s+\mu_s)ds}C_T \middle| \mathcal{F}_t\right].$$

The conditioning on $\mathcal{F}_t$ can be reduced to that on $\mathcal{G}_t$, as shown in App. C. Expression (12) is finally obtained by exploiting the independence of the filtrations $\mathbb{G}^X$ and $\mathbb{G}^Y$. We use the following fact: given two independent filtrations $\mathbb{F}^1 = (\mathcal{F}_t^1)_{t\ge0}$ and $\mathbb{F}^2 = (\mathcal{F}_t^2)_{t\ge0}$, for any bounded random variables V^1 and V^2 that are $\mathcal{F}_\infty^1$- and $\mathcal{F}_\infty^2$-measurable, one has $E[V^1V^2|\mathcal{F}_t^1 \vee \mathcal{F}_t^2] = E[V^1|\mathcal{F}_t^1]E[V^2|\mathcal{F}_t^2]$ for all $t \ge 0$. □

We note that the results in Prop. 5.1 hold also for survival benefits whose value is known at an earlier date u, with $t < u \le T$. This is useful when valuing contracts providing benefits linked to the level attained by some reference fund at discrete times. Similarly, this helps in valuing some forms of with-profit business showing a 'lagging' mechanism between the indexation date and the payment of the benefits.

Proposition 5.2 (Death Benefit). *Let C be a bounded $\mathbb{G}$-predictable process. Then, the time-t fair value $\mathrm{DB}_t(C_\tau;T)$ of the death benefit of amount C_τ, payable in case the insured dies before time T, with $0 \le t \le T$, is given by:*

$$\begin{aligned}\mathrm{DB}_t(C_\tau;T) &= E\left[e^{-\int_t^\tau r_s \mathrm{d}s}C_\tau \mathbb{I}_{\{t<\tau\le T\}} \middle| \mathcal{F}_t\right] \\ &= \mathbb{I}_{\{\tau>t\}} \int_t^T E\left[e^{-\int_t^u (r_s+\mu_s)\mathrm{d}s}\mu_u C_u \middle| \mathcal{G}_t\right] \mathrm{d}u\end{aligned} \tag{13}$$

In particular, if C is $\mathbb{G}^X$-predictable, the following holds:

$$\mathrm{DB}_t(C_\tau;T) = \mathbb{I}_{\{\tau>t\}} \int_t^T E\left[e^{-\int_t^u r_s \mathrm{d}s}C_u \middle| \mathcal{G}_t^X\right] E\left[e^{-\int_t^u \mu_s \mathrm{d}s}\mu_u \middle| \mathcal{G}_t^Y\right] \mathrm{d}u \tag{14}$$

Proof. We first note that the process $(C_u \exp(-\int_t^u r_s \mathrm{d}s))_{u\ge t}$ is predictable with respect to $(\mathcal{G}_u)_{u\ge t}$. Since the predictable σ-algebra on $\Omega \times [t,T]$ is generated by the left-continuous adapted processes on $[t,T]$ (e.g. Kallenberg, 2002, Lemma 25.1(iii)), any predictable process on $[t,T]$ can be generated by the elementary predictable process $(Z_u)_{t\le u\le T}$, where $Z_u \doteq \widetilde{Z}_t \mathbb{I}_{\{u=t\}} + \widetilde{Z}_h \mathbb{I}_{\{h<u\le k\}}$ with $\widetilde{Z}_h \in \mathcal{G}_h$ for each h and for all $t \le h < k \le T$. For $u = \tau$, We can thus write:

$$\begin{aligned}
E\left[\widetilde{Z}_h \mathbb{I}_{\{h<\tau\le k\}} \middle| \mathcal{F}_t\right] &= E\left[\widetilde{Z}_h \left(\mathbb{I}_{\{\tau>h\}} - \mathbb{I}_{\{\tau>k\}}\right) \middle| \mathcal{F}_t\right] \\
&= E\left[\widetilde{Z}_h E\left[\mathbb{I}_{\{\tau>h\}} - \mathbb{I}_{\{\tau>k\}} \middle| \mathcal{F}_t \vee \mathcal{G}_k\right] \middle| \mathcal{F}_t\right] \\
&= \mathbb{I}_{\{\tau>t\}} E\left[\widetilde{Z}_h \left(e^{-\int_0^h \mu_s \mathrm{d}s} - e^{-\int_0^k \mu_s \mathrm{d}s}\right) \middle| \mathcal{F}_t\right] \\
&= \mathbb{I}_{\{\tau>t\}} E\left[\widetilde{Z}_h \int_h^k e^{-\int_0^u \mu_s \mathrm{d}s} \mu_u \mathrm{d}u \middle| \mathcal{F}_t\right] \\
&= \mathbb{I}_{\{\tau>t\}} E\left[\int_t^T e^{-\int_0^u \mu_s \mathrm{d}s} \mu_u Z_u \mathrm{d}u \middle| \mathcal{F}_t\right],
\end{aligned}$$

where we have used the law of iterated expectations in the second line and the doubly stochastic property in the third. The result then follows by a monotone class argument, paired by use of the (conditional) Fubini theorem and by reduction of the conditioning to $\mathcal{G}_t$ (see App. C). Finally, the factorization (14) is obtained by employing the same argument used in the proof of (12) in Prop. 5.1. □

Prop. 5.1 and Prop. 5.2 formalize what was outlined in Sec. 2. The risk-neutral valuation formula (4) extends naturally to securities providing benefits contingent on mortality, provided they are treated as fictitious securities paying out dividends at a rate of μC, under a fictitious (or risk-adjusted) short rate process $r + \mu$. We point out that the whole filtration $\mathbb{G}$ (and not only $\mathbb{G}^X$) may carry information concerning the random benefits, enabling to deal with deferred insurance contracts and with benefits linked to mortality experience as well. This is the case, for example, of insurance contracts providing returns linked in some way to the demographic profit of the insurer. Even when the factorizations (12) and (14) cannot be used, expressions (11) and (13) have a great analytical tractability in the affine setting, as shown, for example, in Biffis and Millossovich (2004) with reference to insurance contracts embedding options to convert survival benefits into annuities.

5.2. **Life Insurance Contracts.** We can cover quite a large class of insurance contracts by suitably combining the building blocks described in Sec. 5.1. First, we note that the valuation of traditional, unit-linked and indexed contracts can be performed with the same instruments. Indeed, the only difference resides in whether the benefit process C mentioned in Prop. 5.1 and Prop. 5.2 is deterministic or random. In the former case, we can value traditional insurance contracts; in the latter, unit-linked or indexed contracts, depending on whether the evolution of C represents the value of a reference fund or an index. With this regard, we underline that (transformations of) suitable traded security prices may be taken as a proxy for inflation or other indexes. An example is provided below with reference to a floating-rate annuity.

We now present an overview of the contracts encompassed by our framework, always referring to contracts taken out at time 0 by a policyholder then aged x years. We try to use standard actuarial notation whenever possible.

Assurances: the fair value ${}_T\bar{A}_x(C)$ of a T-years *term assurance* guaranteeing a benefit C satisfying the assumptions of Prop. 5.2 and payable in case of death in the period $(0, T]$, is simply given by:

$$_T\bar{A}_x(C) = \mathrm{DB}_0(C_\tau; T).$$

For a *whole life assurance*, we just need to extend the maturity to the maximum age x^* (alternatively, to $+\infty$, provided integrability conditions hold), the one to which no one is assumed to survive:

$$\bar{A}_x(C) = \mathrm{DB}_0(C_\tau; x^* - x).$$

Endowments: the fair value $\bar{A}_{x:T\rceil}(C', C'')$ of a T-years generic *endowment* involving a survival benefit C' and a death benefit C'', satisying the conditions of Prop. 5.1 and Prop. 5.2, respectively, is given by the combination of (11) and (13):

$$\bar{A}_{x:T\rceil}(C', C'') = \mathrm{SB}_0(C'_T; T) + \mathrm{DB}_0(C''_\tau; T)$$

The expression includes *pure* ($C'' = 0$) and *standard* ($C' = C''$) endowments, as well as more complicated structures. For example, the fair value of a *double* endowment (involving a whole life assurance and a pure endowment) is simply obtained by summing $\mathrm{SB}_0(C'_T; T)$ and $\mathrm{DB}_0(C''_\tau; x^* - x)$, possibly with $C' = C''$.

Annuities: the fair value of an annuity can be obtained by combining pure endowments. As an example, consider an annuity paying continuously an instantaneous benefit rate C over the policyholder's life time, with the process C satisfying the conditions of Prop. 5.1. Its fair value $\bar{a}_x(C)$ is given by:

$$\bar{a}_x(C) = \int_0^{x^*-x} \mathrm{SB}_0(C_s; s)\mathrm{d}s \tag{15}$$

Expression (15) can be used to value premium cash flows as well. For example, a stream of continuous premium payments at a rate of π over the period $[0, T)$ (with $T \leq x^* - x$), where the process π satisfies the conditions of Prop. 5.1, has fair value: $\int_0^T \mathrm{SB}_0(\pi_s; s)\mathrm{d}s$.

The discrete versions of the insurance contracts considered is straightforward. For example, consider an immediate annuity paying a benefit C (satisfying the conditions of Prop. 5.1) at the beginning of times $0, 1, 2, \ldots$ in case of the insured's survival at each policy date. The analogue of (15) is given by

$$\ddot{a}_x(C) = \sum_{h=0}^{x^*-x-1} \mathrm{SB}_0(C_h; h) \tag{16}$$

We finally present an example of indexed contract. Let us consider a t-years deferred annuity involving a continuous payment of an indexed benefit from time t onwards, conditional on survival of the policyholder at that time. Suppose the benefit is made of a fixed unit amount plus a variable amount equal, say, to a percentage γ of the level of the short rate at each policy date. The fair value of such contract is given by:

$$\begin{aligned}
\mathrm{SB}_0(\ddot{a}_{x+t}(1+\gamma r);t) &= \sum_{h=t}^{x^*-x-1} \mathrm{SB}_0(1+\gamma r_h;h) \\
&= \sum_{h=t}^{x^*-x-1} E\left[e^{-\int_0^h (r_s+\mu_s)\mathrm{d}s}(1+\gamma r_h)\right] \qquad (17)\\
&= \sum_{h=t}^{x^*-x-1} E\left[e^{-\int_0^h r_s \mathrm{d}s}(1+\gamma r_h)\right] E\left[e^{-\int_0^h \mu_s \mathrm{d}s}\right]
\end{aligned}$$

We now assume that under $\mathbb{Q}$ the processes r, μ, X and Y are affine with respect to $\mathbb{F}$, as described in Sec. 4.1 - 4.3. Since X and Y are assumed to be independent, the joint process $Z = (X, Y)$ is clearly affine. If the benefits considered are functions of Z encompassed by the transform (3), all of the expressions mentioned have a great analytical tractability, requiring the (numerical) solution of ODEs and possibly the (numerical) computation of integrals, while in some cases explicit solutions are available. Clearly, the implementation of popular stochastic models for the term structure is greatly simplified in the context of a stochastic mortality environment.

For example, with reference to the annuity computation (17), let us assume that $\rho_0(t) = 0$ and $\rho_1(t) = b \in \mathbb{R}^k$ for all $t \geq 0$ (see Sec. 4.1). Then, the second line in (17) has the following closed-form (up to ODEs solution) expression:

$$\mathrm{SB}_0(\ddot{a}_{x+t}(1+\gamma r);t) = \sum_{h=t}^{x^*-x-1} e^{\alpha_h(0)+\beta_h(0)\cdot Z_0}\left(\hat{\alpha}_h(0) + \hat{\beta}_h(0)\cdot Z_0\right),$$

where, for each $h = t, \ldots, x^* - x - 1$, the functions α_h, β_h, $\hat{\alpha}_h$ and $\hat{\beta}_h$ solve the ODEs (A.1) to (A.4) associated with the transform (3) for the process Z, with discount rate:

$$\Lambda(t,(x,y)) = R(t,x) + M(t,y) = \eta_0(t) + (b, \eta_1(t))\cdot(x,y),$$

and boundary conditions $\alpha_h(h) = 0$, $\beta_h(h) = 0$, $\hat{\alpha}_h(h) = 1$ and $\hat{\beta}_h(h) = (\gamma b, 0) \in \mathbb{R}^{k+d}$. Alternatively, one can focus on the last line in (17) and solve separately two sets of transforms, one for the process X with discount rate R, the other for the process Y with discount rate M. This approach is followed in the next example.

On the lines of Sec. 4.1, suppose that a risky asset is traded continuously in the market, has price process $S = \exp(X^{(i)})$ and satisfies the conditions stated in App. B. We can then easily extend the setup outlined so far to treat unit-linked contracts having reference fund price process represented by S. For example, a unit-linked pure endowment with maturity T has time-0 fair value given by:

$$\begin{aligned}
\mathrm{SB}_0\left(e^{X_T^{(i)}};T\right) &= E\left[e^{-\int_0^T R(s,X_s)\mathrm{d}s}\ e^{\epsilon(i)\cdot X_T}\right] E\left[e^{-\int_0^T M(s,Y_s)\mathrm{d}s}\right] \\
&= e^{\alpha^R(0)+\beta^R(0)\cdot X_0}\ e^{\alpha^M(0)+\beta^M(0)\cdot Y_0},
\end{aligned}$$

where $\epsilon(i)$ is the vector in $\mathbb{R}^k$ with all null components except the i-th one, which is equal to 1, and the functions α^R and β^R solve the ODEs (A.1) - (A.2) for the transform (3) relative to X and R, with boundary conditions $\alpha^R(h) = 0$ and

$\beta^R(h) = \epsilon(i)$, while the functions α^M and β^M solve the same set of ODEs, but relative to Y and M, and with boundary conditions $\alpha^M(h) = 0$ and $\beta^M(h) = 0$.

5.3. **No-arbitrage valuation.** The financial market described in Sec. 4.1 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 purposes of fair valuation purposes, in the sense of IASB (2001, 2004) (see Sec. 1), 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. 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 insured, so that arbitrage pricing results referring to single contracts can only approximately be scaled up 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).

5.4. **Realistic and Risk Neutral Probabilities.** The measure $\mathbb{Q}$ restricted to $\mathcal{G}^X_\infty$ may be inferred from the security prices observable in the market and a vast literature on the subject is available (see Duffie, 2001, and references therein). The specification of $\mathbb{Q}$ restricted to $\mathcal{G}^Y_\infty$, $\mathbb{Q}|_{\mathcal{G}^Y_\infty}$, is not straightforward instead. Some insights into the 'actuarial' meaning of $\mathbb{Q}|_{\mathcal{G}^Y_\infty}$ can be gathered by reasoning in terms of expected discounted cash flow approach, as suggested by IASB (2001). The key point is that in the first line of expressions (11) and (13) expectations are taken with respect to a 'risk-neutral' probability measure which should incorporate any risk-preferences regarding the riskiness of the insurance cash flows, thus enabling the insurer to discount the cash flows at the risk-free rate.

In many models regarding unit-linked contracts (see Bacinello and Persson, 2002, and references therein), in which the financial risk plays a dominant role, that measure is usually assumed to coincide with the 'historical' probability measure based on statistical estimation. The argument justifying such choice is that the mortality risk affecting the insurer is mainly that of random fluctuations of mortality rates around expected values. Such risk could be tamed by holding a large enough portfolio of similar contracts and, being (in principle) fully diversifiable, the associated risk premium would be priced zero in an efficient market. In other insurance contracts, however, random fluctuations are only a secondary aspect of the risk of mortality to which the insurer is exposed. One can think, for example, of annuity contracts, by far the most subject to systematic departures of mortality from expected levels (e.g. Pitacco, 2003a). In such cases, the usual diversification arguments do not apply and a risk premium should indeed be priced by the market. Using conservative pricing or reserving assumptions to specify $\mathbb{Q}$ would not be a solution either, since both usually overestimate the riskiness of the cash flows and must be considered biased, as is recognized in IASB (2001). The choice of $\mathbb{Q}$ should therefore lie 'somewhere' between a purely statistical assumption and a prudential one.

The measure $\mathbb{Q}$ cannot be easily inferred directly from prices of traded assets, since a deep wholesale market for books of insurance contracts does not exist. However, sales of book of contracts do happen sometimes. Moreover, the 'health' of insurance companies is frequently monitored by financial analysts and actuaries in market terms. In such valuations, a methodology called[3] *Embedded Value Method* is commonly used (see Collins and Keeler, 1993, for example). The embedded value of a book of contracts is given by the sum of the shareholders' net assets backing the book and the value of the business in-force at the valuation date. The latter is the expected present value of the future *distributable earnings* (or *free cash flows*), i.e. future profits emerging (as experience unfolds) on the supervisory basis from policies already written, and is computed by using: realistic assumptions to project the liabilities forward; a risk-adjusted rate, or *risk discount rate*, to discount them back. (See Sheard (2000) for a survey on the valuation assumptions used by UK insurance companies.) This adjusted rate usually incorporates the cost of regulatory capital, tax liabilities and any non-diversifiable risk inherent in the cash flows concerned. The latter component is the only one allowed within the IASB's framework.[4]

The results provided by Prop. 5.1 and Prop. 5.2 suggest the possibility of formalizing the use of risk-adjusted discount rates in the Embedded Value Method, as will be shown in a moment. Perhaps more importantly, one can exploit such valuation technique to compute proxies for secondary market prices (when they are not observable) to which calibrate expressions of the type of (12) or (14), in order to infer the dynamics of the processes of interest under $\mathbb{Q}$. Specifically, a time-t value V_t^{EV} can be placed on a book of insurance liabilities by subtracting the in-force value figure from the market value of the assets backing the book:

$$V_t^{EV} = (\text{net assets})_t - (\text{in-force value})_t.$$

Put it in another way, we want to identify the portion of net assets actually representing the insurer's liabilities from a secondary market perspective. The sign of V_t^{EV} may be negative, thus classifying the book as an asset: in what follows, we make reference to situations in which the value of the liability is positive.

As an example, let us consider a book of homogeneous contracts, say pure endowments with maturity $T > 0$ and a deterministic survival benefit $C \in \mathbb{R}$.

[3]To avoid confusion with terminology that might be more common elsewhere, we specify that by Embedded Value Method we mean Actuarial Appraisal Method with no future business taken into account, i.e. with no allowance for the 'goodwill' component.

[4]Under the IASB's framework, the cost of capital is excluded from valuation and a pre-tax discount rate must be used in discounting pre-tax cash flows.

Let us further assume that the book is backed only by the prudential reserves (no additional shareholders' assets). We can then use a set of realistic assumptions and a suitable risk discount rate, netted of any adjustments for cost of capital or taxes, to determine the book's embedded value, and in turn the quantity V_t^{EV}. Furthermore, suppose that the intensity of mortality is modeled through the affine model $\mu_t \doteq M(t, Y_{t-}) = \eta_0(t) + \eta_1(t) \cdot Y_{t-}$ outlined in Sec. 4.2, but specialized to have $\eta_0(t)$ equal to a deterministic intensity, say $m(t)$, reflecting a realistic (best-estimate) assumptions on mortality. Thus, we have:

$$\mu_t = m(t) + \eta_1(t) \cdot Y_{t-}, \tag{18}$$

where the process $\eta_1 \cdot Y_-$ is a component accounting for random departures from the assumption m. Then, if for each $T \geq t \geq 0$ the quantity ${}_{T-t}p_{x+t}$ represents the best-estimate survival probability implied by $m(t)$ (see (9)), the calibration of $\mathrm{SB}_t(C;T)$ to V_t^{EV} enables to extract a risk-neutral adjustment to that very same realistic basis m, exactly as advocated by IASB (2001, principles 5.1, 5.2 and 5.3). Formally, from (12) we get:

$$\begin{aligned} V_t^{EV} &= C\ E^{\mathbb{Q}}\left[e^{-\int_t^T R(s,X_s)\mathrm{d}s}\Big|\mathcal{G}_t^X\right]\ e^{-\int_t^T m(s)\mathrm{d}s}\ E^{\mathbb{Q}}\left[e^{-\int_t^T \eta_1(s)\cdot Y_s\mathrm{d}s}\Big|\mathcal{G}_t^Y\right] \\ &= C\ E^{\mathbb{Q}}\left[e^{-\int_t^T R(s,X_s)\mathrm{d}s}\Big|\mathcal{G}_t^X\right]\ {}_{T-t}p_{x+t}\ \mathrm{ADJ}^{\mathbb{Q}}(t;T,Y_t), \end{aligned} \tag{19}$$

where $\mathrm{ADJ}^{\mathbb{Q}}(t;T,Y_t) = \exp(\alpha^M(t) + \beta^M(t) \cdot Y_t)$ is the risk-neutral adjustment factor to be estimated from V_t^{EV} after the financial component appearing on the right-hand side of (19) has been computed. This is exactly the method adopted (although in a deterministic framework) by Abbink and Saker (2002) for the fair valuation of several types of insurance contracts in accordance with the IASB's principles. We note that the adjustment factor $\mathrm{ADJ}^{\mathbb{Q}}$ admits an explicit expression in the context of one of the intensity models proposed in Sec. 6.3.

Expression (19) makes clear how the use of risk-discount rates in actuarial practice can be given a sound theoretical background within the framework described. From the first line of (19), we see that we are using a realistic assumption m to project the liability forward and a risk-adjusted rate equal to $R + \eta_1 \cdot Y$ to discount it back, the adjustment made to R accounting for the mortality risk affecting the cashflow C. The second line in (19) shows that we can equivalently use a risk-adjusted assumption on mortality, ${}_{T-t}p_{x+t}\ \mathrm{ADJ}^{\mathbb{Q}}(t;T,Y_t)$, with the unadjusted discount rate R. Put it in another way, the stochastic adjustment to the discount rate translates into a multiplicative adjustment (at least in this setup) to the realistic survival probabilities.

6. Mortality Modeling

This section is devoted to show how the stochastic model presented in sec. 4.2 can be implemented successfully for mortality modeling. We will focus on mortality evolution at old (pensionable) ages, demonstrating how the setup can handle the risk of longevity.

In Sec. 6.1 we recall the main features of recent mortality trends in most developed countries. Then, Sec. 6.2 shows how the affine setup enables to easily compute demographic synthetic measures that can be helpful for understanding how the model parameters translate into survival function and death curve[5] shapes. Finally, Sec. 6.3 proposes two worked-out examples of affine models for μ. The first is a unidimensional Poisson-Gaussian process, while the second one is a bidimensional square-root diffusion. Some numerical examples are offered for both models.

6.1. **Mortality Trends.** The most important features of mortality trends in developed countries appear to be the following (e.g. MacDonald, Cairns, Gwilt and Miller (1998), Wilmoth and Horiuchi (1999)):

- a remarkable increase in life expectancy at birth, supported by reductions in mortality particularly at very young ages;
- an increasing concentration of deaths around the mode at old ages of the curve of deaths (the so called Lexis point), so that the survival function $S(\cdot) \doteq \mathbb{P}(\tau_0 > \cdot)$ tends to assume a more and more rectangular shape (*rectangularization* phenomenon);
- a shift of the Lexis point towards older and older ages (*expansion* phenomenon).

We note that the rectangularization and expansion phenomena affect long-term living benefits in opposite ways: while the former reduces the variability of cash flows, the latter increases their duration. As a consequence, the risk of mortality random fluctuations tends to be reduced, while at the same time a systematic risk component (longevity risk) acquires greater importance. A graphical example of the two phenomena is provided in Fig. 1 and Fig. 2, where the survival functions and death curves derived from the Italian contemporary tables SIM-1931 to SIM-1992. It is clear that any candidate model for μ must be able to capture the dynamics just described. Sec. 6.2 and 6.3 treat this issue.

< FIGURE 1 ABOUT HERE >

< FIGURE 2 ABOUT HERE >

6.2. **Demographic Measures.** When dealing with the parametrization of affine intensities of mortality, it is useful to compute several markers providing an indication of how the parameters translate into survival function and death curve shapes.

We can obtain a measure of the rectangularization of the survival function implied by μ by employing the entropy of the survival function $S(\cdot)$, defined as (e.g. Keyfitz, 1985, Ch. 3):

$$H = -\frac{\int_0^{x^*} \log(S(x))S(x)\mathrm{d}x}{\int_0^{x^*} S(x)\mathrm{d}x},$$

where x^* is the maximum age to which no one is assumed to survive. The entropy is a measure of uncertainty of a physical system: in a life table, it measures the extent to which deaths are spread over a broad age range. The closer H is to

[5]In standard actuarial/demographic terminology, the 'curve of deaths' is the density of τ.

zero, the more rectangular the shape of the survival function. Within the joint doubly stochastic and affine framework, the entropy of the 'residual' life table relative to individuals aged x at time 0 boils down to:

$$H_x(Y_0) = -\frac{\int_0^{x^*-x}[\alpha_t(0) + \beta_t(0) \cdot Y_0]\ e^{\alpha_t(0)+\beta_t(0)\cdot Y_0}\ \mathrm{d}t}{\int_0^{x^*-x} e^{\alpha_t(0)+\beta_t(0)\cdot Y_0}\ \mathrm{d}t}, \tag{20}$$

where we have used expression (8) and the transform (3) for the survival function $S_x(\cdot) \doteq \mathbb{P}(\tau_x > \cdot)$, and where for each t the functions $\alpha_t(\cdot) \doteq \alpha(\cdot; 0, t)$ and $\beta_t(\cdot) \doteq \beta(\cdot; 0, t)$ solve the ODEs (A.1) - (A.2) with terminal conditions $\alpha_t(t) = 0$ and $\beta_t(t) = 0$.

Furthermore, the expectation of life at age x can be computed in our setting by using the following expression:

$$\mathring{\mathrm{e}}_x(Y_0) = \int_0^{x^*-x} e^{\alpha_t(0)+\beta_t(0)\cdot Y_0}\mathrm{d}t, \tag{21}$$

where $\mathring{\mathrm{e}}_x(Y_0)$ is the standard demographical and actuarial notation for $E[\tau_x]$. Expressions (20) and (21) require in general a quadrature procedure paired by the recursive solution of ODEs. The results can be employed to have a quick comparison with the benchmarks provided by institutions such as the UN or the OECD.

In case we adopt an affine model of the type (18), we can exploit H_x to gather an approximate indication of the deterministic intensity equivalent to μ in terms of the gain in life expectancy given by the dynamics of $\eta_1 \cdot Y_-$. In particular, let $\mathring{\mathrm{e}}_x(Y_0)$ be the life expectancy computed assuming μ is specified by (18) and let $\mathring{\mathrm{e}}_x$ and H_x be the life expectancy and entropy associated with the deterministic assumption m. Then, the following expression from Keyfitz (1985, p. 62) enables to determine the constant decrease Δ in m associated with the gain in life expectancy provided by the random component $\eta_1 \cdot Y$:

$$\Delta \cong \frac{\mathring{\mathrm{e}}_x(Y_0) - \mathring{\mathrm{e}}_x}{\mathring{\mathrm{e}}_x \bar{t}_x}, \tag{22}$$

where $\bar{t}_x = (\int_0^{x^*-x} t\ {}_tp_x \mathrm{d}t)/(\int_0^{x^*-x} {}_tp_x \mathrm{d}t)$, with ${}_tp_x = \exp(-\int_0^t m(s)\mathrm{d}s)$ for all $t \geq 0$. In other words, we identify the deterministic intensity $m - \Delta$ equivalent to μ in terms of gain in life expectancy. This can be used to understand the balance between the random departures from m, implied by the process $\eta_1 \cdot Y_-$, and a constant shift in the intensity m at all times $t \geq 0$.

We can go further, and obtain also an indication of the constant proportional decrease δ in m associated with the gain in life expectancy provided by $\eta_1 \cdot Y_-$, i.e. the deterministic intensity $m(1-\delta)$ equivalent to μ in terms of gain in life expectancy. This is easily derived through the following expression from Keyfitz (1985, p. 64):

$$\delta \cong \frac{\mathring{\mathrm{e}}_x(\Lambda) - \mathring{\mathrm{e}}_x}{\mathring{\mathrm{e}}_x H_x} \tag{23}$$

6.3. **Examples of affine mortality models.** We begin with a one dimensional example providing explicit solutions for the survival probabilities (8) (see App. D).

On the lines of expression (18), we set $\eta_0(t) = m(t)$ and obtain the following model for μ:

$$\mu_t = m(t) + \eta_1 Y_{t-}, \tag{24}$$

with $\eta_1 \in \mathbb{R}$. The deterministic function m in (24) may represent, for example:

- **(a)** a realistic (best-estimate) assumption on μ representing unbiased expectations about the future based on the best available (company, industry or market) information available, when performing market valuations;
- **(b)** a prudential or pricing demographic basis, when carrying out risk analysis for solvency, reserving or pricing purposes;
- **(c)** an available mortality table, when doing realistic mortality projections of a population of insureds.

In all cases, the component $\eta_1 Y_-$ will represent random departures from the initially chosen basis, capturing random fluctuations as well as systematic deviations, depending on the specification of the dynamics of Y.

We then let the state variable process Y in $\mathbb{R}$ be affine, with dynamics described by the SDE:

$$\mathrm{d}Y_t = \gamma\left(\bar{y}(t) - Y_t\right)\mathrm{d}t + \sigma \mathrm{d}W_t - \mathrm{d}J_t, \tag{25}$$

The bounded measurable function $\bar{y}(t)$ represents a time-varying target to which Y reverts with speed determined by the coefficient $\gamma > 0$, after any shocks due to fluctuations or jumps occur. The diffusion component accounts for fluctuations due to the standard Brownian motion W, whose overall effect on the variations of Y depends on the positive coefficient σ. The jump component is a compound Poisson process independent of W, with constant jump-arrival intensity $k > 0$ and jump sizes exponentially distributed with mean $j > 0$. Since we are interested in mortality evolution at old ages (alternatively, in departures from a demographic basis due to longevity risk), we consider downward jumps, i.e. mortality declines: for any jump time τ_i, we have $\Delta Y_{\tau_i} = Y_{\tau_i} - Y_{\tau_{i-}} = -Z_i$, with $Z_i \sim \mathrm{Exp}(1/j)$ for all $i = 1, 2, \ldots$ (i.e. Z_i has density $\nu(z) = (1/j)\exp(-z/j)$ for $z \in [0, \infty)$).

The extent to which the demographic basis m is 'binding' for the evolution of μ depends on the trade-off between the reversion mechanism to the target $\bar{y}(t)$ and the parametrization of the jump component of (25). In other words, it is crucial the extent to which jumps actually determine a change in the intensity trend. The use of a discontinuous setup for modeling the intensity of mortality may sound unfamiliar. However, we note that neither the size nor the frequency of discontinuity shocks need be unreasonably large. Moreover, the gain in flexibility and distributional richness that can be achieved by adding a discontinuous source of risk in the dynamics of μ can justify the 'abuse' made in terms of path by path behaviour.

The main drawback of the model (24)-(25) is that Y is allowed to take negative values with positive probability. This means that some realizations of $\eta_1 Y_-$ can fall below $-m$, thus violating the non-negativity constraint on the intensity process μ. The problem is similar to that arising for term structure models allowing the nominal short rate to take negative values (e.g. Duffie, 2001, pp. 140-141) and the same as that presented by any credit spreads model based on Gaussian diffusions (e.g. Lando, 1998; Schönbucher, 2003, Sec. 7.1.1). In practice, we would set the parameters $(\eta_1,\ \gamma,\ \bar{y},\ \sigma,\ k,\ j)$ in such way to ensure that our process μ takes negative values with low probability.

By taking expectations in (25), using Fubini's Theorem (e.g. Kallenberg, 2002, p. 14) and differentiating, we get an ODE that can be easily solved yielding the unconditional expectation of Y_t, for each $t \geq 0$:

$$E\left[Y_t\right] = e^{-\gamma t}\left(\gamma \int_0^t e^{\gamma s}\ \bar{y}(s)\ ds + \frac{jk}{\gamma} + E[Y_0]\right) - \frac{jk}{\gamma}$$

From this, by taking limits for $t \to \infty$, we obtain the unconditional long-term mean of Y, which we denote by $\bar{Y}_\infty$. In the simple case of $\bar{y}(t) = \bar{y} \in \mathbb{R}$ for all $t \geq 0$, we get $\bar{Y}_\infty = \bar{y} - jk/\gamma$. This can be used to check the average long-term impact on μ of the parameters $(\gamma,\ k,\ j)$. In particular, we have an immediate indication of how the attraction to the central target of the diffusion component of Y competes with the jump departures, these resulting amplified by smaller values of γ, as expected.

Explicit solutions for the survival probabilities are given in App. D. We provide numerical examples of the model by using three Italian (males) mortality tables to perform projections and make comparisons. We use table SIM92, a period table usually employed to price assurances, the (projected) annuitants table RG48 and the (projected) table relative to Italian males born in 1948, which we call table COH48 in the sequel. We focus on pensionable ages and set $x = 65$. The table choices for m and the parameter values for the process $\eta_1 Y_-$ are all reported in Tab. 1. We note that in the first two cases a constant target $\bar{y}$ for the diffusion component of Y is used, while in the other cases a time-varying central target is employed. The final target value $\bar{y}(x^* - x)$, is also reported. We set $x^* = 114$, consistently with the projection provided by table RG48. One can get an immediate idea of the demographic meaning of the parameter choices by looking at Fig. 3, where the corresponding survival functions are plotted against those relative to tables SIM92, COH48 and RG48. More precise insights can be gathered by examining the measures reported in Tab. 2 and described in Sec. 6.2. We note that $\bar{Y}_\infty$ is computed by replacing in its formula the quantity $\bar{y}$ with $\bar{y}(x^* - x)$ in cases C, D and E. The life expectancy at age 65 is a curtate expectation of life (e.g. Bowers, Gerber, Hickman, Jones and Nesbitt, 1997).

< TABLE 1 ABOUT HERE >

< TABLE 2 ABOUT HERE >

< FIGURE 3 ABOUT HERE >

We now move on to the second worked-out example, which is continuous and ensures that the intensity of mortality stays nonnegative over time. We consider an affine process $Y = (\mu, \bar{\mu})$ in $\mathbb{R}^2_+$, whose first component is the random intensity of mortality μ itself, while the second component represents the stochastic drift $\bar{\mu}$ of such intensity. In other words, we are specializing the setting of Sec. 4.2 to an intensity $\mu_t = M(t, Y_-) = \eta_0(t) + \eta_1(t) \cdot Y$ with $\eta_0 = 0$, $\eta_1 = (1, 0)^\mathsf{T}$ and Y continuous affine in $\mathbb{R}^2_+$. In particular, we assume that Y is a square-root diffusion and that the processes μ and $\bar{\mu}$ have dynamics described by the following SDEs:

$$\mathrm{d}\mu_t = \gamma_1(\bar{\mu}_t - \mu_t)\mathrm{d}t + \sigma_1\sqrt{\mu_t}\,\mathrm{d}W_t^1 \tag{26}$$

$$\mathrm{d}\bar{\mu}_t = \gamma_2(m(t) - \bar{\mu}_t)\mathrm{d}t + \sigma_2\sqrt{\bar{\mu}_t - m^*(t)}\,\mathrm{d}W_t^2, \tag{27}$$

where: $W = (W^1, W^2)$ is a standard Brownian motion in $\mathbb{R}^2$; $\gamma_1, \gamma_2 > 0$ are parameters representing the 'speed of mean reversion' of μ to $\bar{\mu}$ and of $\bar{\mu}$ to m, after any fluctuations due to W occur; the functions m and m^* are bounded and continuous. The function m is a suitable demographic basis (see points (a), (b) and (c) above), such as an available mortality table acting as a time-varying target for the stochastic drift $\bar{\mu}$ of μ. The function m^* is a time-varying lower boundary for the stochastic drift $\bar{\mu}$. It can be interpreted as a more optimistic assumption (in terms of mortality improvements) than that implied by m. Indeed, to make sure that $Y = (\mu, \bar{\mu})$ is well-defined, i.e. that $\mu \geq 0$ a.s. and $\bar{\mu} \geq m^*$ a.s., we have to impose the following conditions: $m \geq m^* \geq 0$, $\bar{\mu}_0 \geq m^*(0)$ and $\mu_0 \geq 0$ (see App. E for details). As a consequence, we see that m and $\bar{\mu}$ always dominate m^*. It seems thus natural to interpret m^* as an asymptotic or as a limiting intensity: the former is an intensity representing asymptotic biological limits by which mortality improvements are constrained; the latter represents mortality rates (e.g. provided by a more advanced, in demographical terms, population) to which the mortality of a population is assumed to converge over time. Note, however, that the conditions stated above only ensure that μ is nonnegative, so that some paths of μ may actually fall below m^*.

We note that the model takes into account the risk of random fluctuations around μ and around the drift's target m, thus enabling to deal with the risk of longevity. The Brownian motions W^1 and W^2 account for random fluctuations in the variations of μ and $\bar{\mu}$ over time, the extent of such fluctuations depending on the state-dependent[6] volatility components with parameters $\sigma_1, \sigma_2 > 0$. Although W^1 and W^2 for simplicity are assumed to be independent (the assumption can be freely relaxed, as App. E shows), μ and $\bar{\mu}$ are clearly correlated with each other through the drift component of μ.

The model proposed is quite flexible in describing the evolution of mortality, as we show through an application to mortality projection at pensionable ages. We assume that the intensity m is that implied by a recent available life table, *from* which we want to model mortality improvements. We take $x = 65$ and use the Weibull law of mortality for the specification of m and m^*, i.e., for example:

$$m(x+t) = \frac{c}{\theta^c}(x+t)^{c-1} \qquad \text{with } \theta, c > 0 \tag{28}$$

and the corresponding survival function is given by $S_x(t) = \exp((-1/\theta^c)[(x+t)^c - x^c])$, for all $t \geq 0$. Several other choices could be made: we opt for this one for demographical and computational reasons. The Weibull law allows for an age-dependent rate of variation in the intensity and admits explicit expressions for a number of markers that can be used to quickly understand the demographic implications of any choice of parameters. For example, if the intensity of mortality is given by (28), then the life expectancy at birth can be obtained through the expression $\mathring{e}_0 = \theta\,\Gamma(1 + 1/c)$, where Γ indicates the Gamma function $\Gamma(a) =$

[6] That is, the model allows for conditional heteroskedacity.

$\int_0^\infty e^{-t} t^{a-1} dt$ (with $a > 0$). Moreover, simple calculus yields that the life expectancy at age x can be expressed as $\mathring{e}_x = \theta[\Gamma(1+1/c) - \gamma(1+1/c, (x/\theta)^c)]$, where $\gamma(a, y) = \int_0^y e^{-t} t^{a-1} dt$ is the Incomplete-Gamma function (e.g. Abramowitz and Stegun, 1974, Sec. 6).

We assume that $x^* = 114$ is the age at which no one survives, consistently with the annuitants life table RG48. We then fit a Weibull survival function to the projected table COH48 and use the corresponding Weibull intensity as our assumption m. We set an asymptotic Weibull intensity m^* which is low enough to allow fluctuations of $\bar{\mu}$ below m, yet large enough to prevent mortality improvements from being unreasonable. Tab. 3 reports the Weibull parameters (c, θ), relative to the demographic basis m, the parameter values (c^*, θ^*), relative to m^*, and the parameters (c_0, θ_0) relative to the 'initial' mortality level μ_0, i.e. that of a generation successive to the 1948 one. The demographic meaning of the parameters is synthesized by the life expectancy at age 65. One should think of m as representing a recent (latest) available life table, and regard μ_0 as representing an estimate of the intensity of mortality of an individual aged 65 belonging to the cohort object of projection.

The values of the drift and volatility parameters are all quoted in Tab. 3 except γ_2, which has been allowed to vary for sensitivity analysis purposes. Fig. 4 depicts several survival functions obtained with γ_2 ranging from 0.50 (lowest rectangularization effect) to 0.05 (greatest rectangularization effect). The survival functions relative to m and to the projected table RG48 (to be used as a benchmark) are also depicted. The effect of the parameter choices on death curves is shown in Fig. 5, where one can appreciate the joint rectangularization and expansion effect on the projected density surface.

< TABLE 3 ABOUT HERE >

< FIGURE 4 ABOUT HERE >

< FIGURE 5 ABOUT HERE >

7. Conclusion

In this work we have shown how mortality can be modeled stochastically in order to provide a rich and flexible framework for actuarial valuations. We have exploited the parallel between life insurance business and credit-sensitive securities, first pointed out in Artzner and Delbaen (1995). By using affine jump-diffusions as driving processes for the evolution of demographic and financial risk factors, we have demonstrated how actuarial valuations can deal simultaneously with different sources of risks. In particular, we have provided closed-form (up to ODEs solution) expressions for the (fair) value of a number of life insurance contracts, either traditional, indexed and unit-linked. The Embedded Value Method, a popular actuarial valuation technique, has been looked at in the context of fair value accounting, in order to understand how risk-adjusted discount rates translate into adjustments to best-estimate assumptions and vice-versa. Finally, two affine models for the intensity of mortality have been proposed, paying particular attention to the demographic meaning of their parameters and to their ability in capturing the dynamics of mortality evolution. Examples of mortality projections at old ages have been provided, showing how the models proposed can handle effectively the risk of longevity.

References

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

Abramowitz, M. and I. Stegun (1974). *Handbook of Mathematical Functions*. Dover Publications, New York.

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

Bacinello, A. and S. Persson (2002). Design and pricing of equity-linked life insurance under stochastic interest rates. *The Journal of Risk & Insurance*, vol. 3(2):6–21.

Bakshi, G. and D. Madan (2000). Spanning and derivative security valuation. *Journal of Financial Economics*, vol. 55(2):205–238.

Bakshi, G., C. Cao and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, vol. 52(5):2003–2049.

Bates, D. (2000). Post-87' crash fears in S&P500 futures options. *Journal of Econometrics*, vol. 94(1-2):181–283.

Biffis, E. (2003). Affine processes in mortality modelling: an actuarial application. Quaderno n. 156, Department of Applied Mathematics, University of Trieste.

Biffis, E. and P. Millossovich (2004). The fair value of guaranteed annuity options. Quaderno n. 162, Department of Applied Mathematics, University of Trieste.

Billingsley, P. (1995). *Probability and Measure*. Wiley, New York, third edn.

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

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

Collins, S. and D. Keeler (1993). Analysis of life company financial performance. Staple Inn Actuarial Society, London.

Cox, J., J. Ingersoll and S. Ross (1985). A theory of the term structure of interest rates. *Econometrica*, vol. 53(2):385–408.

Dahl, M. (2004). Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts. *Insurance: Mathematics* & *Economics*, vol. 35(1):113–136.

Dai, Q. and K. Singleton (2000). Specification analysis of affine term structure models. *Journal of Finance*, vol. 55(5):1943–1978.

Deelstra, G. and F. Delbaen (1998). Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis*, vol. 14(1):77–84.

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., 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(3):984–1053.

Eraker, B., M. Johannes and N. Pohlson (2003). The impact of jumps in volatility and returns. *Journal of Finance*, vol. 58(3):1269–1300.

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.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, vol. 6(2):327–343.

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

IASB (2004). *International Financial Reporting Standard n. 4*. International Accounting Standards Board.

Jeanblanc, M. and M. Rutkowski (2000). Modeling of default risk: an overview. In J. Yong and R. Cont (eds.), *Mathematical Finance: Theory and Practice*, pp. 171–269. Higher Education Press, Beijin.

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer-Verlag, New York, second edn.

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, second edn.

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*, vol. 2(2-3):99–120.

MacDonald, A., A. Cairns, P. Gwilt and K. Miller (1998). An international comparison of recent trends in population mortality. *British Actuarial Journal*, vol. 4(1):3–141.

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

Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics & Economics*, vol. 29(2):231–245.

Pitacco, E. (2003a). Longevity risk in living benefits. In E. Fornero and E. Luciano (eds.), *Developing an Annuity Market in Europe*. Edward Elgar Publishing Ltd.

Pitacco, E. (2003b). Survival models in actuarial mathematics: from Halley to longevity risk. Quaderno n. 2, Department of Applied Mathematics, University of Trieste. Forthcoming in *Insurance: Mathematics & Economics.*

Protter, P. (2004). *Stochastic Integration and Differential Equations*. Springer-Verlag, Heidelberg, second edn.

Schönbucher, P. (2003). *Credit Derivatives Pricing Models*. Wiley Finance.

Scott, L. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Application of Fourier inversion methods. *Mathematical Finance*, vol. 7(4):413–426.

Sheard, M. (2000). Summary and comparison of approaches used to measure life office values. Staple Inn Actuarial Society, London.

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

Wilmoth, J. and S. Horiuchi (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography*, vol. 36(4):475–495.

Appendix A. Affine Processes

We now specify explicitly the affine dependence on $X \in \mathbb{R}^n$ of the coefficients appearing in the SDE (2) of Sec. 3. In particular, we have:

$$\begin{aligned}
&\delta(t,x) = d_0(t) + d_1(t)x \\
&\left(\sigma(t,x)\sigma(t,x)^\intercal\right)_{i,j} = (V_0(t))_{i,j} + (V_1(t))_{i,j} \cdot x \qquad\qquad i,j = 1,\dots,n \\
&\kappa(t,x) = k_0(t) + k_1(t) \cdot x
\end{aligned}$$

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 $k \doteq (k_0, k_1)$ 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 time-dependent jump-size distribution ν_t is determined by its Laplace transform: $\theta(t,c) = \int_{\mathbb{R}^n} e^{c\cdot z} d\nu_t(z)$, defined for $t \in [0,\infty)$, $c \in \mathbb{C}^n$ and such that the integral is finite. The transform θ and the functions d, V and k completely determine the distribution of X, once an initial condition X_0 is given.

Focusing now on the transform (3), when (d,V,k,θ) and the affine function Λ are 'extended well-behaved' in the sense defined by Duffie, Pan and Singleton (2000, App. A), the functions $\alpha(\cdot) \doteq \alpha(\cdot; a, T)$ and $\beta(\cdot) \doteq \beta(\cdot; a, T)$ satisfy the following ODEs:

$$\dot{\beta}(t) = \lambda_1(t) - d_1(t)^\intercal \beta(t) - \frac{1}{2}\beta(t)^\intercal V_1(t)\beta(t) - k_1(t)\left[\theta(t,\beta(t)) - 1\right] \tag{A.1}$$

$$\dot{\alpha}(t) = \lambda_0(t) - d_0(t) \cdot \beta(t) - \frac{1}{2}\beta(t)^\intercal V_0(t)\beta(t) - k_0(t)\left[\theta(t,\beta(t)) - 1\right] \tag{A.2}$$

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:

$$\dot{\hat{\beta}}(t) = -d_1(t)^\intercal \hat{\beta}(t) - \beta(t)^\intercal V_1(t)\hat{\beta}(t) - k_1(t)\left[\Theta(t,\beta(t)) \cdot \hat{\beta}(t)\right] \tag{A.3}$$

$$\dot{\hat{\alpha}}(t) = -d_0(t) \cdot \hat{\beta}(t) - \beta(t)^\intercal V_0(t)\hat{\beta}(t) - k_0(t)\left[\Theta(t,\beta(t)) \cdot \hat{\beta}(t)\right] \tag{A.4}$$

with boundary conditions $\hat{\alpha}(T) = c$ and $\hat{\beta}(T) = b$, where $\Theta(t,c)$ denotes the gradient of $\theta(t,c)$ with respect to $c \in \mathbb{C}^n$, i.e. $\Theta(t,c) = \int_{\mathbb{R}^n} c \exp(c \cdot z)\nu_t(dz)$. 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 $c^\intercal V_1(t) d$.

Solutions to the ODEs above can be found explicitly in some cases, through the use of numerical methods in other cases. The choice of a jump distribution with an explicitly known or easily computed transform θ is clearly important for computational tractability.

Appendix B. No-arbitrage restrictions

We provide here the no arbitrage constraints for any risky security of the form $S = \exp(X_i)$ with affine dividend yield $\zeta(t, X_t) = q_0(t) + q_1(t) \cdot X_t$, as considered in Sec. 4.1. By applying Itô's formula to S (e.g. Protter, 2004, pp. 78-79) 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):

$$
\begin{aligned}
(d_1(t))_i &= \rho_1(t) - q_1(t) - \frac{1}{2}(V_1(t))_{i,i} - k_1(t)\left[\theta(t,\epsilon(i)) - 1\right] \\
(d_0(t))_i &= \rho_0(t) - q_0(t) - \frac{1}{2}(V_0(t))_{i,i} - k_0(t)\left[\theta(t,\epsilon(i)) - 1\right]
\end{aligned}
$$

where $\epsilon(i)$ indicates the vector in $\mathbb{R}^k$ with all null components but the i-th, which is equal to 1.

Appendix C. Construction of the random time of death

We take as given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider a filtration $\mathbb{G}$ satisfying the usual conditions and such that $\mathcal{G}_\infty \subset \mathcal{F}$. We assume that a nonnegative $\mathbb{G}$-predictable process μ is given satisfying $\int_0^t \mu_s \mathrm{d}s < \infty$ a.s. for all $t > 0$. We then fix an exponential random variable Φ with parameter 1, independent of $\mathcal{G}_\infty$, and define the random time of death τ as the first time when the process $\int_0^\cdot \mu_s \mathrm{d}s$ is above the random level Φ, i.e.we set:

$$
\tau \doteq \inf\left\{ t \in \mathbb{R}_+ : \int_0^t \mu_s \mathrm{d}s \geq \Phi \right\}.
$$

We note that $\{\tau > T\} = \{\int_0^T \mu_s \mathrm{d}s < \Phi\}$, for $T \geq 0$. Thus, for $T \geq t \geq 0$, by using the law of iterated expectations and by exploiting the independence of Φ and $\mathcal{G}_T$, one obtains that:

$$
\mathbb{P}\left(\tau > T | \mathcal{G}_t\right) = E\left[\mathbb{P}\left(\Phi > \int_0^T \mu_s \mathrm{d}s \Big| \mathcal{G}_T\right) \Big| \mathcal{G}_t\right] = E\left[e^{-\int_0^T \mu_s \mathrm{d}s} \Big| \mathcal{G}_t\right]. \tag{C.1}
$$

Note that the same result holds for $0 \leq T < t$. The filtration $\mathbb{F} = (\mathcal{F})_{t\geq 0}$ of Sec. 4.3 is finally obtained by defining $\mathcal{F}_t \doteq \mathcal{G}_t \vee \mathcal{H}_t$, where $\mathcal{H}_t = \sigma(\mathbb{I}_{\{\tau \leq s\}}; 0 \leq s \leq t)$. We now observe that $\{\tau > t\}$ is an atom of $\mathcal{H}_t$. We can thus verify (using a slight extension of Billingsley, 1995, ex. 34.4, p. 455) that we have constructed a doubly stochastic $\mathbb{F}$-stopping time driven by $G \subset \mathbb{F}$ in the following way:

$$
\begin{aligned}
\mathbb{P}\left(\tau > T | \mathcal{G}_T \vee \mathcal{F}_t\right) = \mathbb{I}_{\{\tau > t\}} E\left[\mathbb{I}_{\{\tau > T\}} | \mathcal{G}_T \vee \mathcal{H}_t\right] = \mathbb{I}_{\{\tau > t\}} \frac{\mathbb{P}\left(\{\tau > T\} \cap \{\tau > t\} | \mathcal{G}_T\right)}{\mathbb{P}\left(\tau > t | \mathcal{G}_T\right)} = \\
= \mathbb{I}_{\{\tau > t\}} \frac{\mathbb{P}\left(\tau > T | \mathcal{G}_T\right)}{\mathbb{P}\left(\tau > t | \mathcal{G}_T\right)} = \mathbb{I}_{\{\tau > t\}} e^{-\int_t^T \mu_s \mathrm{d}s},
\end{aligned}
$$

Finally, we observe that the independence of Φ and $\mathcal{G}_T$, and in particular of Φ and $\mathcal{G}_t \vee \sigma(\exp(-\int_0^T \mu_s \mathrm{d}s))$, enables to write from (C.1):

$$
E\left[e^{-\int_0^T \mu_s \mathrm{d}s} \Big| \mathcal{G}_t \vee \sigma(\Phi)\right] = E\left[e^{-\int_0^T \mu_s \mathrm{d}s} \Big| \mathcal{G}_t\right],
$$

for all $T \geq t \geq 0$. But $\mathcal{G}_t \subset \mathcal{F}_t \subset \mathcal{G}_t \vee \sigma(\Phi)$ and therefore also the following holds:

$$
E\left[e^{-\int_0^T \mu_s \mathrm{d}s} \Big| \mathcal{F}_t\right] = E\left[e^{-\int_0^T \mu_s \mathrm{d}s} \Big| \mathcal{G}_t\right].
$$

Similar reasoning yields that we can replace the conditioning on $\mathcal{F}_t$ with that on $\mathcal{G}_t$ in Prop. 5.1 and 5.2.

We do not take $\mathbb{G} \vee \sigma(\Phi)$ as our reference filtration $\mathbb{F}$ because, if that were the case, the stopping time τ would be predictable and would not admit an intensity. The construction outlined ensures instead that τ is a totally inaccessible stopping time, a concept intuitively meaning that the insured's death arrives as a total surprise to the insurer (see Protter, 2004, Ch. III.2, for details).

Appendix D. Poisson-Gaussian process: explicit solutions

Since for model (24) the factorization exploited in (19) holds, we can focus on the random component $\eta_1 Y$, where Y has dynamics described through the SDE (25). The parameters of App. A have the following specification in this particular case:

$$d_0(t) = \gamma \bar{y}(t) \qquad d_1 = -\gamma \qquad V_0 = \sigma^2 \qquad V_1 = 0$$
$$k_0 = k \qquad k_1 = 0 \qquad \lambda = (0, \eta_1)$$

Moreover, the jump transform θ is given by:

$$\theta(c) = \int_0^\infty \frac{1}{j} e^{-\left(c+\frac{1}{j}\right)z} \mathrm{d}z = \frac{1}{1+jc} \qquad \text{for} \qquad \Re(c) > -\frac{1}{j},$$

where $\Re(c)$ denotes the real part of $c \in \mathbb{C}$. We can solve (A.1) with boundary condition $\beta(T) = 0$, thus getting:

$$\beta(t) = \eta_1 \frac{e^{-\gamma(T-t)} - 1}{\gamma},$$

which can be plugged into (A.2) to obtain (using the condition $\alpha(T) = 0$ and integrating):

$$\begin{aligned}\alpha(t) =& \frac{\sigma^2}{2\gamma^2}\left[\eta_1^2 (T-t) + \eta_1 \beta(t) - \frac{\gamma}{2}\beta^2(t)\right] - \frac{kj\eta_1}{j\eta_1 - \gamma}(T-t) + \\ &- \frac{k}{j\eta_1 - \gamma} \log|1 + j\beta(t)| + \gamma \int_t^T \bar{y}(u)\beta(u)\mathrm{d}u.\end{aligned}$$

Appendix E. Bidimensional square-root diffusion: technical conditions

We write explicitly the parameters of App. A relative to the dynamics of the process $Y = (\mu, \bar{\mu})^\intercal$ described by the SDEs (26)-(27). Specifically, we have:

$$d_0(t) = \begin{bmatrix} 0 \\ \gamma_2 m(t) \end{bmatrix} \qquad d_1 = \begin{bmatrix} -\gamma_1 & \gamma_1 \\ 0 & -\gamma_2 \end{bmatrix}$$

$$V_0(t) = \begin{bmatrix} 0 & 0 \\ 0 & -\sigma_2^2 m^*(t) \end{bmatrix} \qquad V_1 = \left[\begin{array}{cc} \sigma_1^2 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & \sigma_2^2 \end{array}\right]$$

$$\kappa = (0, (0,0)^\intercal) \qquad \lambda = (0, (1,0)^\intercal)$$

These can be used to solve the ODEs (A.1) to (A.4) when implementing the model.

We now move on to the technical conditions ensuring that the process $Y = (\mu, \bar{\mu})^\intercal$ is well-defined. Let us first focus on the square-root process $\bar{\mu}$. Since the function $f(t,x) = \sigma_2\sqrt{|x - m^*(t)|}$ (with $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$) satisfies the Hölder condition:

$$|f(t,x) - f(t,y)| \leq K\sqrt{|x-y|},$$

with K a positive constant, then strong uniqueness holds for all the SDEs of the type (see Karatzas and Shreve, 1991, pp. 291,310):

$$\mathrm{d}\bar{\mu}_t = k_2(m(t) - \bar{\mu}_t)\mathrm{d}t + \sigma_2\sqrt{|\bar{\mu}_t - m^*(t)|}\mathrm{d}W_t^{(2)}, \tag{E.1}$$

We can therefore use a comparison theorem argument to compare any solution to (E.1) with the trivial solution $\bar{\mu} = m^*$ to the following SDE:

$$\mathrm{d}\bar{\mu}_t = k_2(m^*(t) - \bar{\mu}_t)\mathrm{d}t + \sigma_2\sqrt{|\bar{\mu}_t - m^*(t)|}\mathrm{d}W_t^{(2)}$$

Specifically, we have that imposing $m \geq m^*$ and $\bar{\mu}_0 \geq m^*(0)$ is sufficient to ensure that $\bar{\mu} \geq m^*$ a.s., thus justifying the SDE (27) (e.g Karatzas and Shreve, 1991, p. 293).

We can now focus on the SDE (26) describing the dynamics of μ. Since $\bar{\mu}$ is a measurable adapted process and $\int_0^t \bar{\mu}_u \mathrm{d}u < \infty$ for all $t \in \mathbb{R}_+$, we have that, if $\bar{\mu} \geq 0$ and $\mu_0 \geq 0$, then μ is well-defined and $\mu \geq 0$ (see Deelstra and Delbaen, 1998). We can then take $m \geq m^* \geq 0$, $\bar{\mu}_0 \geq m^*(0)$ and $\mu_0 \geq 0$, thus ensuring that Y is well-defined and that the intensity of mortality μ never becomes negative almost surely. We note that no requirement need to be made on the correlation between μ and $\bar{\mu}$.

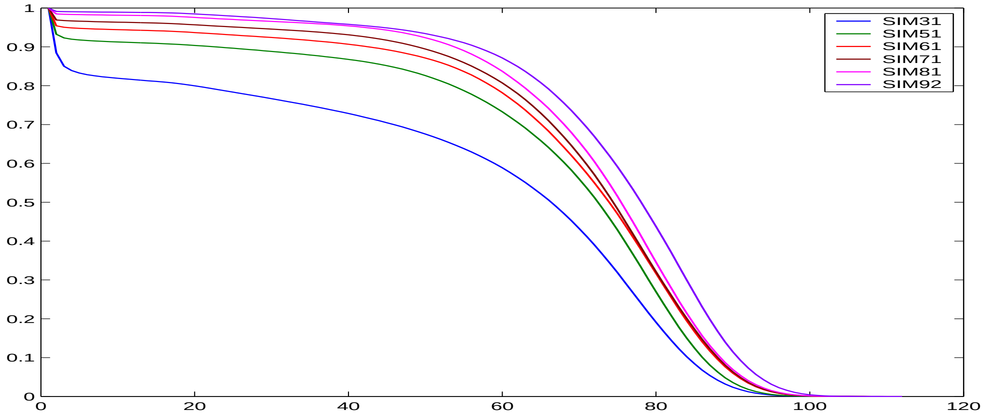

Figure 1: Joint effect of rectangularization and expansion on the survival function.

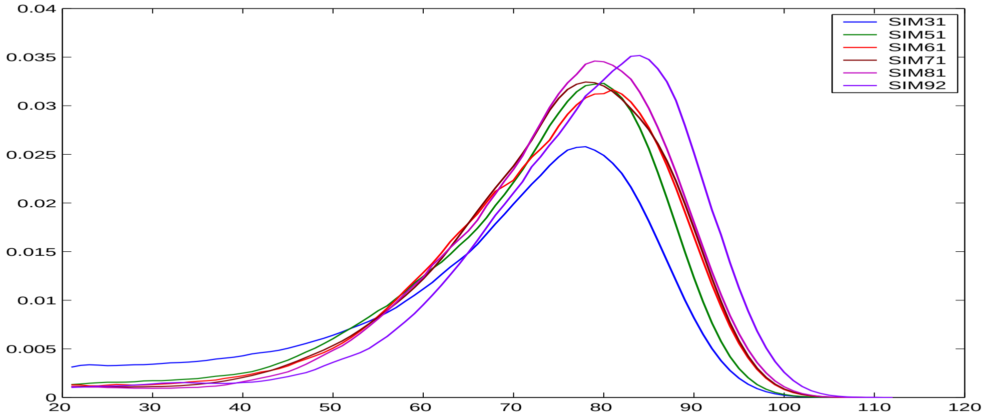

Figure 2: Joint effect of rectangularization and expansion on the curve of deaths.

Projection	$m(t)$	η_1	γ	$\bar{y}(t)$	$\bar{y}(x^* - x)$	σ	j	k	Y_0
A	SIM92	0.5	0.6	-0.005000	-0.0050	0.100	0.100	0.100	-0.016
B	SIM92	0.6	0.6	-0.000900	-0.0009	0.050	0.050	0.400	-0.016
C	Coh48	0.6	0.8	-0.001400 t	-0.0700	0.050	0.005	0.100	-0.009
D	Coh48	0.6	0.6	-0.001800 t	-0.0900	0.075	0.020	0.200	-0.009
E	Coh48	0.7	0.6	-0.000521 t^2	-1.2500	0.070	0.020	0.250	-0.009

Table 1: Parameter values for the Poisson-Gaussian model.

Life Table	$\overset{\circ}{e}_{65}$	H_{65}	Δ	δ	$\bar{Y}_{\infty}$
SIM92	14.639	0.462	-	-	-
Coh48	17.799	0.365	-	-	-
RG48	19.128	0.333	-	-	-
A	16.871	0.406	-0.001400	-33.00%	-0.0108
B	18.215	0.356	-0.002200	-52.87%	-0.0205
C	19.020	0.354	-0.000346	-18.81%	-0.0567
D	20.955	0.298	-0.000895	-48.61%	-0.0565
E	20.221	0.306	-0.000687	-37.32%	-0.7075

Table 2: Demographic measures for life tables and and parameter choices of Tab. 1.

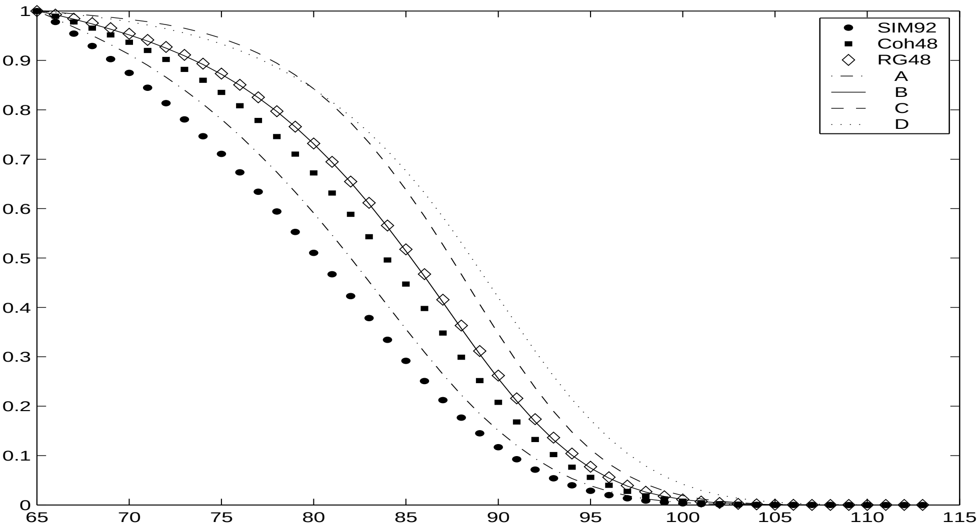

Figure 3: Comparison of the survival functions relative to the base tables SIM92, COH48 and RG48 with those relative to the parameter choices of Tab. 1.

m		$m(0)$		m^*	
c	10.841	c_0	11	c^*	13
θ	86.165	θ_0	90	θ^*	100
$\mathring{e}_{65}$	18.345	$\mathring{e}_{65}$	21.704	$\mathring{e}_{65}$	31.130
k_1	0.6	σ_1	0.15	σ_2	0.35

Table 3: Parameter values for the bidimensional square-root diffusion intensity model.

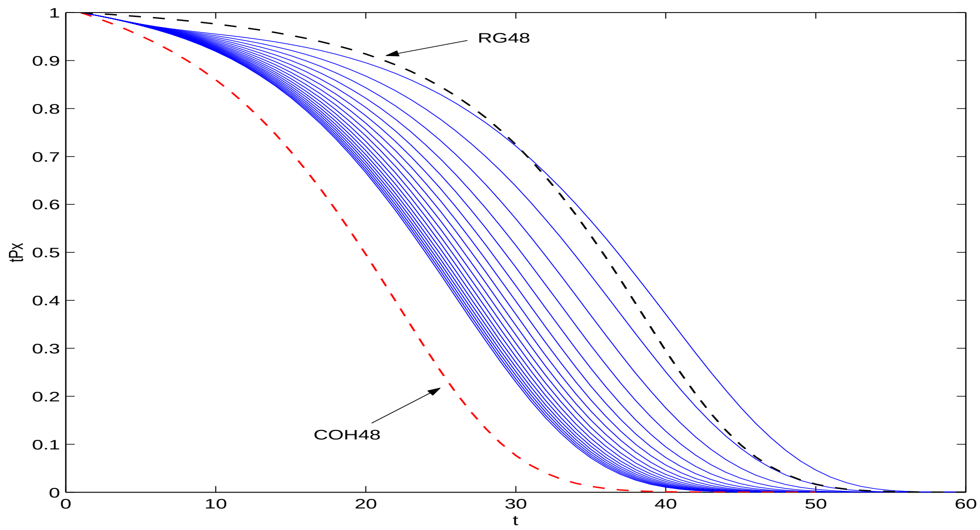

Figure 4: Survival functions obtained with γ_2 ranging from 0.50 to 0.05. All other parameters are as in Tab. 3.

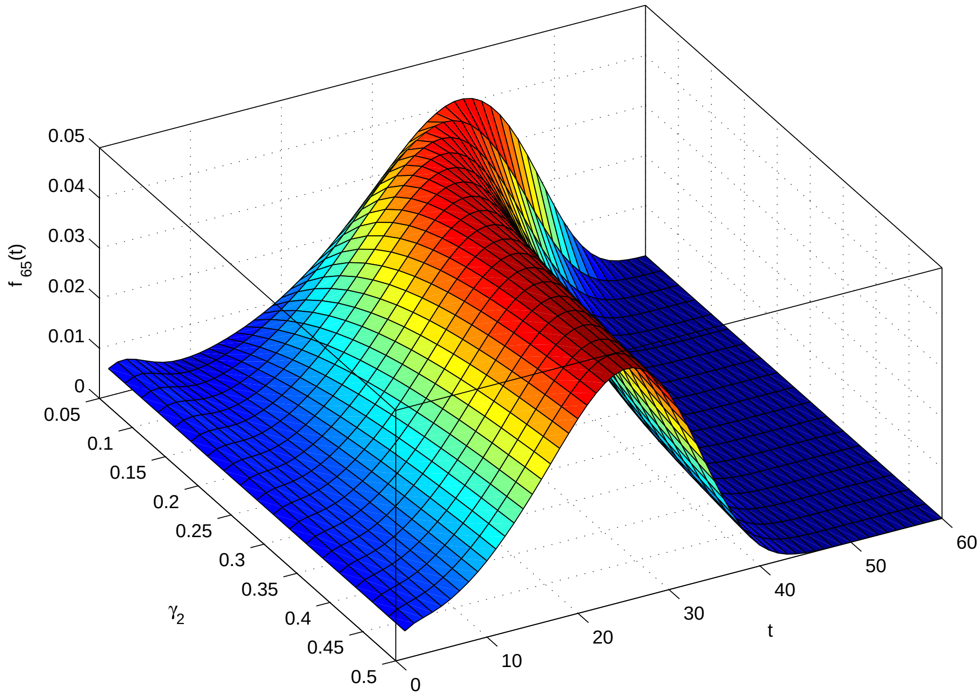

Figure 5: Density surface obtained with γ_2 ranging from 0.50 to 0.05. All other parameters are as in Tab. 3.