Università "LUISS-Guido Carli" di Roma
Facoltà di Economia

Dottorato di Ricerca in
Metodi matematici per l'economia, l'azienda,
la finanza e le assicurazioni
XXII ciclo

An optimal Markovian consumption-investment problem in a market with longevity bonds

Supervisore
Giovanna Nappo

Dottorando
Vanessa Raso

Contents

Chapter 1

Market, portfolio and arbitrage

1.1 Introduction

In this chapter the aim is focused on the mathematical modelling of financial markets. Any financial product which is traded in the market is referred to as an asset. We consider a financial market consisting of $N+1$ financial assets. One of these is instantaneously riskless, and will be called a money market account. Assets 1 through N are different assets such as stocks, bonds with different maturities, or various kinds of financial derivatives. In the following we will give a mathematical definition of basic financial concepts.

We refer to Øksendal [19], Karatzas and Shreve [16], and Björk [2] for the basic notions in stochastic differential theory, the general results in stochastic calculus, and the arbitrage theory in continuous time, respectively.

1.2 Market theory

Definition 1.2.1 (Market Place). *A market place of duration T is a complete probability space $(\Omega, \mathcal{F}, P)$ endowed with a filtration $\mathbb{F} = \{\mathcal{F}_t : t \in [0,T]\}$, such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. We will shortly write $(\Omega, \mathcal{F}, \mathbb{F}, P)$.*

It is clear that the filtration represents the information generated by all observed events up to time t, the information available at time t. In continuous-time, it is often convenient to impose further conditions on the filtration $\mathbb{F}$, i.e., the filtration $\mathbb{F}$ is right-continuous and $\mathcal{F}_0$ contains all the P-negligible sets in $\mathcal{F}$. We will shortly say that $\mathbb{F}$ satisfying the usual conditions[1].

Definition 1.2.2 (Price Process). *A price of an asset is a stochastic process $X = (X(t), t \in [0,T])$, adapted[2] to the filtration $\mathbb{F}$, (shortly X is $\mathbb{F}$-adapted), and such that[3] $X(\omega) \in \mathcal{L}^p(0,T;\mathbb{F})$ for some $p \in [1,\infty]$.*

Definition 1.2.3 (Market Model). *A (finite-dimensional) market $\mathcal{M}(X)$ is a couple composed by a market place and a $N+1$-dimensional process[4] $X = (X_0, \dots, X_N)'$ of assets' prices.*

In this framework the traded assets on the market are stocks (or primary assets) and derivative assets.

[1] Generally, we say that a filtration $\mathbb{S} = \{\mathcal{S}_t : t \in [0,T]\}$ satisfies the usual conditions if $\mathcal{S}_0$ contains all the P-negligible sets in $\mathcal{F}$ and the filtration $\mathbb{S}$ is right-continuous, i.e.

$$\mathcal{S}_{t+} = \mathcal{S}_t, \quad \forall t \geq 0, \tag{1.1}$$

where $\mathcal{S}_{t+} = \bigcap_{\epsilon > 0} \mathcal{S}_{t+\epsilon}$.

[2] Equivalently $X(t)$ is observable on $\mathbb{F}$.

[3] We say that a process $X \in \mathcal{L}^p(a,b;\mathbb{F})$ if $X(t)$ is $\mathbb{F}$-adapted and

$$\begin{aligned} E\left[\int_a^b |X(s)|^p ds\right] &< \infty, \quad \text{for } p \in [1,\infty), \\ E\left[\sup_{s\in[a,b]} |X(s)|\right] &< \infty, \quad \text{for } p = \infty. \end{aligned}$$

Moreover if X is a positive valued process, we will write $X \in \mathcal{L}^p_+(a,b;\mathbb{F})$.

[4] The prime denotes transposition, so that $X(t)$ is a column vectors.

Generally, in the Definition 1.2.3, $X_0 \in \mathcal{L}^\infty(0,T;\mathbb{F})$ represents the money market account and

$$\tilde{X}_i = \frac{X_i}{X_0}, \quad i = 1,\dots,N \tag{1.2}$$

are the discounted prices.

In the sequel we assume that $X_0 = G$, where G as an adapted process of finite variation and with continuous sample paths. For almost all $\omega \in \Omega$, the function $G(t) = G(t,\omega)$ solves the following differential equation

$$dG(t) = r(t)G(t)dt, \tag{1.3}$$

with the conventional initial condition $G(0) = 1$ and where r is an adapted process. Then $G(t)$ is given by the formula

$$G(t) = \exp\left(\int_0^t r(u)du\right). \tag{1.4}$$

In financial interpretation, G represents the price process of a riskless asset whose interest rate at time t is $r(t)$. In another usual interpretation, G represents a model of a bank account at the interest rate r. In the sequel, the process G is referred to as the money market account (also accumulation factor) while $r(t)$ is referred to as the riskless interest rate (also short interest rate or spot interest rate) at time t, accordingly to the following definition.

Definition 1.2.4 (Money Market Account). *A money market account or accumulation factor is a riskless asset whose price process $G \in \mathcal{L}^\infty(0,T;\mathbb{F})$ follow the dynamics (1.3), where $r \in \mathcal{L}^1_+(0,T;\mathbb{F})$ is an $\mathbb{F}$-adapted process and represents the riskless interest rate.*

We now make the following assumptions.

Assumption 1.2.1. *We assume that on $(\Omega,\mathcal{F},P)$ there exists an M-dimensional Wiener process $W = (W_1,\dots,W_M)$, where all the $(W_i(t), t\in[0,T])^{'}$ are independent Wiener processes.*

We observe that the filtration generated by W, $\mathbb{F}^W$, does not satisfy the usual conditions. However, if we replace $\mathcal{F}^W_t$ by $\bar{\mathcal{F}}^W_t = \sigma(\mathcal{F}^W_t \bigcup \mathcal{N})$, ($\mathcal{N}$ is the σ-algebra generated by all the P-negligible sets of $\mathcal{F}$) we obtain a proper filtration, denoted by $\bar{\mathbb{F}}^W$, satisfying the desired conditions (see Section 2.7 of Karatzas and Shreve [16]). We call it the augmented filtration associated to the process W.

Assumption 1.2.2. *We assume that on $(\Omega,\mathcal{F},P)$ the filtration $\mathbb{F}$ is the augmented filtration associated to the process $(W(t), t\geq 0)$, i.e. $\mathbb{F} = \bar{\mathbb{F}}^W$. When we talk about martingale or adapted process without mentioning any filtration, it is assumed that we are dealing with the filtration $\bar{\mathbb{F}}^W$.*

In order to avoid some technical difficulties an augmented filtration is necessary. As an example, let X be an $\mathbb{F}$-adapted process and Y be a process such that $X(t) = Y(t)$ a.s. for every t. In general this does not imply that Y is also an $\mathbb{F}$-adapted process. In fact, the negligible event $\mathcal{N}_t = \{X(t) \neq Y(t)\}$ may not belong to $\mathcal{F}_t$ and therefore $Y(t)$ might not be $\mathcal{F}_t$-measurable. This problem cannot appear if $\mathcal{F}_0$ contains all the P-negligible sets in $\mathcal{F}$. In this case, moreover, every a.s. continuous process has a continuous modification. Also the fact that the filtration is right-continuous is a technical assumption that is often necessary; therefore we shall take care, whenever possible, to prove that our processes of interest are defined on a probability space endowed with a augmented filtration.

Assumption 1.2.3. *For $i = 1,\dots,N$, we assume that* [5]

$$X_i \in \mathcal{L}^2(0,T;\mathbb{F})$$

and that it satisfies a stochastic differential equation of the form

$$\begin{aligned} dX_i(t) &= \mu_i(t)dt + \sum_{j=1}^M \sigma_{ij}(t)dW_j(t) \\ &= \mu_i(t)dt + \sigma_i(t)dW(t) \end{aligned} \tag{1.5}$$

where μ_i and σ_{ij} are adapted. We have used the notation $\sigma_i = (\sigma_{i1},\dots,\sigma_{iM})$.

[5] In order to guarantee the existence of stochastic integral we have to impose some integrability condition on X_i and the class $\mathcal{L}^2(0,T;\mathbb{F})$ turns out to be a natural one.

In the sequel we will call μ_i the drift term (or mean rate of return) of X_i and σ_i the diffusion term (or volatility) of X_i. It is possible to rewrite (1.5) in the follow matrix notation

$$d\bar{X}(t) = \mu(t)dt + \sigma(t)dW(t) \tag{1.6}$$

where

$$\bar{X}(t) = \begin{pmatrix} X^1(t) \\ \vdots \\ X^N(t) \end{pmatrix}, \quad \mu(t) = \begin{pmatrix} \mu_1(t) \\ \vdots \\ \mu_N(t) \end{pmatrix}, \quad \sigma(t) = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1M} \\ \sigma_{21} & \dots & \sigma_{2M} \\ \vdots & \vdots & \vdots \\ \sigma_{N1} & \dots & \sigma_{NM} \end{pmatrix}.$$

Fixing $t_0, \bar{x}_0$, we make the following assumption.

Assumption 1.2.4. *We assume that $\mu(t)$ and $\sigma(t)$ in (1.6) are given by*

$$\mu(t) = \hat{\mu}(t, X(t)), \tag{1.7}$$
$$\sigma(t) = \hat{\sigma}(t, X(t)) \tag{1.8}$$

for some measurable deterministic functions $\hat{\mu}$ and $\hat{\sigma}$. Furthermore, given $\bar{X}(t_0) = \bar{x}_0$, equation (1.6) admits a unique solution. We will denote the unique solution by $X(t) = X(t; t_0, \bar{x}_0)$, $t \geq t_0$.

Recall that the Theorem B.1.1 gives sufficient conditions on $\hat{\mu}$ and $\hat{\sigma}$ to guarantee existence and uniqueness of the solution.

Finally, we observe that the property of G, being a riskless asset, is characterized by the absence of the driving dW-term, while a risk asset is characterized by the presence of a diffusion.

1.3 Self-financing portfolio

In this section the aim is to derive the dynamics of the so called self-financing portfolio. Thus we have the following definitions.

Definition 1.3.1 (Portfolio). *Let the $N+1$-dimensional price process $X = \left(G, \bar{X}\right)$ be given. A portfolio strategy (or simply portfolio) is any $\bar{\mathbb{F}}^X$-adapted $N+1$-dimensional process*

$$h = (h_0, h_1, \dots, h_N),$$

where the component $h_i(t)$ is the number of shares of the i^{th} asset held by the trader at time t and $\bar{\mathbb{F}}^X$ is the augmented filtration associated to the process X.

Definition 1.3.2 (Value Process). *The value process V^h corresponding to the portfolio h is given by*

$$V^h(t) = h(t) \cdot X(t) = \sum_{i=0}^{N} h_i(t) X_i(t). \tag{1.9}$$

A self-financing portfolio is a portfolio with no exogenous infusion or withdrawal of money, in other words the purchase of a new portfolio must be financed only by selling assets already in the portfolio.

In discrete time case, i.e. when $t \in \{t_0, t_1, \dots, t_M\}$ with $t_0 = 0$ and $t_M = T$, an self-financing portfolio $h(t_n) = h_n$ is a portfolio such that

$$h_n \cdot X_n = h_{n+1} \cdot X_n, \quad n = 0, \dots, M, \tag{1.10}$$

where $X_n = X(t_n)$. Adding and subtracting $h_{n+1} \cdot X_{n+1}$ to the left hand side into (1.10), we obtain

$$h_n \cdot X_n + h_{n+1} \cdot X_{n+1} - h_{n+1} \cdot X_{n+1} = h_{n+1} \cdot X_n,$$

and grouping we have

$$h_{n+1} \cdot X_{n+1} - h_n \cdot X_n = h_{n+1} \cdot (X_{n+1} - X_n). \tag{1.11}$$

Analogously to the Definition 1.3.2, we define the value process V^h by

$$V_n^h = h_n \cdot X_n, \tag{1.12}$$

and substituting (1.12) into (1.11) we have

$$V^h_{n+1} - V^h_n = h_{n+1} \cdot (X_{n+1} - X_n). \tag{1.13}$$

If we now consider our continuous time model as a limit of the above discrete time model, as $t_n = t_{n+1} - t_n$ goes to 0, then (with the Itô interpretation of the integral[6]) we obtain the following definition.

Definition 1.3.3 (Self-financing Portfolio). *A portfolio h is self-financing if the value process V^h satisfies the condition*

$$dV^h(t) = h(t) \cdot dX(t) = \sum_{i=0}^{N} h_i(t) dX_i(t). \tag{1.14}$$

or equivalently, for $t_1 < t_2$

$$V^h(t_2) - V^h(t_1) = \int_{t_1}^{t_2} h(t) \cdot dX(t). \tag{1.15}$$

Here the stochastic differential is intended in the Itô's sense[7].

Note that instead of specifying the absolute number of shares held of a certain asset, it may be convenient to specify the relative proportion of the total portfolio value which is invested in the asset. Thus we have the following definition.

Definition 1.3.4 (Relative Portfolio). *For a given portfolio h the corresponding relative portfolio U is given by*

$$U_i(t) = \mathbf{1}_{\{X_i(t)>0\}} \frac{h_i(t) X_i(t)}{V^h(t)} = \mathbf{1}_{\{X_i(t)>0\}} u_i(t), \quad i = 0, 1, \ldots, N. \tag{1.16}$$

where

$$\sum_{i=0}^{N} u_i(t) = 1. \tag{1.17}$$

In terms of the relative portfolio, the dynamics of a self-financing portfolio can be expressed with the following lemma.

Lemma 1.3.1. *A portfolio h is self-financing if and only if*

$$dV^h(t) = V^h(t) \sum_{i=0}^{N} U_i(t) \frac{dX_i(t)}{X_i(t)} = V^h(t) \sum_{i=0}^{N} \mathbf{1}_{\{X_i(t)>0\}} u_i(t) \frac{dX_i(t)}{X_i(t)}, \tag{1.18}$$

where U is the relative portfolio corresponding to h.

Observe that, in (1.18), we have $\mathbf{1}_{\{X_i(t)>0\}}$ in order to consider the case $X_i(t) = 0$, too.

Proof. Equation (1.18) follows immediately from Definitions 1.3.4 and 1.3.3.

□

So far we have considered a situation without any consumption, but if now we consider a situation with some consumption, we have the concepts and results similar to the ones of the previous situation. Taking into account Definitions 1.3.1 1.3.2, we have the following definitions.

Definition 1.3.5 (Consumption Process). *Let the $N+1$-dimensional price process $X = \left(G, \bar{X}\right)$ be given. A consumption process is any $\mathbb{F}^X$-adapted 1-dimensional process $k(t)$, with $t \geq 0$.*

Now we extends the self-financing concept to this setting with the following definition.

Definition 1.3.6 (A Self-financing Portfolio-Consumption). *A portfolio-consumption pair, denoted by (h, k), is called self-financing if the value process V^h satisfies the condition*

$$dV^h(t) = h(t) \cdot dX(t) - k(t)dt = \sum_{i=0}^{N} h_i(t) dX_i(t) - k(t)dt. \tag{1.19}$$

[6] It is important that the increment $\Delta X(t) = X(t_{n+1}) - X(t_n)$ is a forward increment.

[7] We are implicitly assuming the integrability conditions assuring that the stochastic integrals on the right hand side of (1.15) are defined.

or equivalently, for $t_1 < t_2$

$$V^h(t_2) - V^h(t_1) = \int_{t_1}^{t_2} h(t) \cdot dX(t) - \int_{t_1}^{t_2} k(t)dt. \tag{1.20}$$

Here the stochastic differential is intended in the Itô's sense[8].

Observe that the self-financing pairs (h, k) are simply portfolios with no exogenous infusion or withdrawal of money, apart of course from the k-term. In others words, the purchase of a new portfolio, as well as all consumption, must be financed solely by selling assets already in the portfolio.

Finally, we extend Lemma 1.3.1 to this setting with the following lemma.

Lemma 1.3.2. *A portfolio-consumption pair (h, k) is self-financing if and only if*

$$dV^h(t) = V^h(t) \sum_{i=0}^{N} U_i(t) \frac{dX_i(t)}{X_i(t)} - k(t)dt = V^h(t) \sum_{i=0}^{N} \mathbf{1}_{\{X_i(t)>0\}} u_i(t) \frac{dX_i(t)}{X_i(t)} - k(t)dt, \tag{1.21}$$

where U is the relative portfolio corresponding to h.

We now make a further extension of the self-financing concept to a market with, besides the riskless rate $r(t)$, a cost factor associated to the asset. The difference between the present situation (with the cost factor) and the case without the cost factor is that the budget equation (1.10) now has to be modified. Let the N-dimensional cost process $\bar{D}$, where $\bar{D}_i$ denotes the cumulative cost associated to the ith asset, the relevant budget equation is given by

$$h_n \cdot X_n = h_{n+1} \cdot \left(X_n - \bar{D}_n \right), \quad n = 0, 1, \ldots, M, \tag{1.22}$$

where $\bar{D}_n = \bar{D}(t_n)$. Going through the same arguments as above, we end up with the following dynamics for a self-financing portfolio-consumption (h, k)

$$dV^h(t) = \sum_{i=0}^{N} h_i(t) \left(dX_i(t) - d\bar{D}_i(t) \right) - k(t)dt. \tag{1.23}$$

Finally, we extend Lemma 1.3.2 to this setting so that in terms of the relative portfolio, the dynamics of a self-financing portfolio can be expressed as

$$\begin{aligned} dV^h(t) &= V^h(t) \left(\sum_{i=0}^{N} U_i(t) dX_i(t) - dD_i(t) \right) - k(t)dt \\ &= V^h(t) \left(\sum_{i=0}^{N} \mathbf{1}_{\{X_i(t)>0\}} u_i(t) \frac{dX_i(t) - dD_i(t)}{X_i(t)} \right) - k(t)dt. \end{aligned} \tag{1.24}$$

1.4 Financial derivatives, completeness and arbitrage

Financial derivatives are completely defined in terms of some underlying assets already existing on the market. These financial instruments have been created to manage with the risk. They are called derivatives as their evolution depends on the evolution of some primary assets of the market. We will now give the formal definition of a particular derivative: European contingent claim.

Definition 1.4.1 (European Contingent Claim)**.** *An European contingent claim, with date of maturity (exercise date) $S \le T$, (shortly S-claim), on the underlying assets $X_1, \ldots, X_N$ is a random variable $\mathcal{X} \in \mathcal{F}_S^X$. The random variable $\mathcal{X}$ is also called payoff.*

The interpretation of the above definition is that an European contingent claim is a financial contract between two parties, the seller (writer) and the buyer (owner) of the contract. The requirement that $\mathcal{X} \in \mathcal{F}_S^X$ simply means that, at time S, it will actually be possible to determine the payoff.

Generally, there are two main problems concerning derivatives: pricing and hedging. The first problem consists in finding, if it exists, a fair price for a derivative, while the second problem regards the possibility for the writer to minimize the risk associated to the derivative.

Two fundamental concepts in financial theory are the absence of arbitrage and completeness.

[8] We are implicitly assuming the integrability conditions assuring that the stochastic integrals on the right hand side of (1.20) are defined.

Definition 1.4.2 (Arbitrage). *An arbitrage opportunity (shortly arbitrage) on a financial market is a self-financing portfolio h such that*

$$V^h(0) = 0, \quad P(V^h(T) \geq 0) = 1, \quad \text{with} \quad P(V^h(T) > 0) > 0. \tag{1.25}$$

Definition 1.4.3 (Arbitrage-free Market). *A financial market is arbitrage-free if for any self-financing portfolio h such that*

$$V^h(0) = 0 \quad \text{and} \quad P(V^h(T) \geq 0) = 1 \quad \text{imply} \quad P(V^h(T) = 0) = 1. \tag{1.26}$$

An arbitrage is thus equivalent to the possibility to make a profit without any risk of losing money. A market is called efficient, if it is arbitrage-free. The following result shows that, in an efficient market, if a portfolio has a value process whose dynamics contain no driving Wiener process, i.e. a riskless portfolio, then the rate of return of that portfolio must equal the riskless interest rate.

Proposition 1.4.1. *Suppose that there exists a self-financing portfolio h, such that the value process V^h has the dynamics*

$$dV^h(t) = \gamma(t)V^h(t)dt, \quad t \in I \tag{1.27}$$

where q is a adapted cadlag[9] process, and I is an open (non void) time intervall. Then either $\gamma(t) = r(t)$ for all $t \in I$, or it exists an arbitrage.

Proof. For simplicity, we assume that $\gamma(t) > r(t)$ on $I = (t_0, t_1) \subseteq [0, T]$. We can borrow money from the bank and immediately we invest this money in the portfolio h. We follow this strategy on the interval (t_0, t_1), where $\gamma(t) > r(t)$. Thus the net investment at $t = t_0$ is zero, whereas our wealth at $t = t_1$ will be strictly positive. In other words, we have an arbitrage. If instead, $\gamma(t) < r(t)$, we sell the portfolio and we invest immediately this money in the bank, and again there is an arbitrage.

□

Definition 1.4.4 (Attainable Claim). *We say that an S-claim $\mathcal{X}$ is attainable, or financeable, if there exists a self financing portfolio h such that*

$$V^h(S) = \mathcal{X} \quad a.s. \tag{1.28}$$

In this case we say that h is a replicating or hedging portfolio for $\mathcal{X}$. If every contingent claim is attainable we say that the market is complete, otherwise the market is incomplete.

We will give some general results for determining whether a certain model is complete and/or arbitrage-free. These results are obtained by natural applications of martingale theory[10].

Definition 1.4.5 (Equivalent Martingala Measure). *A (probability) measure Q equivalent to P, $Q \sim P$, such that the discounted prices process $\tilde{X}$ are martingales with respect to Q, is called an equivalent martingale or a risk-neutral measure.*

Most of modern finance theory is based on the following theorems, so called first and second fundamental asset pricing theorems.

Theorem 1.4.2 ([14]). *A market is arbitrage-free if and only if there exists an equivalent martingale measure Q.*

Theorem 1.4.3 ([15]). *A market is complete if and only if there is one and only one equivalent martingale measure Q.*

Observe that in order to prove Theorem 1.4.2 and 1.4.3 it is necessary the Assumption 1.2.2. Moreover by assumption of arbitrage-free market, we have the following result, which we will find again in the next chapters.

Theorem 1.4.4 ([19]). *If a financial market[11] $\mathcal{M}(X)$ is arbitrage-free, then the market is complete if and only if the volatility matrix $\sigma(t)$ has a left inverse $\Lambda(t)$ for all t almost surely, i.e., there exist an adapted matrix valued process $\Lambda(t) \in \mathbb{R}^{M \times N}$ such that*

$$\Lambda(t)\sigma(t) = I_M \quad \forall t \ a.s. \tag{1.29}$$

[9] The cadlag process is "continus à droite avec limite à gauche", as the french say, which means right continuous with left limits

[10] The modern theory of financial derivatives is based mainly on martingale theory.

[11] Recall that N is the number of assets and M is the dimension of the underlying Wiener process.

We observe that the property (1.29) is equivalent to the property

$$\mathbf{r}\left(\sigma(t)\right) = M \quad \forall t \text{ a.s.} \tag{1.30}$$

where, for a matrix A, $\mathbf{r}(A)$ is the rank of A.

Corollary 1.4.5 ([19]). *Suppose a financial market $\mathcal{M}(X)$ arbitrage-free.*

1. *If $N = M$ then the market is complete if and only if $\sigma(t)$ is invertible for all t almost surely.*
2. *If the market is complete, then*
$$\mathbf{r}\left(\sigma(t)\right) = M \quad \forall t \ a.s. \ , \tag{1.31}$$
and in particular, $N \geq M$.

Theorem 1.4.6 ([19]). *If a financial market $\mathcal{M}(X)$ is arbitrage-free, then there exists an adapted M-dimensional process $\xi = (\xi_1, \ldots, \xi_M)^{'}$, such that for $i = 1, \ldots, N$*

$$\sum_{j=1}^{M} \sigma_{ij}(t)\xi_j(t) = \mu_i(t) - r(t) \quad \forall t \ a.s. \tag{1.32}$$

or in matrix notation

$$\sigma(t)\xi(t) = \mu(t) - r(t) \quad \forall t \ a.s. \tag{1.33}$$

Conversely, suppose that there exists an M-dimensional process $\xi \in \mathcal{L}^2\left(0, T; \mathbb{F}\right)$ that satisfies (1.33) and such that

$$E\left[\exp\left(\frac{1}{2}\int_0^T \xi(t)^2 dt\right)\right] < \infty. \tag{1.34}$$

Then the market $\mathcal{M}(X)$ is arbitrage-free.

Really (1.34) is Novikov condition which guarantees that $e^{-\int_0^t \xi(s)dW(s) - \frac{1}{2}\int_0^t \xi^2(s)ds}$ is a martingale with mean equals 1.

Before we proceed we have the following useful result, where a Wiener process with respect to Q, $\bar{W}(t)$, can be constructed from a Wiener process with respect to P, $W(t)$, via a change of measure from P to Q.

Lemma 1.4.7. *Suppose there exists an M-dimensional process $\xi \in \mathcal{L}^2(0, T; \mathbb{F})$ that satisfies (1.33). Let*

$$\mathcal{Z}(t) = e^{-\int_0^t \xi(s)dW(s) - \frac{1}{2}\int_0^t \xi^2(s)ds}. \tag{1.35}$$

We assume (1.34) and we consider the probability[12] *measure $Q = Q_T$ on $\mathcal{F}_T$ defined as*

$$dQ = \mathcal{Z}(T)dP. \tag{1.36}$$

Then

$$\bar{W}(t) := \int_0^t \xi(t)dt + W(t) \tag{1.37}$$

is a M-dimensional Wiener process with respect to Q and in terms of $\bar{W}(t)$ we have the following representation of the discounted market

$$d\tilde{X}_i(t) = \tilde{X}_i(t)\sigma_i(t)d\bar{W}(t), \quad i = 1, \ldots, N. \tag{1.38}$$

In particular, if

$$E^Q\left[\int_0^T \left(\tilde{X}_i(t)\sigma_i(t)\right)^2 dt\right] < \infty, \quad i = 1, \ldots, N \tag{1.39}$$

then Q is an equivalent martingale measure.

[12] We observe that, after the assumption (1.34), L is a martingale (Novikov condition). Then we have that

$$Q(\Omega) = \int_\Omega \mathcal{Z}(T)dP = E\left(\mathcal{Z}(T)\right) = E\left(\mathcal{Z}(0)\right) = 1,$$

i.e., Q is a probability measure.

Proof. The first statement follows from Novikov condition (1.34) and Girsanov theorem. Using (1.33), to prove the representation (1.38) we compute

$$\begin{aligned}
d\tilde{X}_i(t) &= d\left(\frac{X_i(t)}{G(t)}\right) = \frac{1}{G(t)}dX_i(t) + d\left(\frac{1}{G(t)}\right)X_i(t) \\
&= \frac{1}{G(t)}\left[\Big(\mu_i(t)X_i(t) - r(t)X_i(t)\Big)dt + X_i(t)\sigma_i(t)dW(t)\right] \\
&= \frac{1}{G(t)}\left[\Big(\mu_i(t)X_i(t) - r(t)X_i(t)\Big)dt + X_i(t)\sigma_i(t)\left(\bar{W}(t) - \xi(t)dt\right)\right] \\
&= \frac{X_i(t)}{G(t)}\sigma_i(t)d\bar{W}(t) \\
&= \tilde{X}_i(t)\sigma_i(t)d\bar{W}(t).
\end{aligned}$$

In particular, by the proprieties of Itô integral, if $E^Q\left[\int_0^T \left(\tilde{X}_i(t)\sigma_i(t)\right)^2 dt\right] < \infty$, then $\tilde{X}_i(t)$ is a martingale with respect to Q.

□

There is a natural economic interpretation of the process ξ. The right hand side of (1.33) is the risk premium of the N-dimensional price process $\bar{X}$. In the left hand side of (1.33) we have $\sigma(t)\xi(t)$, where $\sigma(t)$ is the volatility matrix of the process $\bar{X}$. On the one hand ξ is called the risk premium for unit of volatility, while on the other hand a relation similar to (1.33) appears in the CAPM's theory, so ξ is commonly called the market price of risk. Finally, from (1.32), we see that the risk premium of any asset, $\mu_i(t) - r(t)$, can be written as a linear combination of the volatility components σ_i of the asset. The important point is that the multipliers $\xi_1, \ldots, \xi_N$ are the same for all assets.

In conclusion, in an arbitrage-free market, regardless of whether the market is complete or incomplete, there exists a market price of risk process, ξ, which is common to all assets in the market and satisfies the system of equations (1.33).

More in general, system (1.33) can be used to characterize complete and/or arbitrage-free markets, considering the following three cases.

1. (Unique solution). System (1.33) has a unique solution $\xi(t)$. If (1.34) holds, then from Lemma 1.4.7, we define a unique martingale measure Q, i.e., the market is arbitrage-free and complete.

2. (No solution). System (1.33) has no solution, then there is no martingale measure and the market admits arbitrage.

3. (Multiple solution). System (1.33) has multiple solutions. If (1.34) holds, then there are multiple martingale measures. The market is arbitrage-free, but there are contingent claims which cannot be hedged, i.e., the market is incomplete.

Chapter 2

Short rate models

2.1 Introduction

Most traditional stochastic interest rate models are based on the specification of a riskless interest rate. As we have seen in Chapter 1 (see Definition 1.2.4), the riskless interest rate r is modelled as an adapted process defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ under Assumptions 1.2.1 and 1.2.2, so that $\mathbb{F} = \bar{\mathbb{F}}^{W^r}$. A natural starting point is to give an a priori specification of the dynamics of r. We examine the general case of a riskless interest rate which follows an Itô process under the probability measure P, so we model r as the solution of a stochastic differential equation of the form

$$dr(t) = \mu^r(t)dt + \sigma^r(t)dW^r(t) \tag{2.1}$$

where W^r is a P-Wiener process, $\mu^r(t)$ and $\sigma^r(t)$ are adapted[1]. This model is completely general, subject only to the condition that the paths of the process are integrable with respect to the Lebesgue measure.

As in Section 1.2 (see Assumption 1.2.4), we assume that

$$\mu^r(t) = \hat{\mu}^r(t, r(t)), \tag{2.2}$$

$$\sigma^r(t) = \hat{\sigma}^r(t, r(t)), \tag{2.3}$$

for some measurable deterministic functions $\hat{\mu}^r$ and $\hat{\sigma}^r$, and that there exists a unique solution of equation (2.1). Fixing $r(s) = \bar{r}$, we will denote the unique solution of (2.1) by $r(t) = r(t; s, \bar{r})$, $t \geq s$. We observe that, under the previous assumptions, the $r(t; s, \bar{r})$ is a Itô diffusion[2], which satisfies the important Markov property with respect to the filtration $\mathbb{F}$ (shortly $r(t; s, \bar{r})$ is an $\mathcal{F}_t$-Markov process), i.e., for all Borel measurable, bounded functions f, we have

$$E\left[f(r(t')|\mathcal{F}_t\right] = E\left[f(r(t'))|r(t)\right] = g\left(r(t)\right) \tag{2.4}$$

for fixed t, t' such that $s \leq t \leq t'$ with $g(x) := E\left[f(r(t; s, x))\right]$. In the sequel, for simplicity, we use the following notation[3]

$$E\left[f(r(t; s, \bar{r}))\right] = E_{s,\bar{r}}\left[f(r(t))\right], \tag{2.5}$$

so that the function $g(x)$ can be written as

$$g(x) = E_{s,x}\left[f(r(t))\right]. \tag{2.6}$$

In the market described in Chapter 1, we assume that the riskless interest rate is the only object given a priori, so that the only exogenously given asset is the money account with price process G (see Definition 1.2.4). Let us formulate this as a formalized assumption.

Assumption 2.1.1. *We assume the existence of one exogenously given (riskless) asset. The price, G, of this asset has dynamics given by equation (1.3), where the dynamics of r, under the probability measure P, are given by equation (2.1).*

[1] We assume implicitely the integrability conditions that are necessary to define the right hand side of (2.1).

[2] We refer to Øksendal [19] for the stochastic differential theory

[3] More in general we will use the same kind of notation for functionals of the trajectory $r(t; s, \bar{r})$, $t \geq s$

2.2 Zero coupon bonds: the term structure equation

In this section the aim is focused on the problem of modelling an arbitrage free family of zero coupon bond price processes, and we follow the approach of Björk [2].

By a zero coupon bond (a discount bond) of maturity T we mean a financial security paying to its holder one unit of cash at a fixed date T in the future. Formally, we have the following definition.

Definition 2.2.1 (Zero Coupon Bond). *A zero coupon bond with maturity date T, also called a T-bond, is a contract which guarantees the holder 1 dollar to be paid on the date T. The (random) price at time t of a bond with maturity T is denoted by $B(t,T)$.*

The convention that the payment at the time of maturity, known as the principal value or face value, equals one is made for computational convenience.

Let us first describe briefly the set of general assumptions imposed on our financial market models.

Assumption 2.2.1. *We assume that there exists a market for zero coupon T-bonds for every value of T.*

We thus assume that our market contains all possible bonds (plus, of course, the riskless asset). Consequently it is market containing an infinite numbers of assets, but we again stress the fact that only the riskless asset is exogenously given. In other words, in this model, the riskless asset is considered as the underlying asset whereas all bonds are regarded as derivatives of the "underlying" short rate r, i.e., a zero coupon bond can be thought of as a derivative on the interest rate.

Assumption 2.2.2. *We assume that there is a market for T-bonds for every choice of T and that the market is arbitrage free. We assume furthermore that, for every T, the price of a T-bond has the form*

$$B(t,T) = \hat{B}^T(t, r(t)) \tag{2.7}$$

where $\hat{B}^T$ is a deterministic function of two[4] *real variables. We assume that $\hat{B}^T$ is smooth and strictly positive.*

The aim now is to find the arbitrage free price process of a T-bond, $B(t,T)$. The price of a particular bond is not be completely determined by the specification (2.1) of the r-dynamics and the requirement that the bond market is free arbitrage. To understand the reason why this problem arises, we consider that the arbitrage pricing is pricing a derivative in terms of some underlying assets' prices. In our market we do not have a sufficient number of underlying assets. We thus fail to determine a unique price of a particular bond.

Fortunately this fact does not mean that bond prices can take any form. On the contrary the bond prices with different maturities will have to satisfy certain internal consistency relations in order to avoid arbitrage on the bond market. If we take the price of one particular bond (called benchmark bond) as given, then the prices of all other bonds will be uniquely determined in terms of the price of the benchmark bond (and the r-dynamics), as we will see in Remark 2.2.1. In our market model (see Assumption 2.2.2 and in particular (2.7)) this fact is in complete agreement with the Corollary 1.4.5, since in the a priori given market consisting of one benchmark bond plus the risk free asset we will have $N = M = 1$, thus guaranteeing completeness[5].

Recall that, by Definition 2.2.1, a zero coupon bond pays one unit of cash at a prescribed date T in the future: it is thus clear that, necessarily, we have a simple boundary condition

$$\hat{B}^T(T, r) = 1 \quad \forall r. \tag{2.8}$$

where r is a real variable and denotes a generic outcome of the process $r(t)$.

From Assumption 2.2.2 and the Itô formula we have the price dynamics of the following form for T-bond

$$\frac{dB(t,T)}{B(t,T)} = \mu(t,T)dt + \sigma(t,T)dW^r(t) \tag{2.9}$$

where

$$\begin{aligned} \mu(t,T) &= \hat{\mu}^T(t, r(t)), & (2.10)\\ \sigma(t,T) &= \hat{\sigma}^T(t, r(t)), & (2.11) \end{aligned}$$

for suitable deterministic function $\hat{\mu}$ and $\hat{\sigma}$. The functions $\hat{\mu}^T$ and $\hat{\sigma}^T$ can be expressed by mean of the function $\hat{B}^T$ as shown in the following lemma.

[4] It is convenient to consider $\hat{B}^T$ as a function of only two variables, namely t and r, whereas T is regarded as a parameter.

[5] We are implicitly assuming that the volatility of the benchmark bond is not zero

Lemma 2.2.1. *Under Assumption 2.2.2, the following equalities hold with probability* 1*, for all* t *and for every choice of maturity time* T*.*

$$\hat{\mu}^T(t,r(t)) = \frac{\hat{B}_t^T(t,r(t)) + \hat{B}_r^T(t,r(t))\hat{\mu}^r(t,r(t)) + \frac{1}{2}\hat{B}_{rr}^T(t,r(t))\big(\hat{\sigma}^r(t,r(t))\big)^2}{\hat{B}^T(t,r(t))}, \tag{2.12}$$

$$\hat{\sigma}^T(t,r(t)) = \frac{\hat{B}_r^T(t,r(t))\hat{\sigma}^r(t,r(t))}{\hat{B}^T(t,r(t))}, \tag{2.13}$$

where $\hat{\mu}^r$, $\hat{\sigma}^r$ *are the functions in (2.2),(2.3) respectively, and, where we have used the notation*

$$\hat{B}_t^T(t,r) = \frac{\partial \hat{B}^T}{\partial t}(t,r), \quad \hat{B}_r^T(t,r) = \frac{\partial \hat{B}^T}{\partial r}(t,r), \quad \hat{B}_{rr}^T(t,r) = \frac{\partial^2 \hat{B}^T}{\partial r^2}(t,r). \tag{2.14}$$

In the sequel, when it is convenient, we will use the above notation (2.14).

Proof. (Lemma 2.2.1). The proof of (2.12) and (2.13) follows by observing that

$$d\hat{B}^T(t,r(t)) = \hat{B}_t^T(t,r(t))dt + \hat{B}_r^T(t,r(t))dr(t) + \frac{1}{2}\hat{B}_{rr}^T(t,r(t))\big(\hat{\sigma}^r(t,r(t))\big)^2 dt, \tag{2.15}$$

and inserting the differential form (2.1) of dr into (2.15), we obtain

$$\begin{aligned} \frac{d\hat{B}^T(t,r(t))}{\hat{B}^T(t,r(t))} &= \frac{1}{\hat{B}^T}\hat{B}_t^T dt + \frac{1}{\hat{B}^T}\hat{B}_r^T\hat{\mu}^r dt + \frac{1}{2}\frac{1}{\hat{B}^T}\hat{B}_{rr}^T(\hat{\sigma}^r)^2 dt + \frac{1}{\hat{B}}\hat{B}_r^T\hat{\sigma}^r dW^r(t) \\ &= \frac{1}{\hat{B}^T}(\hat{B}_t^T + \hat{B}_r^T\hat{\mu}^r + \frac{1}{2}\hat{B}_{rr}^T(\hat{\sigma}^r)^2)dt + \frac{1}{\hat{B}^T}\hat{B}_r^T\hat{\sigma}^r dW^r(t) \\ &= \hat{\mu}^T(t,r(t))dt + \hat{\sigma}^T(t,r(t))dW^r(t), \end{aligned} \tag{2.16}$$

where for the notational convenience, the argument $(t,r(t))$ "has been suppressed", so that we have used the shorthand notation of the form

$$\hat{\mu}^r = \hat{\mu}^r(t,r(t)), \quad \hat{\sigma}^r = \hat{\sigma}^r(t,r(t)), \tag{2.17}$$

for the process $r(t)$, and

$$\hat{\mu}^T = \hat{\mu}^T(t,r(t)), \quad \hat{\sigma}^T = \hat{\sigma}^T(t,r(t)), \quad \hat{B}^T = \hat{B}^T(t,r(t)), \tag{2.18}$$

for the process $B(t,T)$, and similarly for the partial derivatives terms. Finally equation (2.16) does not depend on T, thus it must hold with probability 1, for all t and for every choice of maturity time T.

□

Accordingly to Theorem 1.4.6, we have the following central result.

Proposition 2.2.2. *Assume that the bond market is arbitrage free. Then there exists a process* ξ_r *such that the relation*

$$\frac{\mu(t,T) - r(t)}{\sigma(t,T)} = \xi_r(t) \tag{2.19}$$

holds, with probability 1*, for all* t *and for every choice of maturity time* T*.*

By (2.10) and (2.11), we observe that $\xi_r(t)$ can be expressed as a deterministic function of t and $r(t)$, namely

$$\xi_r(t) = \hat{\xi}_r(t,r(t)), \tag{2.20}$$

and (2.19) becomes

$$\frac{\hat{\mu}^T(t,r(t)) - r(t)}{\hat{\sigma}^T(t,r(t))} = \hat{\xi}_r(t,r(t)), \quad \forall t, a.s. \tag{2.21}$$

Proof. (Proposition 2.2.2). We fix two times of maturity T and S, in order to form a portfolio $(h_T(t), h_S(t))$ consisting only of bonds having different times of maturity $B(t,T)$ and $B(t,S)$ respectively, (i.e., in this setting nothing will be invested in the bank or loaned by the bank), and we choose the weights so as to make a riskless portfolio. From general results of Section 1.3, let $h(t) = (h_0, h_1(t), h_2(t))$ be the portfolio associated to $X = (X_0, X_1, X_2)$, where $X_1 = B(t,T)$, $X_2 = B(t,S)$, $h_1(t) = h_T(t)$, $h_2(t) = h_S(t)$ and $h_0 = 0$, so that $h(t) = (h_T(t), h_S(t))$.

Exactly as in (2.16), we have the corresponding equation for the S-bond

$$\frac{d\hat{B}^S(t,r(t))}{\hat{B}^S(t,r(t))} = \hat{\mu}^S(t,r(t))dt + \hat{\sigma}^S(t,r(t))dW^r(t), \tag{2.22}$$

where analogously to (2.12) and (2.13)

$$\begin{aligned} \hat{\mu}^S(t,r(t)) &= \frac{\hat{B}^S_t(t,r(t)) + \hat{B}^S_r(t,r(t))\hat{\mu}^r(t,r(t)) + \frac{1}{2}\hat{B}^S_{rr}(t,r(t))\big(\hat{\sigma}^r(t,r(t))\big)^2}{\hat{B}^S\,(t,r(t))}, \\ \hat{\sigma}^T(t,r(t)) &= \frac{\hat{B}^S_r(t,r(t))\hat{\sigma}^r(t,r(t))}{\hat{B}^S(t,r(t))}. \end{aligned}$$

As we have observed in Chapter 1 (see Definition 1.3.4), often it is convenient to describe a portfolio in relative terms using, instead of $h(t)$, the relative portfolio $U(t)$. From Assumption 2.2.2,

$$B(t,T) > 0, \quad \forall t, \omega$$

then we have that in (1.16) $U(t) = u(t)$. We use the notation $u(t) = (u_T, u_S)$ for the corresponding relative portfolio.

Setting $V(t) = V^h(t) = h_T(t)B(t,T) + h_S(t)B(t,S)$, the value process corresponding to the portfolio h, from Definition 1.3.4 we have that

$$u_T(t) = \frac{h_T(t)B(t,T)}{V(t)} = \frac{h_T(t)\hat{B}^T(t,r(t))}{V(t)}, \tag{2.23}$$

$$u_S(t) = \frac{h_S(t)\hat{B}(t,S)}{V(t)} = \frac{h_S(t)\hat{B}^S(t,r(t))}{V(t)}, \tag{2.24}$$

and

$$u_T(t) + u_S(t) = 1 \quad \forall t. \tag{2.25}$$

Using the self-financing condition in terms of the relative portfolio (see (1.18)), we obtain the following dynamics for the portfolio value V

$$dV(t) = V(t)\left(u_T(t)\frac{d\hat{B}^T(t,r(t))}{\hat{B}^T(t,r(t))} + u_S(t)\frac{d\hat{B}^S(t,r(t))}{\hat{B}^S(t,r(t))}\right). \tag{2.26}$$

Now we insert (2.16) and (2.22) (the expression for $\frac{d\hat{B}^T}{\hat{B}^T}$ and $\frac{d\hat{B}^S}{\hat{B}^S}$ respectively) into (2.26), and we have

$$\frac{dV(t)}{V(t)} = u_T(t)\Big(\hat{\mu}^T(t,r(t))dt + \hat{\sigma}^T(t,r(t))dW^r(t)\Big) + u_S(t)\Big(\hat{\mu}^S(t,r(t))dt + \hat{\sigma}^S(t,r(t))dW^r(t)\Big),$$

so that grouping dt and dW terms, we obtain

$$\frac{dV(t)}{V(t)} = \Big(u_T(t)\hat{\mu}^T(t,r(t)) + u_S(t)\hat{\mu}^S(t,r(t))\Big)dt + \Big(u_T(t)\hat{\sigma}^T(t,r(t)) + u_S(t)\hat{\sigma}^S(t,r(t))\Big)dW^r(t),$$

where the only restriction on the relative portfolio is given by (2.25). If the relative portfolio satisfies the following conditions

$$\begin{cases} u_T(t) + u_S(t) &= 1 \\ u_T(t)\hat{\sigma}^T(t,r(t)) + u_S(t)\hat{\sigma}^S(t,r(t)) &= 0 \end{cases} \tag{2.27}$$

then the value dynamics becomes

$$dV(t) = V(t)\Big(u_T(t)\hat{\mu}^T(t,r(t)) + u_S(t)\hat{\mu}^S(t,r(t))\Big)dt. \tag{2.28}$$

Thus the value process has no driving noise terms and we have obtained a riskless portfolio. From Assumption 2.2.2 the market is arbitrage-free and then the portfolio rate of return and the short rate of interest are equal (see Proposition 1.4.1), namely

$$u_T(t)\hat{\mu}^T(t,r(t)) + u_S(t)\hat{\mu}^S(t,r(t)) = r(t) \quad \forall t, a.s. \tag{2.29}$$

It is easily seen that

$$\begin{aligned} u_T(t) &= -\frac{\hat{\sigma}^S(t,r(t))}{\hat{\sigma}^T(t,r(t)) - \hat{\sigma}^S(t,r(t))} \\ u_S(t) &= \frac{\hat{\sigma}^T(t,r(t))}{\hat{\sigma}^T(t,r(t)) - \hat{\sigma}^S(t,r(t))} \end{aligned}$$

solve system (2.27). Substituting the above expression into (2.29), we have

$$\left(-\frac{\hat{\sigma}^S(t,r(t))}{\hat{\sigma}^T(t,r(t)) - \hat{\sigma}^S(t,r(t))}\right)\hat{\mu}^T(t,r(t)) + \left(\frac{\hat{\sigma}^T(t,r(t))}{\hat{\sigma}^T(t,r(t)) - \hat{\sigma}^S(t,r(t))}\right)\hat{\mu}^S(t,r(t)) = r(t),$$

or equivalently

$$\frac{\hat{\mu}^S(t,r(t))\hat{\sigma}^T(t,r(t)) - \hat{\mu}^T(t,r(t))\hat{\sigma}^S(t,r(t))}{\hat{\sigma}^T(t,r(t)) - \hat{\sigma}^S(t,r(t))} = r(t), \quad \forall t \quad a.s. \tag{2.30}$$

After some reshuffling, equation (2.30) can be rewritten as

$$\frac{\hat{\mu}^S(t,r(t)) - r(t)}{\hat{\sigma}^S(t,r(t))} = \frac{\hat{\mu}^T(t,r(t)) - r(t)}{\hat{\sigma}^T(t,r(t))}, \tag{2.31}$$

i.e.,

$$\frac{\mu(t,S) - r(t)}{\sigma(t,S)} = \frac{\mu(t,T) - r(t)}{\sigma(t,T)}. \tag{2.32}$$

Indeed equation (2.31) can be immediately obtained by grouping $\hat{\sigma}^T$ and $\hat{\sigma}^S$ terms in the right hand side of the following identity

$$\hat{\mu}^S(t,r(t))\hat{\sigma}^T(t,r(t)) - \hat{\mu}^T(t,r(t))\hat{\sigma}^S(t,r(t)) = r(t)\hat{\sigma}^T(t,r(t)) - r(t)\hat{\sigma}^S(t,r(t)).$$

Finally equation (2.32) shows that the left hand side of (2.19) does not depend on T and therefore (2.19) uniquely defines the process ξ_r.

□

Assuming that the support (the set of possible values) of the riskless interest rate $r(t)$ is the entire set $\mathbb{R}_+$, we now can state one of the most important result in the theory of interest rate: for each T the function $\hat{B}^T$ satisfies the so called *term structure equation*.

Theorem 2.2.3. *Assuming that the support of the riskless interest rate $r(t)$ is entire set $\mathbb{R}_+$, in an arbitrage free bond market the function $\hat{B}^T(t,r)$ satisfies the term structure equation*

$$\begin{cases} \hat{B}^T_t(t,r) + \big(\hat{\mu}^r(t,r) - \hat{\xi}_r(t,r)\hat{\sigma}^r(t,r)\big)\hat{B}^T_r(t,r) + \frac{1}{2}\hat{B}^T_{rr}(t,r)\big(\hat{\sigma}^r(t,r)\big)^2 - r\hat{B}^T(t,r) = 0, \\ \hat{B}(T,r;T) = 1, \end{cases} \tag{2.33}$$

where $(t,r) \in (0,T) \times \mathbb{R}_+$.

Remark 2.2.1. *If, for a fixed maturity time T, a T-bond price process $B(t,T)$ is observable, then $B(t,T)$ is called a benchmark of the bond market. If we assume that also $r(t)$ is observable, the obtained results can be interpreted by saying that all bond prices will be determined in terms of the benchmark T-bond and the short rate of interest. Indeed once the market has determined the dynamics of this benchmark $B(t,T)$, then $\mu(t,T)$ and $\sigma(t,T)$ can be considered as known together with $r(t)$, and therefore the market has implicitly specified ξ_r by equation (2.19). Once ξ_r is determined, all other bond prices will be determined by the term structure equation (2.33).*

Proof of Theorem 2.2.3. Inserting (2.12) and (2.13) into (2.21), we obtain

$$\frac{\frac{1}{\hat{B}^T}\left(\hat{B}^T_t + \hat{B}^T_r\hat{\mu}^r + \frac{1}{2}\hat{B}^T_{rr}(\hat{\sigma}^r)^2\right) - r}{\frac{1}{\hat{B}^T}\hat{B}^T_r\hat{\sigma}^r} = \hat{\xi}_r, \tag{2.34}$$

that is

$$\hat{B}^T_t + \hat{\mu}^r\hat{B}^T_r + \frac{1}{2}\hat{B}^T_{rr}(\hat{\sigma}^r)^2 - r\hat{B}^T = \hat{\xi}_r\hat{B}^T_r\hat{\sigma}^r, \tag{2.35}$$

and finally, grouping $\hat{B}^T_r$ terms, we have

$$\hat{B}^T_t + \big(\hat{\mu}^r - \hat{\xi}_r\hat{\sigma}^r\big)\hat{B}^T_r + \frac{1}{2}(\hat{\sigma}^r)^2\hat{B}^T_{rr} - r(t)\hat{B}^T = 0, \tag{2.36}$$

where for the notational convenience, the argument $(t, r(t))$ "has been suppressed", so that we have used the shorthand notation (2.17) and (2.18). Since we have assumed that the support of the process $r(t)$ is the entire set $\mathbb{R}_+$, we can then conclude that the equation (2.36) must also hold identically when we evaluate it at an arbitrary deterministic point (t, r). By Definition 2.2.1, we must also have $\hat{B}(T, r; T) = 1$, so we have proved the result.

$\square$

Before proceeding any further, we observe that the price dynamics of $B(t, T)$ can be expressed by mean of the market price ξ_r. Indeed, by Itô formula and the term structure equations (2.33), we have the price dynamics of the following form

$$\begin{aligned}&\frac{d\hat{B}^T(t, r(t))}{\hat{B}^T(t, r(t))}\\&= \left(r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\frac{\hat{B}_r^T}{\hat{B}^T}(t, r(t))\right) dt + \hat{\sigma}^r(t, r(t))\frac{\hat{B}_r^T}{\hat{B}^T}(t, r(t))dW^r(t).\end{aligned} \tag{2.37}$$

The proof of (2.37) follows by observing that

$$d\hat{B}^T(t, r(t)) = \hat{B}_t^T(t, r(t))dt + \hat{B}_r^T(t, r(t))dr(t) + \frac{1}{2}\hat{B}_{rr}^T(t, r(t))\big(\hat{\sigma}^r(t, r(t))\big)^2 dt, \tag{2.38}$$

and inserting the differential form (2.1), (2.2) and (2.3), into (2.38), we obtain

$$\begin{aligned}d\hat{B}(t, r(t)) &= \hat{B}_t^T(t, r(t))dt + \hat{B}_r^T(t, r(t))dr(t) + \frac{1}{2}\hat{B}_{rr}^T(t, r(t))\big(\hat{\sigma}^r(t, r(t))\big)^2 dt\\&= \left(\hat{B}_t^T(t, r(t)) + \hat{\mu}^r(t, r(t))\hat{B}_r^T(t, r(t)) + \frac{1}{2}\hat{B}_{rr}^T(t, r(t))\big(\hat{\sigma}^r(t, r(t))\big)^2\right) dt\\&\quad + \hat{\sigma}^r(t, r(t))\hat{B}_r^T(t, r(t))dW^r(t)\\&= \left(r(t)\hat{B}^T + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\hat{B}_r^T(t, r(t))\right) dt + \hat{\sigma}^r(t, r(t))\hat{B}_r^T(t, r(t))dW^r(t),\end{aligned}$$

where in the last step we have used the following relation

$$\hat{B}_t^T + \hat{\mu}^r\hat{B}_r^T + \frac{1}{2}\hat{B}_{rr}^T(\hat{\sigma}^r)^2 = \hat{\xi}_r\hat{\sigma}^r\hat{B}_r^T + r(t)\hat{B}^T,$$

given by the term structure equation (2.33) with all terms evaluated at the point $(t, r(t))$.

So far we have found that arbitrage free price process of a T-bond solves the term structure equation (2.33), but we observe that ξ_r is not determined within the model. In order to be able to solve (2.33), we must specify ξ_r exogenously just as we have to specify μ^r and σ^r.

Despite this problem, we can obtain more information by applying the Feynman-Kač representation to the function $\hat{B}^T$.

In the sequel we assume that the process $\xi_r \in \mathcal{L}^2(0, T; \bar{\mathbb{F}}^{W^r})$ and satisfies Novikov condition (1.34), so that the assumptions of Lemma 1.4.7 are satisfied if we choose

$$\xi(t) := \xi_r(t). \tag{2.39}$$

From Lemma 1.4.7 the process $\bar{W}^r(t)$ defined as

$$\bar{W}^r(t) := \int_0^t \xi_r(t)dt + W^r(t) \tag{2.40}$$

is a Wiener process with respect to the measure $Q = Q_T$ defined as

$$dQ = e^{-\int_0^T \xi_r(t)dW^r(t) - \frac{1}{2}\int_0^T \xi_r^2(t)dt}dP. \tag{2.41}$$

Assuming the integrability condition (1.39), let Q be the equivalent martingale such that, under Q, the riskless interest rate follows the dynamics

$$dr(t) \;\; = \;\; \hat{\mu}^r(t, r(t)) + \hat{\sigma}^r(t, r(t))\left[d\bar{W}^r(t) - \hat{\xi}_r(t, r(t))dt\right]. \tag{2.42}$$

Finally we obtain the following stochastic representation formula.

Proposition 2.2.4. *In an arbitrage free bond market, let $\xi_r \in \mathcal{L}^2(0,T;\mathbb{F}^{W^r})$ and satisfies Novikov condition (1.34). Then the bond prices are given by the formula (2.7) with*

$$\hat{B}^T(t,r) = E^Q_{t,r}\left(e^{-\int_t^T r(s)ds}\right), \tag{2.43}$$

where the measure martingale Q and the subscripts t and r denote that the expectation is taken using the dynamics given by (2.42), i.e.,

$$\begin{cases} dr(s) = & \left[\hat{\mu}^r(s,r(s)) - \hat{\xi}_r(s,r(s))\hat{\sigma}^r(s,r(s))\right] ds + \hat{\sigma}^r(s,r(s))d\bar{W}^r(s) \\ r(t) = & r \end{cases} \tag{2.44}$$

where $\bar{W}^r$ is a Wiener process with respect to Q defined in (2.40).

Proof. By Itô's formula we have

$$\begin{aligned} d\hat{B}^T(s,r(s)) &= \hat{B}^T_s ds + \hat{B}^T_r dr(s) + \frac{1}{2}\hat{B}^T_{rr} d\left\langle r(\cdot), r(\cdot)\right\rangle_s \\ &= \left[\hat{B}^T_s + (\hat{\mu}^r - \hat{\xi}_r\hat{\sigma}^r)\hat{B}^T_r + \frac{1}{2}(\hat{\sigma}^r)^2\hat{B}^T_{rr}\right] ds + \hat{\sigma}^r\hat{B}^T_r d\bar{W}^r(s), \end{aligned} \tag{2.45}$$

where we have used the same shorthand notations (2.17) and (2.18), but considering s instead of t. Now, we fix (t,r), set

$$Y(s) = e^{-\int_t^s r(u)du}, \quad s \in [t,T],$$

so that

$$dY(s) = -r(s)Y(s)ds, \tag{2.46}$$

and define the process Z as

$$Z(s) = Y(s)\hat{B}^T(s,r(s)), \quad s \in [t,T]. \tag{2.47}$$

Then, by (2.46) and (2.45), we obtain

$$\begin{aligned} dZ(s) &= Y(s)d\hat{B}^T(s,r(s)) + dY(s)\hat{B}^T(s,r(s)) = e^{-\int_t^s r(u)du}d\hat{B}^T - r(s)e^{-\int_t^s r(u)du}\hat{B}^T ds \\ &= e^{-\int_t^s r(u)du}\left[\hat{B}^T_s - r(s)\hat{B}^T + (\hat{\mu}^r - \hat{\xi}_r\hat{\sigma}^r)\hat{B}^T_r + \frac{1}{2}(\hat{\sigma}^r)^2\hat{B}^T_{rr}\right] ds \\ &\quad + e^{-\int_t^s r(u)du}\hat{\sigma}^r\hat{B}^T_r d\bar{W}^r(s). \end{aligned}$$

or equivalently

$$\begin{aligned} Z(T) &= Z(t) + \int_t^T e^{-\int_t^s r(u)du}\left[\hat{B}^T_s - r(s)\hat{B}^T + (\hat{\mu}^r - \hat{\xi}_r\hat{\sigma}^r)\hat{B}^T_r + \frac{1}{2}\hat{B}^T_{rr}(\hat{\sigma}^r)^2\right] ds \\ &\quad + \int_t^T e^{-\int_t^s r(u)du}\hat{\sigma}^r\hat{B}^T_r d\bar{W}^r(s). \end{aligned}$$

Since, by Theorem 2.2.3, $\hat{B}^T(s,r(s))$ satisfies equation (2.33) evaluated at the point $(s,r(s))$, the time integral will vanish in the above expression and using $\hat{B}^T(T,r) = 1$ we obtain

$$e^{-\int_t^T r(s)ds} = \hat{B}^T(t,r) + \int_t^T e^{-\int_t^s r(u)du}\hat{\sigma}^r(s,r(s))\hat{B}^T_r(s,r(s))d\bar{W}^r(s). \tag{2.48}$$

Taking the expectation of (2.48), we have

$$E^Q_{t,r}\left(e^{-\int_t^T r(s)ds}\right) = \hat{B}^T(t,r),$$

the expected value of the stochastic integral being equal to zero. We have proved the announced result. □

Rewriting formula (2.43) as

$$\hat{B}^T(t,r) = E^Q_{t,r}\left[e^{-\int_t^T r(s)ds} \cdot 1\right], \tag{2.49}$$

we see that the value of a T-bond at time t is given as the expected value of one dollar (final payoff), discount to present value. Thus formula (2.43) is exactly the risk-neutral pricing formula. The main difference between the present situation and the risk-neutral pricing formula setting, is that in the latter model the martingale measure is uniquely determined, while in our model we may have different martingale measures for different choices of ξ_r.

2.2.1 Cox-Ingersoll-Ross interest rate model

The financial literature on interest rate modelling is full of examples of affine processes: the Ornstein-Uhlenbeck process, used by Vasicek (1977), and the so called CIR process, an extension of the Ornstein-Uhlenbeck process, introduced by Cox, Ingersoll and Ross (1985). These models are popularized in finance as Vasicek and Cox-Ingersoll-Ross (CIR) model, respectively.

The Vasicek model specifies that the instantaneous interest rate is a Ornstein-Uhlenbeck process, i.e.,

$$dX(t) = \theta\left(\mu - X(t)\right)dt + \sigma d\bar{W}^r(t),$$

where θ, μ and σ are deterministic positive constants and $\bar{W}^r(t)$ is a Wiener process under a martingale measure Q, while the CIR model specifies that the instantaneous interest rate follows the stochastic differential equation (also named the CIR process)

$$dr(t) = a_r\left(b_r - r(t)\right)dt + \bar{\sigma}_r\sqrt{r(t)}d\bar{W}^r(t), \tag{2.50}$$

where a_r, b_r and $\bar{\sigma}_r$ are deterministic positive constants and $\bar{W}^r(t)$ is a Wiener process under a martingale measure Q. In particular if the initial condition is strictly positive, then $r(t) \geq 0$, but if furthermore $2\,a_r\,b_r > \bar{\sigma}_r^2$ then the process $r(t)$ remains strictly positive, i.e., $P(r(t) > 0) = 1$, $\forall t$, as shown in Shreve [21].

The Vasicek model is often preferred in modelling interest rate since it allows for easy closed form solutions, but the important difference between the Vasicek and CIR model is that in the latter the interest rate is positive, while in the Vasicek model this is not the case, indeed the probability that the interest rate takes negative value is strictly positive.

The convenience of adopting affine processes in modelling the interest rate lies in the fact that, under technical conditions (see Duffie and Singleton [9]), for an affine process $X(t)$ with values in $D \subset \mathbb{R}^d$ we have that

$$E_{t,x}\left[e^{-\int_t^s \Lambda(X(u))du + mX(s)}\right] = e^{\psi_X^0(s-t) + \psi_X(s-t)x}, \tag{2.51}$$

for any affine function $\Lambda : D \to \mathbb{R}$ and any $m \in \mathbb{R}^d$, where the deterministic coefficients $\psi_X^0(\cdot)$ and $\psi_X(\cdot)$ have to be determined. Thanks to this property, we can derive an explicit formula for the price of a zero coupon bond as a function of the interest rate, as shown in the following proposition.

Proposition 2.2.5. *The term structure for the CIR model is given by*

$$\hat{B}^T(t,r) = e^{\psi_r^0(T-t) + \psi_r(T-t)r}, \tag{2.52}$$

where

$$\psi_r(s) = \frac{1 - e^{\alpha_r s}}{\beta_r + \gamma_r\, e^{\alpha_r s}}, \tag{2.53}$$

$$\psi_r^0(s) = -\frac{2\,a_r b_r}{\bar{\sigma}_r^2}\ln\left(\frac{\beta_r + \gamma_r\, e^{\alpha_r s}}{\alpha_r}\right) + \frac{a_r b_r}{\beta_r}\, s \tag{2.54}$$

and

$$\alpha_r = -\sqrt{a_r^2 + 2\bar{\sigma}_r^2}, \quad \beta_r = \frac{\alpha_r - a_r}{2}, \quad \gamma_r = \frac{\alpha_r + a_r}{2}. \tag{2.55}$$

Observe that this result is well known in literature, but for notational convenience we report the proof of this classic result. To this end we refer to Shreve [21].

Proof. By property (2.51) and by formula (2.43), given $r(t)$, we have that

$$\hat{B}^T(t,r) = E_{t,r}^Q\left[e^{-\int_t^T r(u)du}\right] = e^{\psi_r^0(T-t) + \psi_r(T-t)r}. \tag{2.56}$$

Furthermore, under the usual regularity conditions for the Feynman-Kač approach, we can get the functions $\psi_r^0(\cdot)$ and $\psi_r(\cdot)$, by recalling that $\hat{B}^T$ solves the partial differential equation

$$\begin{cases} \hat{B}_t^T(t,r) + a_r\left(b_r - r\right)\hat{B}_r^T(t,r) + \frac{\bar{\sigma}_r^2}{2}\, r\,\hat{B}_{rr}^T(t,r) = r\hat{B}^T(t,r), \\ \hat{B}^T(T,r) = 1. \end{cases} \tag{2.57}$$

By the property (2.52) we obtain

$$\hat{B}^T_t = \hat{B}^T\left(-\dot{\psi}^0_r(T-t) - r\dot{\psi}_r(T-t)\right), \tag{2.58}$$

$$\hat{B}^T_r(t,r) = \hat{B}^T(t,r)\psi_r(T-t), \quad \hat{B}^T_{rr}(t,r) = \hat{B}^T(t,r)\psi^2_r(T-t), \tag{2.59}$$

where $\dot{\psi}^0_r(u) = \frac{d\psi^0_r}{du}(u)$ and $\dot{\psi}_r(u) = \frac{d\psi_r}{du}(u)$. Substituting into the partial differential equation (2.57) and dividing each term by the common factor $\hat{B}^T$, we have

$$-\dot{\psi}^0_r(T-t) - \dot{\psi}_r(T-t)r + a_r(b_r - r)\psi_r(T-t) + \frac{\bar{\sigma}^2_r}{2}\, r\, \psi^2_r(T-t) = r,$$

so that, grouping the terms multiplying r, we obtain

$$\left(\dot{\psi}_r(T-t) + a_r\psi_r(T-t) - \frac{\bar{\sigma}^2_r}{2}\psi^2_r(T-t) + 1\right) r + \dot{\psi}^0_r(T-t) - a_r b_r \psi_r(T-t) = 0. \tag{2.60}$$

Since (2.60) holds for all $r \geq 0$, we have that the terms multiplying r are equal to zero, as well as the other term, so that we obtain two ordinary differential equations in $T-t=s$ given by

$$\dot{\psi}^0_r(s) = a_r b_r \psi_r(s), \tag{2.61}$$

$$\dot{\psi}_r(s) = \frac{\bar{\sigma}^2_r}{2}\psi^2_r(s) - a_r\psi_r(s) - 1, \tag{2.62}$$

with the initial conditions

$$\psi^0_r(0) = 0, \qquad \psi_r(0) = 0,$$

which are derived from the terminal condition $\hat{B}^T(T,r) = 1$. Observe that the equation (2.62) is called Riccati equation and the solution is given by[6] (2.53). Now substituting the expression for $\psi_r(s)$ into (2.61), we obtain the expression for $\psi^0_r(s)$ given by (2.54).

$\square$

2.3 A two-dimensional market model: Bond and Stock

In this section we consider a market model consisting, besides the money market account $G(t)$, of only two assets, i.e. a zero coupon bond, defined as in Section 2.1, with price process $B(t,T)$, and one stock, with price process $S(t)$, where $S(t)$ is only[7] influenced by a 2-dimensional Wiener process, $\left(W^r, W^S\right)'$. We will shortly write (G,B,S) to denote this market.

Accordingly to Section 2.1, we consider the riskless interest rate $r(t)$, evolving as in (2.1), (2.2) and (2.3). Recall that $r(t)$ is modelled as an adapted process defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, (see Definition 1.2.4), where in this setting, by Assumptions 1.2.1 and 1.2.2,

$$\mathbb{F} = \bar{\mathbb{F}}^W,$$

with $W = \left(W^r, W^S\right)$. We need the following assumption.

---

[6] A method to solve the equation (2.62) is given by

$$y(s) = e^{-\int_0^s \frac{2}{2}\,\psi_r(u)\,du}.$$

Using the expression for $\dot{\psi}_r(s)$, we obtain a homogeneous second order constant coefficient linear ordinary differential equation for $y(s)$

$$\begin{cases} \ddot{y} + a_r\dot{y} - \frac{\bar{\sigma}^2_r}{2}\, y = 0, \\ \dot{y}(0) = 0 \\ y(0) = 1 \end{cases}$$

whose the solution is given by

$$y(s) = -\frac{\xi_-}{\xi_+ - \xi_-}\, e^{\xi_+\, s} + \frac{\xi_+}{\xi_+ - \xi_-}\, e^{\xi_-\, s},$$

where $\xi_\pm$ are the characteristic polynomial solutions given by

$$\xi_+ = \frac{-a_r + \sqrt{a_r^2 + 2\bar{\sigma}^2_r}}{2} \quad \xi_- = \frac{-a_r - \sqrt{a_r^2 + 2\bar{\sigma}^2_r}}{2}.$$

Substituting $y(s) = e^{-\int_0^s \frac{\bar{\sigma}^2_r}{2}\,\psi_r(u)\,du}$ in the above expression, by taking logarithms and next deriving with respect to s, we obtain the solution $\psi_r(s)$.

[7] Observe that the stock S has its own risk source.

Assumption 2.3.1. *Assume that the market (G, B, S) is arbitrage free. Assume furthermore that $B(t,T)$ is given by (2.7) and Lemma 2.2.1, while the price process S is given by*

$$\frac{dS(t)}{S(t)} = \mu^S dt + \sigma_r^S dW^r(t) + \sigma_S^S dW^S(t), \tag{2.63}$$

where μ^S, σ_r^S and σ_S^S are deterministic constants.

From (2.63) we observe that the price process S is assumed to be a geometric Brownian motion. In particular we have that the process $S(t)$ is strictly positive, i.e., $S(t) > 0$, for all ω, t.

Since the dynamics of $B(t,T)$ are given by (2.9), we have the following market structure

$$\begin{aligned} \frac{dB(t,T)}{B(t,T)} &= \mu(t,T)dt + \sigma(t,T)dW^r(t) \\ \frac{dS(t)}{S(t)} &= \mu^S dt + \sigma_r^S dW^r(t) + \sigma_S^S dW^S(t). \end{aligned}$$

Then by Theorem 1.4.6, we know that there exists an adapted 2-dimensional process $(\xi_r(t), \xi_S(t))'$, such that

$$\Sigma(t)\left(\xi_r(t), \xi_S(t)\right)' = \left(\hat{\mu}^T(t), \mu^S\right)' - r(t) \quad \forall t \text{ a.s.}, \tag{2.64}$$

where

$$\Sigma(t) = \begin{pmatrix} \hat{\sigma}^T(t) & 0 \\ \sigma_r^S & \sigma_S^S \end{pmatrix}.$$

From (2.64), we obtain that

$$\xi_r(t) = \frac{\mu(t,T) - r(t)}{\sigma(t,T)} \tag{2.65}$$

$$\xi_S(t) = \frac{\mu^S - \sigma_r^S \xi_r(t) - r(t)}{\sigma_S^S}. \tag{2.66}$$

From (2.65) we immediately have that ξ_r coincides exactly with the process given by (2.19), indeed ξ_r is a market price for the riskless interest rate, while ξ_S is a market price for the stock.

Furthermore by (2.65) we obtain again (2.20), i.e. $\xi_r(t) = \hat{\xi}_r(t, r(t))$, while by (2.66) we have that $\xi_S(t)$ can be expressed as a deterministic function of t and $r(t)$, namely

$$\xi_S(t) = \hat{\xi}_S(t, r(t)), \tag{2.67}$$

but not of the process $S(t)$. Then (2.66) becomes

$$\hat{\xi}_S(t, r(t)) = \frac{\mu^S - \sigma_r^S \hat{\xi}_r(t, r(t)) - r(t)}{\sigma_S^S}. \tag{2.68}$$

Now, substituting (2.68) into (2.63) we obtain

$$\frac{dS(t)}{S(t)} = \left(r(t) + \sigma_r^S \hat{\xi}_r(t, r(t)) + \sigma_S^S \hat{\xi}_S(t, r(t))\right) dt + \sigma_r^S dW^r(t) + \sigma_S^S dW^S(t), \tag{2.69}$$

i.e., the price process S can be expressed by mean of the market price $\left(\hat{\xi}_r, \hat{\xi}_S\right)'$.

Remark 2.3.1. *In particular we observe that similar results hold also if μ^S, σ_r^S and $\sigma_S^{r,S}$ depend on t, $r(t)$ and $S(t)$. More precisely we have that*

1. *if μ^S, σ_r^S and σ_S^S depend on t and $r(t)$, then $\xi_S(t)$ may still be expressed as a deterministic function of t and $r(t)$, as in (2.67);*

2. *if $\mu^S(t) = \hat{\mu}^S(t, r(t), S(t))$, $\sigma_r^S = \hat{\sigma}_r^S(t, r(t), S(t))$ and $\sigma_S^S = \hat{\sigma}_S^S(t, r(t), S(t))$, where $\hat{\mu}^S$, $\hat{\sigma}_r^S$ and $\hat{\sigma}_S^S$ are deterministic functions, then $\xi_S(t)$ may still be expressed as a deterministic function of t, $r(t)$, and $S(t)$, i.e.,*

$$\xi_S(t) = \hat{\xi}_S(t, r(t), S(t)), \tag{2.70}$$

where by slight abuse of notation we have used the same symbol $\hat{\xi}_S$ to denote the deterministic function in (2.70) and (2.67). In this case the dynamics of $S(t)$ is described by the following stochastic differential equation

$$\frac{dS(t)}{S(t)} = \Big(r(t) + \hat{\sigma}_r^S(t, r(t), S(t))\hat{\xi}_r(t, r(t)) + \hat{\sigma}_S^S(t, r(t), S(t))\hat{\xi}_S(t, r(t), S(t))\Big)\, dt$$
$$+ \hat{\sigma}_r^S(t, r(t), S(t)) dW^r(t) + \hat{\sigma}_S^S(t, r(t), S(t)) dW^S(t). \tag{2.71}$$

2.4 Discrete-time Rolling Bonds

In this section the aim is focused on the problem of modelling a discrete-time rolling bond price process, and we refer to Rutkowski [20].

Now we fix a discrete set of times $\mathcal{T} = \{t_k\}_{k\geq 0}$ such that $t_k \leq t_{k+1}$ and consider a self-financing strategy such that, its total wealth is reinvested at any fixed date $t \in \mathcal{T}$ in discount bonds maturing at time $t+T$ (i.e., no cash component is present). For a fixed T, the price process of this strategy is referred to as the discrete-time rolling bond. In particular here we fix $\Delta \in (0, T)$ and take $t_k = k\Delta$, for $k = 0, 1, 2, \ldots$, and we denote $U^{\Delta}(t, T)$ the corresponding price process.

Recalling that the price of a T-sliding bond is the price at time t of a $T+t$-bond, i.e., $B(t, T+t)$, we observe that, in contrast to the rolling bond, it is not possible to trade in arbitrage-free market a sliding bond, since it does not represent a self-financing trading strategy, so that it cannot be considered as a tradable security in an arbitrage-free market (see Rutkowski [20]).

Assume that at time $t \in [t_0, t_1) = [0, \Delta)$ we hold 1 bond, so that $U^{\Delta}(0, T) = B(0, T)$ and

$$U^{\Delta}(t, T) = B(t, T) = \hat{B}^T(t, r(t)) \qquad 0 \leq t < \Delta.$$

At time $t_1 = \Delta$, the wealth $B(\Delta, T)$ is reinvested in bonds maturing at time $T + \Delta$ and we keep it until time $t_2 = 2\Delta$, so that

$$U^{\Delta}(t, T) = \frac{B(\Delta, T)}{B(\Delta, T+\Delta)} B(t, T+\Delta), \qquad \Delta \leq t < 2\Delta.$$

Consequently, we have

$$U^{\Delta}(t, T) = \frac{B(\Delta, T)}{B(\Delta, T+\Delta)} \frac{B(2\Delta, T+\Delta)}{B(2\Delta, T+2\Delta)} B(t, T+2\Delta) \qquad 2\Delta \leq t < 3\Delta,$$

and for $k\Delta \leq t < (k+1)\Delta$ we have

$$\begin{aligned} U^{\Delta}(t, T) &= \frac{B(\Delta, T)}{B(\Delta, T+\Delta)} \frac{B(2\Delta, T+\Delta)}{B(2\Delta, T+2\Delta)} \cdots \frac{B^{T+(k-1)\Delta}(k\Delta, r(k\Delta))}{B(k\Delta, T+k\Delta)} B(t, T+k\Delta) \\ &= \frac{\hat{B}^T(\Delta, r(\Delta))}{\hat{B}^{T+\Delta}(\Delta, r(\Delta))} \frac{\hat{B}^{T+\Delta}(2\Delta, r(2\Delta))}{\hat{B}^{T+2\Delta}(2\Delta, r(2\Delta))} \cdots \frac{\hat{B}^{T+(k-1)\Delta}(k\Delta, r(k\Delta))}{\hat{B}^{T+k\Delta}(k\Delta, r(k\Delta))} \hat{B}^{T+k\Delta}(t, r(t)). \end{aligned} \tag{2.72}$$

Simple induction arguments show that, for any $t \geq 0$, the price process of the discrete-time rolling bond satisfies

$$U^{\Delta}(t, T) = \prod_{k=1}^{\lfloor t/\Delta \rfloor} \frac{B(k\Delta, T+(k-1)\Delta)}{B(k\Delta, T+k\Delta)} B(t, T+\lfloor t/\Delta \rfloor \Delta) = U^{\Delta}(\lfloor t/\Delta \rfloor \Delta)\, \hat{B}^{T+\lfloor t/\Delta \rfloor \Delta}(t, r(t)). \tag{2.73}$$

The last formula leads to the following result.

Proposition 2.4.1. *Let $B(t, T)$ be a zero coupon bond with price processes given by (2.37). For any fixed T, the price process $U^{\Delta}(\cdot, T)$ of the discrete-time rolling bond satisfies*

$$\frac{dU^{\Delta}(t, T)}{U^{\Delta}(t, T)} = \mu_{U^{\Delta}}(t, T)dt + \sigma_{U^{\Delta}}(t, T)dW^r(t), \tag{2.74}$$

where

$$\mu_{U^{\Delta}}(t, T) = \hat{\mu}^T_{U^{\Delta}}(t, r(t)) = r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t)) \frac{\hat{B}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{B}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, r(t)), \tag{2.75}$$

$$\sigma_{U^{\Delta}}(t, T) = \hat{\sigma}^T_{U^{\Delta}}(t, r(t)) = \hat{\sigma}^r(t, r(t)) \frac{\hat{B}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{B}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, r(t)). \tag{2.76}$$

Proof. By the formula for the price process U^Δ (2.73) we have that

$$\begin{aligned} dU^\Delta(t,T) &= U^\Delta(\lfloor t/\Delta\rfloor\Delta, T)\, d\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t)) \\ &= U^\Delta(\lfloor t/\Delta\rfloor\Delta)\, \hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t)) \frac{d\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))}{\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))} \\ &= U^\Delta(t,T)\, \frac{d\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))}{\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))}, \end{aligned}$$

so that

$$\frac{dU^\Delta(t,T)}{U^\Delta(t,T)} = \frac{d\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))}{\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t, r(t))},$$

and, since $B(t,T) = \hat{B}^T(t, r(t))$ satisfies the stochastic differential equation (2.37), we have proved the announced result.

$\square$

In particular, let us consider the CIR model introduced in Section 2.2.1. By the explicit formula for bonds (2.52) and the expression (2.73) we obtain the following explicit formula

$$U^\Delta(t,T) = \prod_{k=1}^{\lfloor t/\Delta\rfloor} \frac{e^{\psi_0(T-\Delta)+r(k\Delta)\psi_r(T-\Delta)}}{e^{\psi_0(T)+r(k\Delta)\psi_r(T)}} e^{\psi_0(T+\lfloor t/\Delta\rfloor\Delta-t)+r(t)\psi_r(T+\lfloor t/\Delta\rfloor\Delta-t)}, \qquad t > 0 \tag{2.77}$$

taking into account that (see 2.72) for $k\Delta \le t < (k+1)\Delta$

$$U^\Delta(t,T) = \\ \frac{e^{\psi_0(T-\Delta)+r(\Delta)\psi_r(T-\Delta)}}{e^{\psi_0(T)+r(\Delta)\psi_r(T)}} \frac{e^{\psi_0(T-\Delta)+r(2\Delta)\psi_r(T-\Delta)}}{e^{\psi_0(T)+r(2\Delta)\psi_r(T)}} \cdots \frac{e^{\psi_0(T-\Delta)+r(k\Delta)\psi_r(T-\Delta)}}{e^{\psi_0(T)+r(k\Delta)\psi_r(T)}} e^{\psi_0(T+k\Delta-t)+r(t)\psi_r(T+k\Delta-t)}. \tag{2.78}$$

Furthermore, in this framework, $\mu_{U^\Delta}(t,T)$ and $\sigma_{U^\Delta}(t,T)$ given by (2.75) and (2.76) become

$$\mu_{U^\Delta}(t,T) = \hat{\mu}^T_{U^\Delta}(t, r(t)) = r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\psi_r(T + \lfloor t/\Delta\rfloor\Delta - t), \tag{2.79}$$

$$\sigma_{U^\Delta}(t,T) = \hat{\sigma}^T_{U^\Delta}(t, r(t)) = \hat{\sigma}^r(t, r(t))\psi_r(T + \lfloor t/\Delta\rfloor\Delta - t). \tag{2.80}$$

Chapter 3

Survival models

3.1 Introduction

In the last decades, significant improvements in the duration of life have been experienced in most developed countries. The mortality risk and, in particular, the longevity risk has been largely studied in recent years when dealing with the pricing of insurance products. It is well known that the price of any insurance product on the duration of life depends on two main basis: demographical and financial assumptions. Traditionally, actuaries have been treating both the demographic and the financial assumptions in a deterministic way, by considering available mortality tables for describing the future evolution of mortality.

More recently, stochastic models have been adopted to describe the uncertainty linked both to mortality and to financial factors. In this chapter we focus on the mortality risk and on modelling the survival function of the individual, leaving a stochastic approach of both mortality and financial risks to Chapter 4.

We refer to Brémaud [6] and Duffie [8], for basic theory of point processes with a stochastic intensity, and modelling the dynamic mortality, respectively.

3.2 Mortality risk

We consider an individual aged $x = 0$, i.e., a new-born individual, and denote by τ the random variable that describes his duration of life on a space probability $(\Omega, \mathcal{F}, P)$. The survival function, denoted by $\bar{F}(t)$, is defined as follows

$$\bar{F}(t) = P(\tau > t) = 1 - F(t), \tag{3.1}$$

with

$$\bar{F}(0) = P(\tau > 0) = 1, \tag{3.2}$$

where F is the distribution function of τ.

Two indicators are typically used to describe the mortality of an individual: the survival function and the mortality intensity. The survival function indicates the probability that a new-born individual will survive at least t years.

Analogously we consider an individual aged $x \geq 0$ and τ_x is the random variable that describe his future lifetime. Then τ_x is the life's duration of an individual aged x, given that he is alive at that age, i.e.,

$$\tau_x \stackrel{\mathcal{L}}{=} \tau - x | \tau > x \tag{3.3}$$

Via the survival function, we can derive the distribution function of τ_x, given that he/she is alive at that age, as

$$\begin{aligned} F_x(t) &= P(\tau_x \leq t) = P(\tau \leq t + x \,|\tau > x\,) \\ &= \frac{P\,(x < \tau \leq t + x)}{P(\tau > x)} = \frac{F(t+x) - F(x)}{1 - F(x)} = 1 - \frac{\bar{F}(x+t)}{\bar{F}(x)}. \end{aligned}$$

As in (3.1) we have also that the survival function of an individual aged x is

$$\bar{F}_x(t) = 1 - F_x(t).$$

The (deterministic) mortality intensity (or[1] mortality force) is defined as

$$\mu(x) = \lim_{\Delta x \to 0} \frac{P(x \le \tau < x + \Delta x \,|\tau \ge x)}{\Delta x}, \tag{3.4}$$

i.e., the probability of dying in a short period of time after x, between age x and age $x+\Delta x$, can be approximated by $\mu(x)\Delta x$, when Δx is small. For large values of the age, the mortality force is increasing as x increases, as the probability of imminent death increases when ageing[2].

From (3.4), if there is the density function of τ, denoted by f, we have

$$\mu(x)dx = P(x \le \tau < x + dx \,|\tau \ge x) = \frac{f(x)dx}{\bar{F}(x)},$$

or equivalently

$$\mu(x)dx = \frac{-\bar{F}'(x)}{\bar{F}(x)} = -\frac{d}{dx} \log \bar{F}(x).$$

Thus, since $\bar{F}(0) = 1$, we can write

$$\bar{F}(x) = e^{-\int_0^x \mu(s)ds}$$

or analogously

$$\bar{F}_x(t) = e^{-\int_0^t \mu_x(s)ds}, \tag{3.5}$$

where $\mu_x(t) = \mu(x+t)$, $t \ge 0$.

3.3 The mathematical framework

Before proceeding any further, in this section we focus on some necessary mathematical tools for a different and more appropriate approach to modelling the mortality risk, which includes the adoption of stochastic models. Now we describe a brief review of the theory of counting process, doubly stochastic Poisson process and their stochastic intensities. A realization of a point process over $[0, \infty)$ can be described by a sequence of random variable $\{T_n : n \in \mathbb{N}\}$, defined on a probability space $(\Omega, \mathcal{F}, P)$, with values in $[0, \infty]$, where

$$T_0 = 0, \tag{3.6}$$

and increasing in the following sense

$$T_n < \infty \quad \text{imply} \quad T_n < T_{n+1}. \tag{3.7}$$

This realization is, by definition, nonexplosive if

$$T_\infty = \lim_{n \to \infty} T_n = \infty. \tag{3.8}$$

To each realization T_n corresponds a counting function $N(t)$ defined by

$$N(t) = \begin{cases} n & \text{if } t \in [T_n, T_{n+1}), \\ +\infty & \text{if } t \ge T_\infty, \end{cases} \tag{3.9}$$

or analogously

$$N(t) = \sum_{n \ge 1} \mathbf{1}_{\{T_n \le t\}}. \tag{3.10}$$

$N(t)$ is therefore a right-continuous step function such that $N(0) = 0$, and its jumps are upward jumps of magnitude 1.

Thus we have the following definition.

[1] Mortality force is used particularly in demography and actuarial science.

[2] There are exceptions, like very small values of x (due to the infant mortality) and values around 20-25 (due to the young mortality).

Definition 3.3.1 (Point Process). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A sequence of increasing random variable $\{T_n : n \in \mathbb{N}\}$ on a probability space $(\Omega, \mathcal{F}, P)$, satisfying (3.6) and (3.7), is called a point process. The associated counting process $N = (N(t) : t \geq 0)$, defined as in (3.9), is also called a point process, by abuse of notation*[3]. *The point process is also said to be nonexplosive if, for all $t \geq 0$, $N(t) < \infty$ almost surely (or equivalently if $T_\infty = \infty$ almost surely.). Moreover, when the condition*

$$E[N(t)] < \infty, \quad t \geq 0 \tag{3.11}$$

holds, the point process N is said to be integrable.

Clearly one can consider T_n as the n^{th} jump time of the process N, and $N(t)$ as the number of jumps occurred up to time t, including time t.

Now we introduce a particularly important class of models, the Poisson processes. There exist several equivalent definitions of a Poisson process, the one adopted here is given in terms of counting process.

Definition 3.3.2 (Poisson Process). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\lambda(t)$ be a positive measurable (deterministic) function such that*

$$\int_0^t \lambda(u)du < \infty, \quad t \geq 0. \tag{3.12}$$

The nonexplosive counting process N is called a Poisson process with the intensity function $\lambda(t)$ if the following conditions are satisfied.

1. *For all s and $t > s$, the random variable $N(t)-N(s)$ has the Poisson distribution with parameter $\int_s^t \lambda(u)du$;*
2. *The process N has independent increments, i.e., for all $n \in \mathbb{N}$, and for any choice of mutually disjoint intervals $(s_i, t_i]$, $(1 \leq i \leq n)$, the random variables $N(t_i) - N(s_i)$, $(1 \leq i \leq n)$, are independent.*

If in addition $\lambda(t) = \bar{\lambda}$, N is called a homogeneous Poisson process with intensity $\bar{\lambda}$.

Now it is important to make a distinction between an adapted process and a predictable process. Intuitively, a process is predictable if, at any time t, it depends only on the information in the underlying filtration that is aviable up to, but not including, time t.

We have the following definition.

Definition 3.3.3 ($\mathbb{F}$-Predictable Process). *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space*[4] *satisfying the usual conditions. A process Y is said to be predictable if $Y : \Omega \times [0, \infty] \to \mathbb{R}$ is measurable with respect to the σ-algebra on $\Omega \times [0, \infty]$ generated by the set of all left-continuous adapted processes. We will shortly write $\mathbb{F}$-predictable process.*

In all practical applications, the predictable processes to be encountered are adapted processes and left-continuous, in fact any left-continuous adapted process is predictable, as is, in particular, any continuous adapted process (see Theorem T5 in Section I.3 of Brémaud [6]).

Now we are ready for the notion of the stochastic intensity, and in particular the stochastic intensity takes into account the dynamics of a counting process. It is a local description that tells what is expected to happen in the next infinitesimal interval given the past of the point process. The efficient formulation of this notion is in terms of martingales.

Definition 3.3.4 ($\mathbb{F}$-Stochastic Intensity). *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual conditions. Let λ be a positive $\mathbb{F}$-predictable process such that for all $t \geq 0$*

$$\int_0^t \lambda(s)ds < \infty \quad a.s. \tag{3.13}$$

A nonexplosive $\mathbb{F}$-adapted counting process N is said to admit the intensity λ if the compensator of N admits the representation $\int_0^t \lambda(s)ds$, i.e., if

$$M(t) = N(t) - \int_0^t \lambda(s)ds, \quad t \geq 0, \tag{3.14}$$

is a $\mathbb{F}$-local martingale. We will shortly write $\mathbb{F}$-stochastic intensity.

---

[3] An innocuous one, since N and $\{T_n : n \geq 0\}$ obviously carry the same information.

[4] We observe that in this chapter we consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where $\mathbb{F}$ is a generic filtration of sub-σ-algebras of $\mathcal{F}$, i.e., Assumption 1.2.2 is not a priori holds.

We observe that the requirement of predictability allows us to consider the intensity as essentially unique (see Theorem T12 in Section I.3 of Brémaud [6]) and, obviously, a counting process with a deterministic intensity is a Poisson process.

Now we have the following result linking point processes and martingales.

Proposition 3.3.1. *Suppose N is an $\mathbb{F}$-adapted counting process and λ is a positive $\mathbb{F}$-predictable process such that for all $t \geq 0$*

$$E\left[\int_0^t \lambda(s)ds\right] < \infty. \tag{3.15}$$

Then the following results are equivalent

1. *N is nonexplosive and λ is the $\mathbb{F}$-stochastic intensity of N;*
2. *$\left\{M(t) = N(t) - \int_0^t \lambda(s)ds : t \geq 0\right\}$ is a $\mathbb{F}$-martingale.*

Proof. From Theorems T8 and T9 in Section II.3 of Brémaud [6] we have that property 1. implies property 2. and the converse respectively.

□

As a consequence of Proposition 3.3.1, when N is a nonexplosive point process with the intensity λ satisfying (3.15), for all $0 \leq s \leq t$, we have that

$$E\left[N(t) - N(s)\,|\mathcal{F}_s\right] = E\left[M(t) - M(s)\,|\mathcal{F}_s\right] + E\left[\int_s^t \lambda(v)dv\,|\mathcal{F}_s\right] = E\left[\int_s^t \lambda(v)dv\,|\mathcal{F}_s\right],$$

M being a $\mathbb{F}$-martingale, and so[5]

$$E\left[N(t) - N(s)\,|\mathcal{F}_s\right] = E\left[\int_s^t \lambda(v)dv\,|\mathcal{F}_s\right]. \tag{3.16}$$

In particular, if $\lambda(t)$ is bounded and right-continuous from (3.16) we have that

$$\lim_{t \to s^+} \frac{1}{t-s} E\left[N(t) - N(s)\,|\mathcal{F}_s\right] = \lambda(s) \quad \text{a.s.}, \tag{3.17}$$

by application of the Lebesgue averaging theorem and the Lebesgue dominated-convergence theorem successively. Equation (3.17) (see the analogy with equation (3.4)) stresses the importance of the process λ in giving information about the average number of jumps of the process under observation in a small period of future time. The idea is that, at time t, the jump intensity $\lambda(t)$ gives information about the expected number of jumps in the next future or, in other words, about the likelihood of a jump in the immediate future. It cannot predict the actual occurrence of a jump, that comes as a "sudden surprise".

The following type of point processes is very common in applications. It is a “doubly stochastic” Poisson process, in the sense that it can be constructed in two steps. First one draws a random intensity function, that is a real positive measurable locally stochastic process, $\lambda = \{\lambda(t) : t \geq 0\}$, and having done so, one generates a Poisson process N with the intensity function $\lambda(t)$. Formally we have the following definition.

Definition 3.3.5 (Doubly Stochastic Poisson Process)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{G}$ be a σ-algebra such that $\mathcal{F} \supseteq \mathcal{G} \supseteq \mathcal{F}^\lambda_\infty$, where $\mathcal{F}^\lambda_\infty = \sigma\left(\lambda(t) : t \geq 0\right)$ and λ is a real positive measurable locally stochastic process such that (3.15) holds. A point process N is called a doubly stochastic Poisson process (or Cox process) with respect to $\mathcal{G}$ with the intensity function λ if, conditionally on $\mathcal{G}$, N is a Poisson process with the intensity function λ, i.e., for all $0 \leq s \leq t$ and all $u \in \mathbb{R}$,*

1. $E\left[e^{iu(N(t)-N(s))}\,|\mathcal{G}\right] = \exp\left\{\left(e^{iu} - 1\right)\int_s^t \lambda(v)dv\right\}$;
2. *The process N has, conditionally on $\mathcal{G}$, independent increments, i.e., for all $n \in \mathbb{N}$, and for any choice of mutually disjoint intervals $(s_i, t_i]$, $(1 \leq i \leq n)$, the random variables $N(t_i) - N(s_i)$, $(1 \leq i \leq n)$, are, conditionally on $\mathcal{G}$, independent.*

In the sequel we make the following assumption.

[5] This result reminds of a more classical definition of the intensity.

Assumption 3.3.1. *Let N be a doubly stochastic Poisson process with the intensity function λ on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. We assume that*

$$\mathcal{F}_t^N \vee \mathcal{F}_t^\lambda \subseteq \mathcal{F}_t, \quad \forall t. \tag{3.18}$$

As we shall see later, from Proposition 3.3.4, under Assumption 3.3.1, a doubly stochastic Poisson process with respect to $\mathcal{G}$ with the intensity function λ, admits λ as $\mathcal{G} \vee \mathcal{F}_t^N$-stochastic intensity.

Before proceeding, we need the following lemmas.

Lemma 3.3.2. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathcal{S} \subseteq \mathcal{A}$ be a σ-algebra and $X_0, X_1, \ldots, X_n$ be real random variables such that, conditionally on $\mathcal{S}$, are independent. Then*

$$\mathcal{L}\left(X_0 \,|\mathcal{S} \vee \sigma\left(X_1, \ldots, X_n\right)\right) = \mathcal{L}\left(X_0 \,|\mathcal{S}\right). \tag{3.19}$$

Proof. The random variables $X_0, X_1, \ldots, X_n$ are, conditionally on $\mathcal{S}$, independent, i.e., for all $n \geq 0$, for all $f_0, f_1, \ldots, f_n$ bounded borelian function

$$E\left[f_0(X_0) f_1(X_1) \ldots f_n(X_n) \,|\mathcal{S}\right] = E\left[f_0(X_0) \,|\mathcal{S}\right] E\left[f_1(X_1) \,|\mathcal{S}\right] \cdots E\left[f_n(X_n) \,|\mathcal{S}\right]. \tag{3.20}$$

Equality (3.19) can be rewritten as

$$E\left[f_0(X_0) \,|\mathcal{S} \vee \sigma\left(X_1, \ldots, X_n\right)\right] = E\left[f_0(X_0) \,|\mathcal{S}\right], \quad \forall f_0,$$

i.e., for all $C \in \mathcal{S}$, for all $g_1, \ldots, g_n$ bounded borelian function

$$E\left[f_0(X_0) g_1(X_1) \cdots g_n(X_n) \mathbf{1}_C\right] = E\left[E\left[f_0(X_0) \,|\mathcal{S}\right] g_1(X_1) \cdots g_n(X_n) \mathbf{1}_C\right]. \tag{3.21}$$

We obtain (3.21), and then the announced result, by using (3.20) and observing that

$$\begin{aligned}
E\left[f_0(X_0) g_1(X_1) \cdots g_n(X_n) \mathbf{1}_C\right] &= E\left[E\left[f_0(X_0) g_1(X_1) \cdots g_n(X_n) \,|\mathcal{S}\right] \mathbf{1}_C\right] \\
&= E\left[E\left[f_0(X_0) \,|\mathcal{S}\right] \prod_{i=1}^n E\left[g_i(X_i) \,|\mathcal{S}\right] \mathbf{1}_C\right] \\
&= E\left[\left(E\left[f_0(X_0) \,|\mathcal{S}\right] \mathbf{1}_C\right) E\left[\prod_{i=1}^n g_i(X_i) \,|\mathcal{S}\right]\right] \\
&= E\left[E\left[f_0(X_0) \,|\mathcal{S}\right] \mathbf{1}_C \prod_{i=1}^n g_i(X_i)\right],
\end{aligned} \tag{3.22}$$

where in the last step we have used that $\mathbf{1}_C E\left[f_0(X_0) \,|\mathcal{S}\right]$ is $\mathcal{S}$-measurable.

□

Lemma 3.3.3. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual conditions and $M(t)$ be an $\mathbb{F}$-martingale. If $\mathbb{A} = \{\mathcal{A}_t : t \in [0, T]\}$ is a filtration such that*

$$\mathcal{F}_t^M \subseteq \mathcal{A}_t \subseteq \mathcal{F}_t, \forall t, \tag{3.23}$$

then $M(t)$ is an $\mathbb{A}$-martingale.

Observe that Lemma 3.3.3 also extends to the local martingale.

Proof. It is easily seen that M is a $\mathbb{A}$-martingale. In fact

- M is $\mathbb{A}$-adapted, being $\mathcal{F}_t^M \subseteq \mathcal{A}_t$, for all t;
- $M(t)$ is integrable for all t, M being a $\mathbb{F}$-martingale;
- using (3.23) and M being a $\mathbb{F}$-martingale, we have for all $t_2 > t_1$

$$E\left[M(t_2) \,|\mathcal{A}_{t_1}\right] = E\left[E\left[M(t_2) \,|\mathcal{F}_{t_1}\right] |\mathcal{A}_{t_1}\right] = E\left[M(t_1) \,|\mathcal{A}_{t_1}\right] = M(t_1).$$

□

Now we can show[6] that Definition 3.3.5 coincides with the definition of doubly stochastic Poisson process as given in Duffie [8] with the following proposition.

[6]The definition of stochastic intensity is not uniform in the literature.

Proposition 3.3.4. *Let N be a doubly stochastic Poisson process with respect to $\mathcal{G}$ with intensity function λ. Then we have the following results.*

1. *For all s and $t > s$, conditional on $\mathcal{G} \vee \mathcal{F}_s^N$, the random variable $N(t) - N(s)$ has the Poisson distribution with parameter $\int_s^t \lambda(u)du$, i.e., for all $0 \leq s < t$ and all $k \geq 0$,*

$$P\left[N(t) - N(s) = k \left| \mathcal{F}_s^N \vee \mathcal{G}\right.\right] = e^{-\int_s^t \lambda(u)du} \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!}. \tag{3.24}$$

2. *If*

$$E\left[\int_0^t \lambda(s)ds\right] < \infty, \quad \forall t \tag{3.25}$$

then $M(t) = N(t) - \int_0^t \lambda(u)du$ is a $\mathcal{F}_t^N \vee \mathcal{G}$-martingale;

3. *If $\mathbb{H} = \{\mathcal{H}_t : t \in [0,T]\}$ is a filtration such that*

$$\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t \subseteq \mathcal{F}_t^N \vee \mathcal{G}, \quad \forall t \tag{3.26}$$

and λ is an $\mathbb{H}$-predictable process, then N admits λ as $\mathbb{H}$-stochastic intensity;

4. *If $\mathbb{A} = \{\mathcal{A}_t : t \in [0,T]\}$ is a filtration such that $\mathcal{A}_t$ is independent of $\mathcal{F}_t^N \vee \mathcal{G}$, for all t, then $M(t)$ is also a $\mathcal{H}_t \vee \mathcal{A}_t$-martingale.*

Proof. (Proposition 3.3.4). First we prove 1. From Lemma 3.3.2, taking $\mathcal{S} = \mathcal{G}$, $X_0 = N(t) - N(s)$ and $X_i = N(s_i) - N(s_{i-1}) = \Delta N(s_i)$, with $0 = s_0 \leq s_i \leq s_n \leq s$, for $i = 1, \ldots, n$, we obtain

$$P\left(N(t) - N(s) = k \left| \mathcal{G} \vee \mathcal{F}_s^N\right.\right) = P\left(N(t) - N(s) = k \left| \mathcal{G}\right.\right) = e^{-\int_s^t \lambda(u)du} \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!},$$

where we use that N is a doubly stochastic Poisson process with respect to $\mathcal{G}$, that

$$\sigma\left(X_1, \ldots, X_n\right) = \sigma\left(\Delta N(s_1), \ldots, \Delta N(s_n)\right),$$

and that

$$\sigma\left(\Delta N(s_1), \ldots, \Delta N(s_n) : 0 = s_0 \leq s_i \leq s_n \leq s\right) = \mathcal{F}_s^N. \tag{3.27}$$

Secondly we prove 2. From (3.24) we observe that

$$E\left[N(t) - N(s) \left| \mathcal{F}_s^N \vee \mathcal{G}\right.\right] = \int_s^t \lambda(u)du, \tag{3.28}$$

then by (3.25) we obtain

$$E\left[N(t) - \int_0^t \lambda(u)du \left| \mathcal{F}_s^N \vee \mathcal{G}\right.\right] = N(s) - \int_0^s \lambda(u)du, \tag{3.29}$$

i.e., $M(t) = N(t) - \int_0^t \lambda(u)du$ is a $\mathcal{F}_t^N \vee \mathcal{G}$-martingale.

Now we prove 3. By Proposition 3.3.1, it is sufficient to prove that

$$M(t) = N(t) - \int_0^t \lambda(s)ds \tag{3.30}$$

is a $\mathbb{H}$-martingale. By result in 2. we have that $M(t)$ is a $\mathcal{F}_t^N \vee \mathcal{G}$-martingale. Then from Lemma 3.3.3 we obtain that M is a $\mathbb{H}$-martingale, since (3.26) hold.

Finally we prove 4. Since $\mathbb{H}$ is a filtration such that (3.26) holds, we obtain that $\mathcal{A}_t$ is also independent of $\mathcal{F}_t^N \vee \mathcal{F}_t^\lambda$. Then by a property of conditional expectation we have that

$$E\left[M(t) \left| \mathcal{H}_s \vee \mathcal{A}_s\right.\right] = E\left[M(t) \left| \mathcal{H}_s\right.\right] = M(s), \tag{3.31}$$

where we have used that M is a $\mathbb{H}$-martingale (see property 3.). Thus $M(t)$ is a $\mathcal{H}_t \vee \mathcal{A}_t$-martingale.

We have proved the announced results.

□

Observe that by result in 3. we obtain as in Duffie [8], for all $0 \le s < t$,

$$E\left[N(t) - N(s) \left| \mathcal{F}_s^N \vee \mathcal{F}_t^\lambda \right.\right] = \int_s^t \lambda(u)du, \tag{3.32}$$

since $\mathcal{F}_s^N \vee \mathcal{F}_s^\lambda \subseteq \mathcal{F}_s^N \vee \mathcal{F}_t^\lambda \subseteq \mathcal{F}_s^N \vee \mathcal{G}$. Summarizing, the idea of the doubly stochastic assumption is that, conditional on λ, N is a Poisson process with intensity λ. In particular, from (3.24), we know for any time $t > s$, conditional on the σ-algebra $\mathcal{G} \vee \mathcal{F}_t^N$, generated by the events in $\mathcal{G} \cup \mathcal{F}_t^N$, the number $N(t) - N(s)$ of jumps (or arrivals) between s and t is distributed as a Poisson random variable with parameter $\int_s^t \lambda(u)du$. Before proceeding any further, now we want build an doubly stochastic Poisson process with respect to $\mathcal{G} \supseteq \mathcal{F}_\infty^\lambda$, with the stochastic intensity λ, through a Poisson process independent of the intensity λ. Let $\hat{N}(t)$ be a homogeneous Poisson process with intensity $\bar{\lambda} = 1$ independent of $\mathcal{G}$, then we define

$$N(t) = \hat{N}\left(\int_0^t \lambda(u)du\right). \tag{3.33}$$

Note that by construction $\hat{N}$ is independent of λ. Furthermore observe that

$$\mathcal{F}_t^N \subseteq \mathcal{F}_\infty^{\hat{N}} \vee \mathcal{F}_t^\lambda, \ \forall t. \tag{3.34}$$

Then we have that $N(t)$ is an doubly stochastic Poisson process as we shown in the following proposition.

Proposition 3.3.5. *Let $N(t)$ be defined as in (3.33). Then $N(t)$ is an doubly stochastic Poisson process with respect to $\mathcal{G}$, with the stochastic intensity λ.*

Proof. By Lemma A.1.1 in Appendix A with $\mathcal{M} = \mathcal{G} \supseteq \mathcal{F}_\infty^\lambda$, $\mathcal{A} = \mathcal{F}_\infty^{\hat{N}}$ and

$$\Psi(\omega) = \psi(\int_0^s \lambda(u)du, \int_0^t \lambda(u)du, \hat{N}(\cdot)) = \mathbf{1}_{\left\{\hat{N}\left(\int_0^t \lambda(u)du\right) - \hat{N}\left(\int_0^s \lambda(u)du\right) = k\right\}},$$

with $k \in \mathbb{R}$, we have that

$$\begin{aligned} P\left[N(t) - N(s) = k \left|\mathcal{G}\right.\right] &= P\left[\hat{N}\left(\int_0^t \lambda(u)du\right) - \hat{N}\left(\int_0^s \lambda(u)du\right) = k \left|\mathcal{G}\right.\right] \\ &= P\left[\hat{N}(x) - \hat{N}(y) = k\right]\Big|_{(x,y) = (\int_0^t \lambda(u)du, \int_0^s \lambda(u)du)} \\ &= e^{-\int_s^t \lambda(u)du} \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!}. \end{aligned} \tag{3.35}$$

Then by Definition 3.3.5 and Proposition 3.3.4, it is sufficient to prove that

$$P\left[N(t) - N(s) = k \left|\mathcal{G} \vee \mathcal{F}_s^N\right.\right] = e^{-\int_s^t \lambda(u)du} \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!}. \tag{3.36}$$

We apply Lemma 3.3.2, with $\mathcal{S} = \mathcal{G}$, $X_0 = N(t) - N(s)$ and $X_i = N(s_i) - N(s_{i-1}) = \Delta N(s_i)$, where $0 = s_0 \le s_i \le s_n \le s$, for $i = 1, \ldots, n$, and $\mathcal{A} = \mathcal{F}$. Taking into account that $\hat{N}$ is Poisson process and $N(t)$ satisfies (3.35), that $\sigma(X_1, \ldots, X_n) = \sigma(\Delta N(s_1), \ldots, \Delta N(s_n))$, and that

$$\sigma\left(\Delta N(s_1), \ldots, \Delta N(s_n) : 0 = s_0 \le s_i \le s_n \le s\right) = \mathcal{F}_s^N, \tag{3.37}$$

we obtain

$$P\left(N(t) - N(s) = k \left|\mathcal{G} \vee \mathcal{F}_s^N\right.\right) = P\left(N(t) - N(s) = k \left|\mathcal{G}\right.\right) = e^{-\int_s^t \lambda(u)du} \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!},$$

i.e., the announced result.

□

As we will see below, we model the death time of an individual as the first jump time T_1 of a nonexplosive counting process $N(t)$, i.e., the counting process N is a process that jumps for the first time when the individual dies. Furthermore, since the death time of an individual is finite, a good model should have the property that T_1 is a.s. finite. Finally, on the basis of demographic considerations, it could be desirable to assume that the death time of an individual is uniformly bounded, i.e., the model should have the property that $T_1 \le L$ a.s., for some constant L. To this end we state and prove the following results.

Theorem 3.3.6. *Under hypothesis of Proposition 3.3.4, let T_1 be the first jump time of $N(t)$. Then for each $0 \le s \le t$*

$$P\left(T_1 > t \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right) = \mathbf{1}_{\{T_1 > s\}} e^{-\int_s^t \lambda(u)du}. \tag{3.38}$$

If $\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t \subseteq \mathcal{G} \vee \mathcal{F}_t^N$, then

$$P\left(T_1 > t \left| \mathcal{H}_s \right.\right) = \mathbf{1}_{\{T_1 > s\}} E\left[e^{-\int_s^t \lambda(u)du} \left| \mathcal{H}_s \right.\right]. \tag{3.39}$$

Proof. Letting A be the event that $N(t) - N(s) = 0$, we have

$$\begin{aligned} P\left(T_1 > t \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right) &= E\left[\mathbf{1}_{\{N(s)=0\}} \mathbf{1}_{\{N(t)-N(s)=0\}} \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right] \\ &= \mathbf{1}_{\{N(s)=0\}} E\left[\mathbf{1}_A \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right] \\ &= \mathbf{1}_{\{T_1 > s\}} E\left[\mathbf{1}_A \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right]. \end{aligned}$$

Now, using (3.24) with $k = 0$, we obtain

$$\begin{aligned} E\left[\mathbf{1}_A \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right] &= E\left[\mathbf{1}_A \left| \mathcal{G} \vee \mathcal{F}_s^N \right.\right] \\ &= E\left[P\left(N_t - N_s = 0 | \mathcal{F}_s^N \vee \mathcal{G}\right)\right] \\ &= e^{-\int_s^t \lambda(u)du}. \end{aligned}$$

Finally (3.39) is an obvious consequence of (3.38) and the iterated conditional expectations property. □

We observe that

$$P\left(T_1 > t \left| \mathcal{H}_s \right.\right) = \mathbf{1}_{\{T_1 > s\}} P\left(T_1 > t \left| \mathcal{H}_s \right.\right),$$

then (3.39) is equivalent to

$$P\left(T_1 > t \left| \mathcal{H}_s \right.\right) = E\left[e^{-\int_s^t \lambda(u)du} \left| \mathcal{H}_s \right.\right], \qquad \text{on } \{T_1 > s\}, \tag{3.40}$$

(analogously for (3.38)).

Lemma 3.3.7. *Let T_1 be the first jump time of a doubly stochastic Poisson process $N(t)$ with intensity $\lambda(t)$.*

I *The following conditions are equivalent.*

- **I.1** *T_1 is finite a.s., i.e.,*

$$P\big(T_1 < \infty\big) = 1; \tag{3.41}$$

- **I.2** *The process $\lambda(t)$ satisfies the following property*

$$P\left(\int_{t_0}^\infty \lambda(u)du = \infty\right) = 1, \quad \textit{for a fixed } t_0 \ge 0. \tag{3.42}$$

- **I.3** *The process $\lambda(t)$ satisfies the following property*

$$P\left(\int_{t_0}^\infty \lambda(u)du = \infty\right) = 1, \quad \forall t_0 \ge 0. \tag{3.43}$$

II *For any deterministic constant $L < \infty$, the following conditions are equivalent.*

- **II.1** *T_1 in bounded above by L a.s., i.e.,*

$$P\big(T_1 \le L\big) = 1; \tag{3.44}$$

- **II.2** *The process $\lambda(t)$ satisfies the following property*

$$P\left(\int_{t_0}^L \lambda(u)du = \infty\right) = 1, \quad \textit{for a fixed } t_0 \in [0, L). \tag{3.45}$$

- **II.3** *The process $\lambda(t)$ satisfies the following property*

$$P\left(\int_{t_0}^L \lambda(u)du = \infty\right) = 1, \quad \forall t_0 \in [0, L). \tag{3.46}$$

Proof. The process $\lambda(t)$ being positive, conditions I.2 and I.3 are clearly equivalent. To prove the equivalence of I.1 and I.3 observe that, by relation (3.38), for each $0 \leq t_0 \leq t$

$$P(T_1 > t) = E\Big(E\Big(\mathbf{1}_{T_1>t}\Big|\mathcal{G} \vee \mathcal{F}_{t_0}^N\Big)\Big) = E\Big(\mathbf{1}_{T_1>t_0} E\Big(e^{-\int_{t_0}^t \lambda(u)du}\Big)\Big), \tag{3.47}$$

and

$$\begin{aligned} P(T_1 = \infty) &= \lim_{t\to\infty} P(T_1 > t) = \lim_{t\to\infty} E\Big(\mathbf{1}_{T_1>t_0} E\left(e^{-\int_{t_0}^t \lambda(u)du}\right)\Big) \\ &= E\Big(\mathbf{1}_{T_1>t_0} E\left(e^{-\int_{t_0}^\infty \lambda(u)du}\right)\Big). \end{aligned} \tag{3.48}$$

so that

$$P(T_1 < \infty) = 1 - P(T_1 = \infty) = 1 - E\Big(\mathbf{1}_{T_1>t_0} E\left(e^{-\int_{t_0}^\infty \lambda(u)du}\right)\Big).$$

Then by the above expression we accomplish the proof of part I, i.e.

$$P(T_1 < \infty) = 1 - E\Big(\mathbf{1}_{T_1>t_0} E\left(e^{-\int_{t_0}^\infty \lambda(u)du}\right)\Big) = 1$$

if and only if (3.43) holds.
The proof of part II being similar, we just observe that for each $0 \leq t_0 \leq t$

$$P(T_1 > L) = E\Big(E\Big(\mathbf{1}_{T_1>L}\Big|\mathcal{G} \vee \mathcal{F}_{t_0}^N\Big)\Big) = E\Big(\mathbf{1}_{T_1>t_0} E\Big(e^{-\int_{t_0}^L \lambda(u)du}\Big)\Big) = 0. \tag{3.49}$$

$\square$

As a final remark we note that by Proposition 3.3.5 and the previous Lemma 3.3.7, in order to model a death time τ as the first time of a doubly stochastic Poisson process, it is sufficient to have a strictly positive process $\lambda(t)$ and an exponential variable E_1, independent of $\mathcal{F}_\infty^\lambda$, and define $\tau = \inf\{t > 0 : \int_0^t \lambda(s)ds \geq E_1\}$. More generally, in a probability space $(\Omega, \mathcal{F}, P)$ endowed with a filtration $\mathbb{G} = \{\mathcal{G}_t\}$ it is sufficient to have a strictly positive, $\mathbb{G}$-adapted process $\lambda(t)$ and and exponential variable E_1, independent of $\mathcal{G} \supseteq \mathcal{G}_\infty$, and define τ as above, i.e., as the first time such that the integral $\int_0^t \lambda(s)ds$ reaches E_1.

3.4 Modelling mortality risk

In this section, we focus on mortality risk and modelling the survival function of the individual of a given population. In particular we present a model with a financial and mortality risk, where the interest rate $r(t)$ and the stochastic mortality intensity $\lambda(t)$ are dependent, but with uncorrelated driving noises.

In financial literature, for a long time, usually only the deterministic mortality intensity has been considered, while, more recently, the stochastic mortality intensity has been introduced using doubly stochastic Poisson processes. In many financial applications a useful assumption is that the stochastic intensity is an affine process. As already seen in Section 2.2.1, the convenience of adopting such processes in modelling the intensity is given by the key property of affine processes, i.e., the property (2.51).

Turning to the problem of modelling adequately the mortality dynamics, we will now use some of the mathematical tools presented in the previous section. We consider an individual aged x at time $t = t_0$, and model[7] her/his death time τ_x as the first jump time of a nonexplosive counting process $N(t)$, i.e., the counting process N is a process that jumps whenever the individual dies. Thus

$$\begin{cases} N(t) = 0 & \text{if} \quad t_0 \leq t < \tau_x \\ N(t) > 0 & \text{if} \quad t \geq \tau_x > t_0. \end{cases}$$

Moreover if we assume that $N(t)$ is a doubly stochastic Poisson process with respect to $\mathcal{G}$, with $\mathbb{H}$-stochastic intensity $\lambda_x(t)$, where $\mathcal{G} \supset \mathcal{F}_\infty^\lambda$ and $\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t \subseteq \mathcal{G} \vee \mathcal{F}_t^N$, then, according to (3.39), the (conditional) survival probability is given by

$$\bar{F}_x(t|t_0) = P\left(\tau_x > t\,|\mathcal{H}_{t_0}\right) = E\left[e^{-\int_{t_0}^t \lambda_x(u)du}\,|\mathcal{H}_{t_0}\right]. \tag{3.50}$$

[7] In this section we consider a counting process starting at time $t = t_0$ instead of $t = 0$. The appropriate modifications due to this fact are evident.

The similarity with the survival probability until time t for an individual aged x, expressed in terms of the mortality force, $\mu_x(t)$, is strong as we can see in (3.5). Nevertheless, comparing (3.5) and (3.50), we deduce[8] that $\mu_x(t) \neq E[\lambda_x(t)]$. We notice that, when t changes, the process λ_x describes the future mortality intensity for any age $x+t$ of an individual aged x at time t_0. In other words, our process λ_x captures the mortality intensity for a particular generation and a particular initial age. For notational convenience in the sequel we omit the initial age x and the intensity is denoted by $\lambda(t)$. Finally, the specification of intensity process $\lambda(t)$ is obviously crucial for the solution of equation (3.50).

Generally, contrary to the interest rate, in modelling the stochastic intensity, the non negativity of the model is necessary, since it is an intensity process. Furthermore, in the financial literature, the stochastic intensity is usually assumed independent of the riskless interest rate, while the interest rate growth may affect the active population mortality intensity (for instance, a large interest rate may diminish health care and prevention). Thus we now present a stochastic intensity model depending on the interest rate.

We take the CIR model for the interest rate $r(t)$, (see (2.50) of Section 2.2.1), and the stochastic intensity $\lambda(t) = \lambda^{(c)}(t)$, where

$$\begin{cases} d\lambda^{(c)}(t) = a_\lambda \left(b_\lambda - \lambda^{(c)}(t) + c\, r(t)\right) dt + \bar{\sigma}_\lambda \sqrt{\lambda^{(c)}(t)} dW^\lambda(t), \\ \lambda^{(c)}(t_0) = \lambda^{(c)} \end{cases} \tag{3.51}$$

where W^λ is a 1-dimensional Wiener process independent of W^r, a_λ, b_λ, $\bar{\sigma}_\lambda$ are strictly positive deterministic constants such that $2\,a_\lambda\, b_\lambda > \bar{\sigma}_\lambda^2$, c is a positive deterministic constant, and $P(\lambda^{(c)} > 0) = 1$. Observe that the process $\lambda^{(c)}(t)$ dependens on $r(t)$, in the sense that the drift of $\lambda^{(c)}(t)$ is a function of $r(t)$ (when $c > 0$), and the processes $r(t)$ and $\lambda(t) = \lambda^{(c)}(t)$ are uncorrelated since the driving noises W^r and W^λ are independent.

Furthermore, ***(i)*** $\lambda^{(c)}(t)$ is strictly positive (as should be for a mortality intensity process), ***(ii)*** $\tau^{(c)}$ is finite a.s., where $\tau^{(c)}$ is the first jump time of a doubly stochastic Poisson process with intensity $\lambda^{(c)}(t)$. Before proving this result (see Proposition 3.4.2) we consider the case $c = 0$

$$\begin{cases} d\lambda^{(0)}(t) = a_\lambda \left(b_\lambda - \lambda^{(0)}(t)\right) dt + \bar{\sigma}_\lambda \sqrt{\lambda^{(0)}(t)}\, dW^\lambda(t), \\ \lambda^{(0)}(t_0) = \lambda^{(0)} \end{cases} \tag{3.52}$$

The model $\lambda^{(0)}(t)$ is then a CIR process, which is the simplest positive model, (see Section ese.CIR.B). Moreover, in this case the processes $r(t)$ and $\lambda(t) = \lambda^{(0)}(t)$ are independent. Furthermore $\tau^{(0)}$ is a.s. finite, as shown in the following proposition.

Proposition 3.4.1. *Let $\tau^{(0)}$ be the first jump time of a doubly stochastic Poisson process with intensity $\lambda^{(0)}(t)$ given by (3.52). Then*

$$P\left(\int_{t_0}^{\infty} \lambda^{(0)}(u) du = \infty\right) = 1, \quad \forall t_0 \geq 0. \tag{3.53}$$

and

$$P\left(\tau^{(0)} < \infty\right) = 1. \tag{3.54}$$

Proof. We prove only (3.53), since (3.54) is equivalent to (3.53) (see Lemma 3.3.7). Since $\lambda^{(0)}(t)$ is a CIR model, by the property of affine processes, i.e. (2.51), we have that

$$E_{t_0,\lambda}\left(e^{-\int_{t_0}^T \lambda^{(0)}(u) du}\right) = e^{\psi_\lambda^0(T-t_0) + \psi_\lambda(T-t_0)\lambda},$$

where the functions $\psi_\lambda(s)$ and $\psi_\lambda^0(s)$ solve the equations

$$\dot{\psi}_\lambda^0(s) = a_\lambda\, b_\lambda \psi_\lambda(s), \tag{3.55}$$

$$\dot{\psi}_\lambda(s) = -a_\lambda \psi_\lambda(s) + \frac{\bar{\sigma}_\lambda^2}{2}\psi_\lambda^2(s) - 1 \tag{3.56}$$

with the initial conditions $\psi_\lambda^0(0) = 0$, $\psi_\lambda(0) = 0$, and are given by[9]

$$\psi_\lambda(s) = \frac{1 - e^{\alpha_\lambda s}}{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}}, \tag{3.57}$$

$$\psi_\lambda^0(s) = -\frac{2\, a_\lambda b_\lambda}{\bar{\sigma}_\lambda^2} \ln\left(\frac{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}}{\alpha_\lambda}\right) + \frac{a_\lambda b_\lambda}{\beta_\lambda}\, s \tag{3.58}$$

[8] By applying the Jensen inequality to (3.50), and comparing with (3.5)

[9] The process $\lambda^{(0)}$ being a CIR model, to get ψ_λ^0 and ψ_λ we can use (2.53) and (2.54), with a_λ, b_λ, $\bar{\sigma}_\lambda$, α_λ, β_λ, γ_λ instead of a_r, b_r, $\bar{\sigma}_r$ α_r, β_r, γ_r.

with

$$\alpha_\lambda = -\sqrt{a_\lambda^2 + 2\,\bar\sigma_\lambda^2}, \quad \beta_\lambda = \frac{\alpha_\lambda - a_\lambda}{2}, \quad \gamma_\lambda = \frac{\alpha_\lambda + a_\lambda}{2}. \tag{3.59}$$

Since

$$\lim_{s\to\infty} \psi_\lambda(s) = \lim_{s\to\infty} \frac{1 - e^{\alpha_\lambda s}}{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}} = \frac{1}{\beta_\lambda}, \tag{3.60}$$

$$\begin{aligned}\lim_{s\to\infty} \psi_\lambda^0(s) &= \lim_{s\to\infty} \left[-\frac{2\,a_\lambda b_\lambda}{\bar\sigma_\lambda^2} \ln\left(\frac{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}}{\alpha_\lambda}\right) + \frac{a_\lambda b_\lambda}{\beta_\lambda}\, s\right] \\ &= -\frac{2\,a_\lambda b_\lambda}{\bar\sigma_\lambda^2} \ln\left(\frac{\beta_\lambda}{\alpha_\lambda}\right) - \lim_{s\to\infty} \frac{2\,a_\lambda b_\lambda}{\sqrt{a_\lambda^2 + 2\,\bar\sigma_\lambda^2} + a_\lambda}\, s = -\infty, \end{aligned} \tag{3.61}$$

we can conclude that

$$\lim_{T\to\infty} E_{t_0,\lambda}\left(e^{-\int_{t_0}^T \lambda^{(0)}(u)du}\right) = \lim_{T\to\infty} e^{\psi_\lambda^0(T-t_0) + \psi_\lambda(T-t_0)\lambda} = 0$$

so that

$$E_{t_0,\lambda}\left(e^{-\int_{t_0}^\infty \lambda^{(0)}(u)du}\right) = 0,$$

i.e., the process $\lambda^{(0)}(t)$ satisfies the property (3.53).

□

We now turn to the case $c > 0$.

Proposition 3.4.2. *Let $\lambda^{(c)}(t)$ be a process with dynamics given by (3.51), with a_λ, b_λ, c and $\bar\sigma_\lambda$ strictly positive deterministic constants such that $2\,a_\lambda\,b_\lambda > \bar\sigma_\lambda^2$. Then*

$$\lambda^{(c)}(t) > 0 \qquad a.s. \tag{3.62}$$

and

$$P\left(\int_{t_0}^\infty \lambda^{(c)}(u)du = \infty\right) = 1, \quad \forall t_0 \geq 0. \tag{3.63}$$

As a consequence, if $\tau^{(c)}$ is the first jump time of a doubly stochastic Poisson process with intensity $\lambda^{(c)}(t)$, then

$$P\big(\tau^{(c)} < \infty\big) = 1. \tag{3.64}$$

Observe that (3.51) is a particular case of the model studied by Deelstra and Delbaen [7]. In [7] the authors suggest that extending comparison results as in Karatzas and Shreve [16], it is easy to check that the solution of (3.51) remains positive a.s., i.e., $P(\lambda^{(c)}(t) \geq 0) = 1$. The slightly stronger condition (3.62) relies on the following lemma.

Lemma 3.4.3. *Let $\lambda^{(c)}(t)$ be a process with dynamics given by (3.51), and $\lambda^{(0)}(t)$ be the CIR process given by (3.52). If*

$$P\left(\lambda^{(0)} \leq \lambda^{(c)}\right) = 1, \tag{3.65}$$

then

$$P\left(\lambda^{(0)}(t) \leq \lambda^{(c)}(t)\right) = 1, \quad t \geq t_0. \tag{3.66}$$

Proof of Lemma 3.4.3. Let $h(s) = \sqrt{s}$ be the function in point 2. of Theorem B.2.1. Then condition (B.23) is satisfied, indeed

$$|\sqrt{y} - \sqrt{x}| \leq \sqrt{|y-x|} \tag{3.67}$$

and $h(s)$ is a strictly increasing function with $h(0) = 0$ and

$$\int_0^\infty \frac{1}{(\sqrt{u})^2} du = \infty. \tag{3.68}$$

Setting $b(t,x) = a_\lambda\,(b_\lambda - x)$, since

$$|b(t,x) - b(t,y)| = a_\lambda\; |x-y| \tag{3.69}$$

condition (B.24) holds. Therefore, since $r(t) > 0$ a.s., we have that also the condition 5. is satisfied. Then by comparison Theorem B.2.1 with $X^1(t) = \lambda^{(0)}(t)$ and $X^2(t) = \lambda^{(c)}(t)$, we obtain (3.66).

□

Proof of Proposition 3.4.2. Observe that $P(\lambda^{(0)} > 0) = 1$ and the condition $2\,a_\lambda\,b_\lambda > \bar{\sigma}_\lambda^2$ ensures that $P\left(\lambda^{(0)}(t) > 0\right) = 1$, for all $t \geq t_0$. Then by Lemma 3.4.3 we can conclude that

$$P\left(\lambda^{(c)}(t) \geq \lambda^{(0)}(t) > 0\right) = 1, \quad \forall t \geq t_0. \tag{3.70}$$

Furthermore, since (3.66) holds, and $\lambda^{(0)}(t)$ satisfies (3.53), we have that the mortality intensity $\lambda^{(c)}(t)$ also satisfies the same property

$$P\Big(\int_{t_0}^{\infty} \lambda^{(c)}(u)du = \infty\Big) = 1, \tag{3.71}$$

i.e., $\tau^{(c)}$ is finite a.s. (see Lemma 3.3.7).

□

Chapter 4

Financial and mortality risk models

4.1 Introduction: The Longevity Bond

Longevity bonds are the first financial products to offer longevity protection by hedging the trend in longevity. Longevity bonds are needed because lifetime has been constantly increasing (medical improvements, better life standards, etc) and so there is a demand for instruments hedging the longevity risk, i.e., the risk that members of some reference population might live longer, on average, than anticipated, for example, in the life companies' mortality tables (assuming constant longevity can lead to a bankrupt of a pension plan or a life insurer). The uncertainty of longevity projections is illustrated by the fact that life expectancy for men aged 60 is more than 5 years' longer in 2005 than it was anticipated to be in mortality projections made in the 1980 (we refer to Hardy [13]).

To meet this demand, the Capital markets offer longevity bonds with coupons depending on the survival rate of a given population. They can be used to hedge a big portion of the longevity risk. The longevity bonds can take a large variety of forms which can vary enormously in their sensitivities to longevity shocks. This is the ideal asset for hedging the longevity risk of a pension fund. In fact, while the population which subscribed to the fund increases its longevity, the fund risks to have to pay pensions for longer and longer period. Nevertheless, the increasing in longevity also means a lower decreasing rate in the longevity bond coupons. In this way, the higher pensions can be faced through the less decreasing coupons.

Longevity bonds were first proposed by Blake and Burrows [3], and the first operational mortality-linked bond appeared in 2003. A second mortality-linked bond was announced in 2004, the longevity bond offered by European Investment Bank (EIB) and BNP Paribas (although it failed to come to market). In November 2004, the European Investment Bank (EIB) unveiled plans to issue the first longevity bond that offers a partial longevity risk hedge to UK pension schemes and life insurers. For the longevity expertise and reinsurance capacity, the EIB relies on PartnerRe[1], while the financial component of the longevity bond is managed by the BNP Paribas.

In Azzopardi-BNP Paribas [10], a longevity bond is defined as an asset paying a coupon which is strictly proportional to the survival rate of a given population taken in a given moment. As its name suggests, the survival rate is the proportion of some initial reference population aged x at time t who are still alive at some future time s, with $s > t$. We will refer to it as the BNP-Paribas longevity bond. The main characteristics of this bond are therefore:

- The bond was designed to be a hedge to the holder.
- The issuer gains if the survival rate is lower than anticipated (and conversely, the buyer gains if the survival rate is higher than anticipated).
- The bond is a hedge against a portfolio dominated by annuity (rather than life insurance) policies.
- The bond is designed to protect the holder against any unanticipated improvement in mortality up to the maturity date of the bond.
- The value of the survival rate between t and s (with $s > t$) involves a single national survivor index.
- All coupons are at risk longevity shocks, more precisely, the coupon payments are directly proportional to the survival rate.

[1] PartnerRe Ltd. (PartnerRe) is an international reinsurance group. The Company provides reinsurance on a worldwide basis through its wholly owned subsidiaries, Partner Reinsurance Company Ltd. (Partner Reinsurance), Partner Reinsurance Europe Limited (PartnerRe Europe) and Partner Reinsurance Company of the United States (PartnerRe U.S.).

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space and we denote by τ^j, with $j = 1, \dots, m$, the death time of the jth element of the given population, m being the size of the population. We take the following assumptions.

Assumption 4.1.1. *We assume that there exists a strictly positive process λ and a σ-algebra $\mathcal{G}$, with $\mathcal{G} \supset \mathcal{F}^{\lambda}_{\infty}$, such that*

- *τ^j, $j = 1, \dots, m$, are, conditionally on $\mathcal{G}$, independent and identically distributed random variables on $(\Omega, \mathcal{F}, P)$;*
- *τ^j is, accordingly to Section 3.2 and 3.4, the first jump time of a doubly stochastic Poisson process $N^j(t)$ with respect to $\mathcal{G}$ with the intensity function[2] $\lambda(t)$;*

Finally, and without loss of generality, we assume that

- *there exists a random time τ such that*

$$\mathcal{L}^P\left(\tau^j \,|\mathcal{G}\right) = \mathcal{L}^P\left(\tau \,|\mathcal{G}\right), \quad j = 1, \dots, m. \tag{4.1}$$

In the sequel the process λ is referred to as the stochastic mortality intensity. Observe that τ^j, for $j = 1, \dots, m$,share the same stochastic intensity $\lambda(t)$, and that

$$\mathcal{L}^P\left(\tau^1, \dots, \tau^m \,|\mathcal{G}\right) = \left(\mathcal{L}^P\left(\tau \,|\mathcal{G}\right)\right)^m, \quad j = 1, \dots, m. \tag{4.2}$$

In the next sections we present the generalities of the financial-mortality risk models and face the problem of modelling an arbitrage free family of zero coupon longevity bond price processes, respectively.

Finally, in the last section, we present another type of mortality-linked bonds, i.e., we take into account a zero coupon longevity bond, defined as a financial security whose single payout occurs at maturity T if holder is alive at time T. In this setting the payment at the time of maturity, known as the principal value or face value, equals one if the holder is alive at time T, else zero. In this case we present the problem of modelling an arbitrage free family of zero coupon longevity bond price processes.

In the last section, we introduce a new zero coupon longevity bond different from the Azzopardi-BNP Paribas [10], nevertheless, under suitable conditions (see Proposition 4.5.1), we obtain the arbitrage free price process. The latter one coincides with the price process of the longevity bond under a special condition (see Assumption 4.5.4). In the latter case the new bond is an alternative bond with respect to the previous one: indeed, as it is easy to see, if both bonds are traded in the market, then there are arbitrage opportunities.

We refer to Azzopardi-BNP Paribas [10], Menoncin [18] for the longevity bond theory, and we follow the methodological approach taken through the use and the construction of locally riskless portfolios as in Björk [2] for finding the arbitrage-free price process.
The death times are modelled as the first jump times of a doubly stochastic Poisson process, as is usual in the literature (see, e.g., Biffis [1], and Luciano and Vigna [17]).

4.2 Financial and mortality risk

In Chapter 2, we have studied the simplest possible incomplete market, namely a market where the only randomness comes from a scalar stochastic process, i.e., short rate r, which is not the price of a traded asset.

In the setting proposed here, we also consider a stochastic process λ representing the stochastic mortality intensity of the given population, and study a model with two non-priced underlying asset, i.e., r and λ.

In order to model the evolution of the stochastic mortality intensity, $\lambda(t)$, let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space (see Assumption 4.1.1). Furthermore, by Assumption 3.3.1, the filtration $\mathbb{F}$ must contains $\mathcal{F}^{\lambda}_t \vee \mathcal{F}^N_t$, where $N = \left(N^1, \dots, N^m\right)$.

To be consistent with Assumptions 1.2.1 and 1.2.2, as in Chapter 2, in the sequel, we assume that

$$\bar{\mathbb{F}}^{W^r} \subset \mathbb{F}, \tag{4.3}$$

where $\bar{\mathbb{F}}^{W^r}$ denotes the augmented filtration associated to the process W^r. Observe that if $\mathcal{F}^{\lambda}_t \subset \mathcal{F}^{\lambda}_{t_0} \vee \mathcal{F}^{W^{\lambda}}_t$, then we can take $\mathcal{F}_t = \mathcal{F}^{\lambda}_{t_0} \vee \mathcal{F}^{r}_{t_0} \vee \mathcal{F}^{W^z}_t$.

[2] We recall that in the literature the intensity is usually denoted by $\lambda_x(t)$ when x is the age of each member of the population, but for notational convenience we omit the initial age x.

As already discussed in Chapter 3, the crucial point is the filtration with respect to which the process λ is a stochastic mortality intensity, i.e., thanks to Assumption 4.1.1 on the death time τ^j,

$$M^j(t) = N^j(t) - \int_0^t \lambda(s)ds, \ t \geq 0, \quad j = 1, \ldots, m$$

is a martingale. The same holds for the stopped martingales

$$M^{\tau^j}(t) = M^j(t \wedge \tau^j) \left(= 1_{\{\tau^j \leq t\}} - \int_0^{t\wedge\tau^j} \lambda(s)ds \right), \quad j = 1, \ldots, m. \tag{4.4}$$

Usually, the stochastic intensity is considered with respect to a filtration $\mathbb{H}$ satisfying the usual conditions and such that $\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t \subseteq \mathcal{F}_t \subseteq \mathcal{F}_t \vee \mathcal{G}$, for all $t \in [0,T]$ and $j = 1, \ldots, m$. For example we can take $\mathcal{H}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda$. By Proposition 3.3.4 we know that the $\mathbb{H}$- stochastic intensity is still λ. Let us formulate this as a formalized assumption.

Assumption 4.2.1. *We assume that*

$$\mathcal{F}_t^N \vee \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \subseteq \mathcal{F}_t, \quad \forall t \in [0,T], \quad j = 1, \ldots, m. \tag{4.5}$$

As in Chapter 2, a natural starting point is to give an a priori specification of the dynamics of r and λ. We examine the general case of the riskless interest rate $r(t)$ of Chapter 2, evolving as in (2.1), while we consider the stochastic mortality intensity $\lambda(t)$ evolving accordingly to

$$d\lambda(t) = \mu^\lambda(t)dt + \sigma^\lambda(t)dW^\lambda(t), \tag{4.6}$$

where W^λ is a 1-dimensional Wiener process independent of W^r. The latter assumption implies that the processes $r(t)$ and $\lambda(t)$ are uncorrelated, nevertheless in the sequel the stochastic mortality intensity λ is not assumed to be independent of r, in the sense that $\mu^\lambda(t)$ and $\sigma^\lambda(t)$ may depend also on $r(t)$: similarly to Section 1.2 (see Assumption 1.2.4) we assume that

$$\mu^\lambda(t) = \hat{\mu}^\lambda(t, r(t), \lambda(t)), \tag{4.7}$$
$$\sigma^\lambda(t) = \hat{\sigma}^\lambda(t, r(t), \lambda(t)), \tag{4.8}$$

for some measurable deterministic functions $\hat{\mu}^\lambda$ and $\hat{\sigma}^\lambda$ such that the conditions for existence of a unique solution are verified (see (2.2) and (2.3)). In the sequel, for the sake of simplicity, we denote $z = (r, \lambda)'$. According to (2.1), (2.2), (2.3), and (4.6), (4.7), (4.8), the dynamics of z are given by

$$dz(t) = \mu^z(t)dt + \Sigma^z(t)dW^z(t), \tag{4.9}$$

where $W^z = \left(W^r, W^\lambda\right)'$ is a 2-dimensional Wiener process, and

$$\mu^z(t) = \tilde{\mu}^z(t, z(t)) = \begin{pmatrix} \hat{\mu}^r(t, r(t)) \\ \hat{\mu}^\lambda(t, z(t)) \end{pmatrix}, \tag{4.10}$$

$$\Sigma^z(t) = \tilde{\Sigma}^z(t, z(t)) = \begin{pmatrix} \hat{\sigma}^r(t, r(t)) & 0 \\ 0 & \hat{\sigma}^\lambda(t, z(t)) \end{pmatrix}. \tag{4.11}$$

Fixing $z(s) = \bar{z}$, we will denote the unique solution $z(t)$ of (4.9) also by $z(t; s, \bar{z})$, $t \geq s$. We observe that if (4.10) and (4.11) are assumed, then $z(t; s, \bar{z})$ is an Itô diffusion, which satisfies the important Markov property with respect to the filtration $\mathbb{F}$ (shortly $z(t; s, \bar{z})$ is an $\mathcal{F}_t$-Markov process), i.e., for all Borel measurable, bounded functions f, we have

$$E\left[f(z(t')|\mathcal{F}_t\right] = E\left[f(z(t'))|z(t)\right] = g\left(z(t)\right) \tag{4.12}$$

for fixed t, t' such that $s \leq t \leq t'$, with[3] $g(y) := E\left[f(z(t; s, y))\right]$, $y \in \mathbb{R}^2$. Note that since $\mathcal{F}_t^z \subset \mathcal{F}_t$ this implies that $z(t)$ is also a $\mathcal{F}_t^z$-Markov process.

[3] In the sequel the equalities analogous to (2.5) and (2.6) hold, substituting $r(t)$ with $z(t)$, i.e.,

$$E\left[f(z(t; s, \bar{z}))\right] = E_{s,\bar{z}}\left[f(z(t))\right],$$

and

$$g(y) = E_{s,y}\left[f(z(t))\right].$$

More in general we will use the same kind of notation for functionals of the trajectory $z(t; s, \bar{z})$, $t \geq s$.

In Chapter 2 we have studied a market where the only randnomness comes from a stochastic process $r(t)$, which is not the price of a traded asset, and we have discussed the problems which arise when pricing derivatives written in terms of the underlying process $r(t)$. In our setting there are two non priced underling assets, $r(t)$ and $\lambda(t)$, and we will discuss the term structure of derivatives written in terms of these two underlying processes. To this aim, we assume that $r(t)$ and $\lambda(t)$ are two objects given a priori, and the exogenously given assets are the money account with price process G (see Definition 1.2.4), and a benchmark bond, defined as in Section 2.2 (see Remark 2.2.1) plus a new benchmark longevity bond to be define in the next sections. Apart from the assumptions on the benchmark longevity bond, we summarize the assumptions as follows.

Assumption 4.2.2. *The only objects which are a priori given are the following.*

- *An empirically observable 2-dimensional stochastic process $z = (r, \lambda)'$ with dynamics given by (4.9), (4.10) and (4.11). Note that we assume that r and λ are not price processes of any traded asset.*
- *A money market account $G(t)$, i.e., with $dG(t) = r(t)G(t)dt$, as in Definition 1.2.4.*
- *Fixed $T = T_0$, a benchmark bond $B(t, T_0)$, defined as in Section 2.2 (see Remark 2.2.1), whose price process is given by (2.7).*

4.3 Zero Coupon Longevity Bond: the term structure equation

In Azzopardi-BNP Paribas [10], a longevity bond is defined as an asset paying a coupon which is strictly proportional to the (cumulative) survival rate of a population taken in a given moment.

As an example we recall that for the BNP-Paribas longevity bond a population of Welsh males, all with the same age[4] x.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space and Assumption 4.1.1 holds. Let τ^j, with $j = 1, \ldots, m$, the death time of the jth element of the given population, m being the size of the population, and τ the random time such that (4.1) holds. We consider

$$\frac{\sum_{i=1}^m 1_{\{\tau^i > s\}}}{\sum_{j=1}^m 1_{\{\tau^j > t\}}}, \tag{4.13}$$

and refer to it as the survival rate between t and s (with $s > t$), given by the number of survived people in s with respect to the number of survived people in t, in the given population.

By BNP-Paribas longevity bond definition, a zero coupon longevity bond is a financial security that pays, at time T, the value of the survival rate, given by (4.13) for $s = T$. The price process at time t of such longevity bond, with maturity T, is denoted by $L_{(m)}(t, T)$. The aim is to prove that, under suitable conditions, the price $L_{(m)}(t, T)$ converge to a price $L(t, T)$ (see the subsequent formula (4.18)). The latter price is the price of the longevity bond considered here. From equation (4.18) one could get the term structure, nevertheless we will get the term structure by the same method as in Chapter 2.

Now we will show that this bonds value can be well approximated by a T-zero coupon bond with the addition that the holder of longevity bonds must pay costs due to the longevity risk.

From the risk-neutral pricing formula we have that

$$L_{(m)}(t, T) = E^Q \left[\frac{\sum_{i=1}^m 1_{\{\tau^i > T\}}}{\sum_{j=1}^m 1_{\{\tau^j > t\}}} e^{-\int_t^T r(s)ds} \big| \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \right], \tag{4.14}$$

where Q is a risk neutral measure.

As a preliminary result, in the next proposition we will show that, conditionally on $\mathcal{G}$, the joint law of τ_j, $j = 1, ..., m$ is the same under P and Q, under the assumption that for every time T, there exists a $\mathcal{G}$-measurable random variable such that the risk-neutral measure restricted to $\mathcal{F}_T$, i.e. $Q|_{\mathcal{F}_T}$, is given by

$$dQ|_{\mathcal{F}_T} = \mathcal{Z}(T)\, dP|_{\mathcal{F}_T}. \tag{$*$}$$

To be concrete note that the above assumption is satisfied if

$$\mathcal{Z}(T) = e^{-\int_0^T \xi_z(t) dW^z(t) - \frac{1}{2}\int_0^T \xi_z' \xi_z(t) dt}, \tag{$**$}$$

for some 2-dimensional $\mathcal{G}$-measurable process $\xi_z(t)$. In its turn the previous condition holds when $\xi_z(t)$ is $\mathcal{F}_t^r \vee \mathcal{F}_t^\lambda$-adapted and

$$\mathcal{G} \supset \mathcal{F}_\infty^{W_r} \vee \mathcal{F}_\infty^{W_\lambda} \vee \sigma(\lambda_0) \vee \sigma(r_0) \supset \mathcal{F}_\infty^r \vee \mathcal{F}_\infty^\lambda. \tag{4.15}$$

[4] In the first population considered by BNP-Paribas in 2003, the age was $x = 65$.

Proposition 4.3.1. *Assume that* $(*)$ *holds. Then, conditional to* $\mathcal{G}$*, the random variables* τ^j*,* $j = 1, \dots, m$*, have the same joint law under* P *and* Q*.*

Proof. In order to prove it, first of all we introduce the processes $H_j(t) = 1_{\tau_i > t}$. The aim is to prove that for every $I \subset \{1, ..., m\}$ and $E_i \in \mathcal{F}_\infty^{H_i}$

$$E^Q \left[\prod_{i \in I} 1_{\{E_i\}} | \mathcal{G} \right] = E^P \left[\prod_{i \in I} 1_{\{E_i\}} | \mathcal{G} \right] = \prod_{i \in I} E^P \left[1_{\{E_i\}} | \mathcal{G} \right],$$

where the last equality is due to Assumption 4.1.1.
Since $\mathcal{F}_\infty^{H_1, \dots, H_m} = \vee_{T \geq 0} \mathcal{F}_T^{H_1, \dots, H_m}$, it is sufficient to prove that the previous equality holds for every T, $I \subset \{1, ..., m\}$ and $E_i \in \mathcal{F}_T^{H_i}$. For notational convenience we will considered only $|I| \leq 2$, but the same proof holds true if we consider $|I| \geq 2$: Let E_i and E_j be such that $E_i \in \mathcal{F}_T^{N_i}$ and $E_j \in \mathcal{F}_T^{N_j}$. By the generalized Bayes formula,

$$E^Q \left[1_{\{E_i\}} 1_{\{E_j\}} | \mathcal{G} \right] = \frac{E^P \left[\mathcal{Z}(T) \, 1_{\{E_i\}} 1_{\{E_j\}} \middle| \mathcal{G} \right]}{E^P \left[\mathcal{Z}(T) \, | \mathcal{G} \right]} \tag{4.16}$$

Since $\mathcal{Z}(T)$ is $\mathcal{G}$-measurable, from (4.16) we have that

$$E^Q \left[1_{\{E_i\}} 1_{\{E_j\}} | \mathcal{G} \right] = E^P \left[1_{\{E_i\}} 1_{\{E_j\}} | \mathcal{G} \right]. \tag{4.17}$$

Note that, when $i = j$, the expression (4.17) implies that

$$E^Q \left[1_{\{E_i\}} \middle| \mathcal{G} \right] = E^P \left[1_{\{E_i\}} \middle| \mathcal{G} \right],$$

while if $i \neq j$

$$\begin{aligned} E^Q \left[1_{\{E_i\}} 1_{\{E_j\}} \middle| \mathcal{G} \right] &= E^P \left[1_{\{E_i\}} \middle| \mathcal{G} \right] E^P \left[1_{\{E_j\}} \middle| \mathcal{G} \right] \\ &= E^Q \left[1_{\{E_i\}} \middle| \mathcal{G} \right] E^Q \left[1_{\{E_j\}} \middle| \mathcal{G} \right]. \end{aligned}$$

□

We are now in a position to prove that the price process in the original definition of the BNP-Paribas longevity bond, (see Azzopardi-BNP Paribas [10]), we can be well approximated by a price process $L_\infty(t, T)$, called a limit zero coupon BNP-Paribas longevity bond, given by

$$L_\infty(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} e^{-\int_t^T \lambda(u) du} | \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \right]. \tag{4.18}$$

Observe that, by Assumption 4.2.2, the process (r, λ) is a Markovian diffusion, and therefore $L_\infty(t, T)$ is a function of $z(t) = (r(t), \lambda(t))$.

Proposition 4.3.2. *Let* $L_{(m)}(t, T)$ *and* $L_\infty(t, T)$ *be defined by* (4.14) *and* (4.18)*, respectively. Assume the same conditions of Proposition 4.3.1, together with condition* (4.15)*. Then the sequence* $L_{(m)}(t, T)$ *converges a.s. to* $L_\infty(t, T)$*.*

Proof. Using the iterated conditional expectations property, we obtain

$$\begin{aligned} L_{(m)}(t, T) &= E^Q \left[\frac{\sum_{i=1}^m 1_{\{\tau^i > T\}}}{\sum_{j=1}^m 1_{\{\tau^j > t\}}} e^{-\int_t^T r(s) ds} | \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \right] \\ &= E^Q \left[E^Q \left[\frac{\sum_{i=1}^m 1_{\{\tau^i > T\}}}{\sum_{j=1}^m 1_{\{\tau^j > t\}}} e^{-\int_t^T r(s) ds} | \mathcal{G} \right] | \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \right] \\ &= E^Q \left[e^{-\int_t^T r(s) ds} E^Q \left[\frac{\sum_{i=1}^m 1_{\{\tau^i > T\}}}{\sum_{j=1}^m 1_{\{\tau^j > t\}}} | \mathcal{G} \right] | \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \right] \end{aligned} \tag{4.19}$$

where $\mathcal{G} \supset \mathcal{F}_\infty^\lambda \vee \mathcal{F}_\infty^r$.

By Assumption 4.1.1, the random variables τ^j are, under P, independent and identically distributed, conditionally on $\mathcal{G}$ and they are, conditionally on $\mathcal{G}$, independent copies of a random time τ. As we have seen in the proof of Proposition 4.3.1, the same holds under Q, and furthermore $\mathcal{L}^Q(\tau_j|\mathcal{G}) = \mathcal{L}^P(\tau_j|\mathcal{G}) = \mathcal{L}^P(\tau|\mathcal{G})$ (see Proposition 4.3.1). Then, for all $u > 0$

$$E^Q\left[\frac{1}{m}\sum_{j=1}^m 1_{\{\tau^j>u\}}|\mathcal{G}\right] = P(\tau > u|\mathcal{G}) = P\left(N(u) = 0\,|\mathcal{G}\vee\mathcal{F}_0^N\right) = e^{-\int_0^u \lambda(s)ds}. \tag{4.20}$$

Then, the law of large numbers implies that the joint law

$$\mathcal{L}^Q\left[\frac{1}{m}\sum_{j=1}^m 1_{\{\tau^j>t\}}, \frac{1}{m}\sum_{i=1}^m 1_{\{\tau^i>T\}}|\mathcal{G}\right] \xrightarrow[m\to\infty]{} \delta_{(P[\tau>t|\mathcal{G}],P[\tau>T|\mathcal{G}])}, \tag{4.21}$$

where $\delta_{(x_1,x_2)}$ is the Dirac measure centered on (x_1, x_2).

Finally, from (4.19) and (4.21), we obtain that

$$\begin{aligned}
L_{(m)}(t,T) &= E^Q\left[e^{-\int_t^T r(s)ds}E^Q\left[\frac{\frac{1}{m}\sum_{j=1}^m 1_{\{\tau^j>T\}}}{\frac{1}{m}\sum_{j=1}^m 1_{\{\tau^j>t\}}}|\mathcal{G}\right]|\mathcal{F}_t^r\vee\mathcal{F}_t^\lambda\right] \\
&\xrightarrow[m\to\infty]{} E^Q\left[e^{-\int_t^T r(s)ds}\frac{e^{-\int_0^T \lambda(u)du}}{e^{-\int_0^t \lambda(u)du}}|\mathcal{F}_t^r\vee\mathcal{F}_t^\lambda\right] \\
&= E^Q\left[e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(u)du}|\mathcal{F}_t^r\vee\mathcal{F}_t^\lambda\right].
\end{aligned} \tag{4.22}$$

□

As already said, we will now get again expression (4.18) for the longevity bond, but starting a different set of assumptions and with the method of Chapter 2. The above result suggests that we can view the price process of the longevity bond as the price process of a T-zero coupon bond with a cost due to the longevity risk depending to the stochastic mortality intensity $\lambda(t)$. Formally, we have the following definition.

Definition 4.3.1 (Limit Zero Coupon BNP-Paribas Longevity Bond)**.** *A limit zero coupon BNP-Paribas longevity bond with maturity date T, also called a T-longevity bond, is a zero coupon bond, which guarantees the holder 1 dollar to be paid on the date T; furthermore, besides the riskless rate $r(t)$, there is a cost, due to the longevity risk, depending on the stochastic mortality intensity $\lambda(t)$. The (random) price at time t of such longevity bond, with maturity T, is denoted by $L(t,T)$, while the cumulative cost over the interval $[t,T]$ is denoted by $D(t,T)$.*

In order to simplify the notation, we will write "zero coupon longevity bond" instead of "limit zero coupon BNP-Paribas longevity bond".

The aim is focused on the problem of finding an arbitrage-free price process of a T-zero coupon longevity bonds starting by the Definition 4.3.1. As we will see, following the approach of Chapter 2, we will obtain the same formula (4.18).

Let us first describe briefly the set of general assumptions imposed on our financial market models.

Assumption 4.3.1. *In addition to the Assumptions 2.2.1 and 2.2.2 on the bond market, and the Assumption 4.2.2, we assume that there exists a market for zero coupon T-longevity bonds for every value of T.*

We thus assume that our market contains bonds and longevity bonds with all possible maturity times, but we stress that only benchmark assets, besides the riskless asset, is exogenously given.

In the other words, in this setting, the benchmark asset $B(t, T_0)$ is considered as the underlying asset whereas all the other bonds are uniquely determined in terms of the price of this benchmark and the dynamics of "underlying" $r(t)$, while all the longevity bonds are regarded as derivative of the "underlying" $z(t)$. Clearly in this market, since Assumptions 2.2.1 and 2.2.2 are valid, the relations (2.19) between two bonds with different maturities still hold. Analogously, we expect that similar relations hold for the longevity bonds. Therefore our main goal is broadly to investigate the relations among the price processes of longevity bonds with different maturities in an free-arbitrage market. To this aim we will use the approach of Chapter 2 and therefore we need a further assumption for the longevity bonds, which generalizes Assumption 2.2.1 for the T-bonds.

Assumption 4.3.2. *In addition to Assumptions 2.2.1, 2.2.2, and 4.2.2, we assume that the market for T bonds T-longevity bonds is arbitrage free. We assume furthermore that*

- *for every T, the price of T-longevity bonds has the form*

$$L(t,T) = \hat{L}^T(t, z(t)), \tag{4.23}$$

where $\hat{L}^T$ is a deterministic function of three real variables; furthermore we assume that $\hat{L}$ is smooth and positive;

- *the cumulative cost $D(t,T)$ of Definition 4.3.1 has the form*

$$dD(t,T) = L(t,T)\lambda(t)dt. \tag{4.24}$$

As in the case of T-bonds, $\hat{L}^T$ is a function of only three variables, namely t and $z = (r,\lambda)'$, whereas T is regarded as a parameter. Moreover, according to the condition $L(T,T) = 1$, we have the following boundary

$$\hat{L}^T(T,z) = 1 \quad \forall z, \tag{4.25}$$

where z denotes a generic outcome of the process $z(t)$.

By condition (4.23) of Assumption 4.3.2 and the multidimensional Itô formula, we have that the dynamics of $L(t,T)$ has the following form

$$\frac{dL(t,T)}{L(t,T)} = \mu_L(t,T)dt + \sigma_L(t,T)dW^z(t), \tag{4.26}$$

where $\sigma_L(t,T) = (\sigma_{L,r}(t,T), \sigma_{L,\lambda}(t,T))$ is a 2-dimensional row vector, and

$$\mu_L(t,T) = \hat{\mu}_L^T(t, z(t)) = \hat{\mu}_L^T(t, r(t), \lambda(t)), \tag{4.27}$$

$$\sigma_L(t,T) = \tilde{\sigma}_L^T(t, z(t)) = \left(\hat{\sigma}_{L,r}^T(t, z(t)), \hat{\sigma}_{L,\lambda}^T(t, z(t))\right), \tag{4.28}$$

for suitable deterministic function $\hat{\mu}^T$ and $\tilde{\sigma}^T$. As in Lemma 2.2.1, the functions $\hat{\mu}_L^T$ and $\tilde{\sigma}_L^T$ can be expressed by mean of the function $\hat{L}^T$ as shown in the following Lemma 4.3.3.

Lemma 4.3.3. *Under condition (4.23) of Assumption 4.3.2, the following equalities hold with probability* 1*, for all t and for every maturity time T.*

$$\hat{\mu}_L^T(t, z(t)) = \frac{\hat{L}_t^T(t, z(t)) + \hat{L}_z^T(t, z(t))\tilde{\mu}^z(t, z(t)) + \frac{1}{2} tr\left[\left(\tilde{\Sigma}^z\right)' \hat{L}_{zz}^T \tilde{\Sigma}^z\right](t, z(t))}{\hat{L}^T(t, z(t))} \tag{4.29}$$

$$\tilde{\sigma}_L^T(t, z(t)) = \frac{\hat{L}_z^T(t, z(t))\tilde{\Sigma}^z}{\hat{L}^T(t, z(t))} = \left(\frac{\hat{L}_r^T \hat{\sigma}^r}{\hat{L}^T}, \frac{\hat{L}_\lambda^T \hat{\sigma}^\lambda}{\hat{L}^T}\right)(t, z(t)), \tag{4.30}$$

where $tr[A]$ denotes the trace of a square matrix A, $\tilde{\mu}^z$, $\tilde{\Sigma}^z$ are the functions in (4.10),(4.11) respectively, and, where we have used the notation

$$\hat{L}_t^T(t,z) = \frac{\partial \hat{L}^T}{\partial t}(t,z), \; \hat{L}_z^T(t,z) = \left(\frac{\partial \hat{L}^T}{\partial r}(t,z), \frac{\partial \hat{L}^T}{\partial \lambda}(t,z)\right),$$

$$\hat{L}_{zz}^T(t,z) = \begin{pmatrix} \frac{\partial^2 \hat{L}^T}{\partial r^2}(t,z) & \frac{\partial \hat{L}^T}{\partial r \partial \lambda}(t,z) \\ \frac{\partial \hat{L}^T}{\partial \lambda \partial r}(t,z) & \frac{\partial^2 \hat{L}^T}{\partial \lambda^2}(t,z) \end{pmatrix}$$

In the sequel, when it is convenient, we will use the above notation.

Proof. After some reshuffling the multidimensional Itô formula gives us (4.29) and (4.30), similar to the proof of Lemma 2.2.1.

□

We can now apply the approach of Chapter 2 to this setting. As observed above, the a priori given market consists of the benchmark bond $B(t,T_0)$ and the money market account $G(t)$. Observe that in this market the number M of random sources equals two (the 2-dimensional Wiener process, W^z), while the number N of traded assets (besides $G(t)$) equals one. From Corollary 1.4.5, we may thus expect that the market is arbitrage-free, but not complete[5].

[5]Another way of seeing this problem appears if we try to price a certain T-longevity bond, using the technique in Chapter 2 generalized to the case where we have two underlying objects, i.e. $z = (r,\lambda)'$. This generalization is necessary since, as already observed, all the longevity bonds are regarded as derivatives of the underlying process z, in other words a zero coupon longevity bond can be thought of as a derivative on z.

As recalled in Section 2.2, introducing the benchmark bond $B(t, T_0)$ in our market, allows us to replicate[6] a zero coupon bond. Analogously, given the presence of random source W^λ, if we would consider one benchmark longevity bond, we might obtain a unique arbitrage-free price process also for longevity bonds.

Summarizing, in accordance to Chapter 2, we expect the following.

- We cannot say anything precise about the price process of any particular longevity bond, i.e., the price of a particular longevity bond will not be completely determined by the z-dynamics and the requirement that the market is arbitrage-free.

- Different longevity bonds will, however, have to satisfy certain internal consistency requirements in order to avoid arbitrage on the market.

- More precisely, since we now have on our market a 2-dimensional Wiener process, i.e. two random sources, we can specify, besides the benchmark bond $B(t, T_0)$, the price processes of one benchmark longevity bond. The price processes of all other longevity bonds will then be uniquely determined by the prices of this benchmark. For the sake of simplicity of notation we will assume that the maturity of the benchmark longevity bond is T_0, but one could take any other time $T_0' > t$.

The following central result is similar to Theorem 1.4.6, and extends Proposition 2.2.2 to this setting.

Proposition 4.3.4. *Under Assumption 4.3.2, fix one benchmark bond, $B(t, T_0)$, with price processes given by (2.9), (2.10) and (2.11) with $T = T_0$, and one benchmark longevity bond, $L(t, T_0)$ with price processes given by (4.26), (4.27) and (4.28), with $T = T_0$. Assume furthermore that $B(t, T_0)$ and $L(t, T_0)$ are such that*

$$\hat{\sigma}^{T_0}(t, r(t)) \neq 0, \quad \hat{\sigma}^{T_0}_{L,\lambda}(t, z(t)) \neq 0, \quad \forall t \leq T_0. \tag{4.31}$$

Then there exists a process $\xi_z(t) = (\xi_r(t), \xi_\lambda(t))'$ such that the relations

$$\sigma(t, T)\xi_r(t) = \mu(t, T) - r(t), \tag{4.32}$$

$$\sigma_L(t, T)\xi_z(t) = \mu_L(t, T) - \lambda(t) - r(t) \tag{4.33}$$

hold for all t a.s. and for every maturity time T.

Observe that the condition (4.31) is the mathematical formulation of the requirement that the family of benchmark derivatives is rich enough to span the entire derivative space, as we will see from proof of Proposition 4.3.4.

Observe that (4.32) and (4.33) are called market price of risk equations, and the process ξ_z is the market price (vector) of risk due to W^z. In particular the component $\xi_r(t)$ is given by (4.32), exactly as in (2.19) of Section 2.1, while the component $\xi_\lambda(t)$ is given by

$$\xi_\lambda(t) = \frac{\mu_L(t, T) - \lambda(t) - r(t) - \xi_r(t)\sigma_L^r(t, T)}{\sigma_L^\lambda(t, T)}, \tag{4.34}$$

i.e., analogously to ξ_r, the component ξ_λ has the dimension "risk premium per unit of λ-type volatility", so that we called ξ_λ the market price for the longevity risk due to W^λ, Finally, we observe that $\xi_z(t)$ can be expressed as a deterministic function of t and $z(t)$, namely

$$\xi_z(t) = \tilde{\xi}_z(t, z(t)) = \left(\hat{\xi}_r(t, z(t)), \hat{\xi}_\lambda(t, z(t))\right)',$$

(see (4.29)-(4.30)).

Proof of Proposition 4.3.4. We have already proved (4.32) in Section 2.1 (see Proposition 2.2.2), then we turn to prove (4.33). Fix one benchmark bond and one benchmark longevity bond with price process of the form

$$B(t, T_0) = \hat{B}^{T_0}(t, r(t)),$$

$$L(t, T_0) = \hat{L}^{T_0}(t, z(t)),$$

where $B(t, T)$ is a zero coupon bond of Section 2.1 (see (2.7) and Lemma 2.2.1). In order to simplify the notation, we will write T instead of T_0.

Considering a zero coupon longevity bond of maturity $S \neq T$, we have the corresponding equation for the S-longevity bond

$$dL(t, S) = L(t, S)\left[\mu_L(t, S)dt + \sigma_L(t, S)dW^z(t)\right]. \tag{4.35}$$

[6] We obtain a unique arbitrage-free price process since we can replicate our derivative.

where analogously to (4.27) and (4.28)

$$\mu_L(t,S) = \hat{\mu}_L^S(t,z(t)), \quad \sigma_L(t,S) = \tilde{\sigma}_L^S(t,z(t)) \tag{4.36}$$

We now form a portfolio based only on $B(t,T)$, $L(t,T)$, and $L(t,S)$, and as in the proof of Proposition 2.2.2, (see also Section 1.3) let $h(t) = (h_0(t), h_1(t), h_2(t), h_3(t))$ be the portfolio associated to $X = (X_0, X_1, X_2)$, where

$$X_0 = G(t), \quad X_1 = B(t,T), \quad X_2 = L(t,T), \quad X_3 = L(t,S), \tag{4.37}$$

and

$$h_0(t) = h^G(t) = 0, \quad (h_1(t), h_2(t), h_3(t)) = (h_T(t), h_T^L(t), h_S^L(t)), \tag{4.38}$$

i.e., nothing is invested in the bank or loaned by the bank. Similarly to Section 1.3, instead of specifying for each asset the absolute number of shares held, i.e. $h(t)$, it may be convenient to consider the corresponding relative portfolio $(U_T(t)U_T^L(t), U_S^L(t))$. Under Assumptions 2.2.2 and 4.3.2, setting $u(t) = (u_T(t), u_T^L(t), u_S^L(t))'$, by (1.16) and (1.17), we have

$$\begin{aligned} U_T(t) &= 1_{\{B(t,T)>0\}} u_T(t) = u_T(t) \\ U_T^L(t) &= 1_{\{L(t,T)>0\}} u_T^L(t) = u_T^L(t) \\ U_S^L(t) &= 1_{\{L(t,S)>0\}} u_S^L(t) = u_S^L(t), \end{aligned}$$

for the relative portfolio corresponding to $B(t,T)$, $L(t,T)$, and $L(t,S)$, with

$$u_T(t) + u_T^L(t) + u_S^L(t) = 1. \tag{4.39}$$

The dynamics of the value process for the corresponding self-financing portfolio (see (1.24)) are given by

$$\frac{dV(t)}{V(t)} = u_T(t)\frac{dB(t,T)}{B(t,T)} + u_T^L(t)\frac{dL(t,T) - dD(t,T)}{L(t,T)} + u_S^L(t)\frac{dL(t,S) - dD(t,S)}{L(t,S)}, \tag{4.40}$$

where the gain differential for the T-longevity bond is given by

$$dL(t,T) - dD(t,T) = L(t,T)\left(\left(\mu_L(t,T) - \lambda(t)\right)dt + \tilde{\sigma}_L(t,T)dW^z(t)\right), \tag{4.41}$$

and the same expressions applies to $dL(t,S) - dD(t,S)$ replacing T with S. The price processes for T-bond, (see (2.9), (2.10) and (2.11)), with respect to dW^z are given by

$$\frac{dB(t,T)}{B(t,T)}\Big(= \hat{\mu}^T(t,r(t))dt + \hat{\sigma}^T(t,r(t))dW^r(t)\Big) = \hat{\mu}^T(t,r(t))dt + \tilde{\sigma}^T(t,r(t))dW^z(t), \tag{4.42}$$

with $\tilde{\sigma}^T(t,r(t)) = (\hat{\sigma}^T(t,r(t)), 0)$ (here we are using the same notations of Proposition 2.2.2).

Then, inserting in (4.40) the dynamics (4.41) and (4.42) of the price processes involved we get

$$\begin{aligned} \frac{dV(t)}{V(t)} =& \Big[u_T(t)\hat{\mu}^T + u_T^L(t)\Big(\hat{\mu}_L^T - \lambda(t)\Big) + u_S^L(t)\Big(\hat{\mu}_L^S - \lambda(t)\Big)\Big]dt \\ &+ \Big[u_T(t)\tilde{\sigma}^T + u_T^L(t)\tilde{\sigma}_L^T + u_S^L(t)\tilde{\sigma}_L^S\Big]dW^z(t), \end{aligned}$$

where for the notational convenience, the arguments $(t, r(t))$ and $(t, z(t))$ “have been suppressed“, so that we have used the shorthand notations of the form

$$\hat{\mu}^T = \hat{\mu}^T(t,r(t)), \quad \tilde{\sigma}^T = \tilde{\sigma}^T(t,r(t)), \tag{4.43}$$

for the process $B(t,T)$, and

$$\hat{\mu}_L^T = \hat{\mu}_L^T(t,z(t)), \quad \tilde{\sigma}_L^T = \tilde{\sigma}_L^T(t,z(t)), \tag{4.44}$$

for the process $L(t,T)$, and similarly for the process $L(t,S)$. Here, when it is convenient, we will use the above notations (4.43) and (4.44). We try to choose $u_T(t)$, $u_T^L(t)$, and $u_S^L(t)$ so that the market is arbitrage-free. By Proposition 1.4.1 the portfolio rate of return and the short rate of interest must be equal, namely[7]

$$u_T(t)\hat{\mu}^T + u_T^L(t)\Big(\hat{\mu}_L^T - \lambda(t)\Big) + u_N^L(t)\Big(\hat{\mu}_L^N - \lambda(t)\Big) + u_S^L(t)\Big(\hat{\mu}_L^S - \lambda(t)\Big) = r(t), \tag{4.45}$$

[7]For the notational convenience we are using the notations (4.43) and (4.44).

necessarily holds for all t, with probability 1, and then, using (4.39), we obtain for all t

$$\begin{aligned} &u_T(t)\Big(\hat{\mu}^T(t,r(t))-r(t)\Big)+u_T^L(t)\Big(\hat{\mu}_L^T(t,z(t))-\lambda(t)-r(t)\Big)\\ &+u_N^L(t)\Big(\hat{\mu}_L^N(t,z(t))-\lambda(t)-r(t)\Big)+u_S^L(t)\Big(\hat{\mu}_L^S(t,z(t))-\lambda(t)-r(t)\Big)=0. \end{aligned} \tag{4.46}$$

Moreover we look for a portfolio minimizing the risk associated to the derivative, i.e., such that the corresponding value process has no driving Wiener process, W^z. This means that we want to solve the equation

$$u_T(t)\tilde{\sigma}^T(t,r(t))+u_T^L(t)\tilde{\sigma}_L^T(t,z(t))+u_N^L(t)\tilde{\sigma}_L^N(t,z(t))+u_S^L(t)\tilde{\sigma}_L^S(t,z(t))=0, \tag{4.47}$$

In order to see some structure, let H be the following matrix

$$H(t,z)=H(t,r,\lambda)=\begin{pmatrix} \hat{\mu}^T-r & \hat{\mu}_L^T-\lambda-r & \hat{\mu}_L^S-\lambda-r\\ \hat{\sigma}^T & \hat{\sigma}_{L,r}^T & \hat{\sigma}_{L,r}^S\\ 0 & \hat{\sigma}_{L,\lambda}^T & \hat{\sigma}_{L,\lambda}^S \end{pmatrix} \tag{4.48}$$

so that we now write (4.46) and (4.47) in matrix form as

$$H\big(t,z(t)\big)u(t)=H\big(t,r(t),\lambda(t)\big)u(t)=0, \tag{4.49}$$

where we have used the notations (4.43) and (4.44). If H were invertible, then the system (4.49) would have a unique solution, i.e., the null solution, but this solution does not satisfy the condition (4.39), then H must be singular. For readability reasons, we study $H^{'}$, the transpose of H, i.e.,

$$H'=H'(t,r,\lambda)=\begin{pmatrix} \hat{\mu}^T-r & \hat{\sigma}^T & 0\\ \hat{\mu}_L^T-\lambda-r & \hat{\sigma}_{L,r}^T & \hat{\sigma}_{L,\lambda}^T\\ \hat{\mu}_L^S-\lambda-r & \hat{\sigma}_{L,r}^S & \hat{\sigma}_{L,\lambda}^S \end{pmatrix}. \tag{4.50}$$

The matrix H' being singular, the columns are linearly dependent. Since under the conditions (4.31), i.e., $\hat{\sigma}^T(t,r(t))\neq 0$, and $\hat{\sigma}_L^{\lambda,T}(t,z(t))\neq 0$, the matrix

$$\sigma=\begin{pmatrix} \hat{\sigma}^T(t,r(t)) & 0\\ \hat{\sigma}_{L,r}^T(t,z(t)) & \hat{\sigma}_{L,\lambda}^T(t,z(t)) \end{pmatrix}.$$

is invertible (with probability 1 for each t), the first column of $H^{'}$ can be written as a linear combination of the other columns. We thus deduce the existence of the 2-dimensional process $\xi_z=(\xi_r,\xi_\lambda)'$ such that setting $\mathbf{1}_r=(1,1)'$, and $\mathbf{1}_\lambda=(0,1)'$

$$\sigma\xi_z=\mu-\lambda\mathbf{1}_\lambda-r\mathbf{1}_r,\qquad \text{i.e.,}\qquad \begin{cases} \tilde{\sigma}^T(t,r(t))\xi_z(t) &= \hat{\mu}^T(t,r(t))-r(t)\\ \tilde{\sigma}_L^T(t,z(t))\xi_z(t) &= \hat{\mu}_L^T(t,z(t))-\lambda(t)-r(t) \end{cases} \tag{4.51}$$

and taking into account (2.31)

$$\begin{cases} \tilde{\sigma}^S(t,r(t))\xi_z(t) &= \hat{\mu}^S(t,r(t))-r(t)\\ \tilde{\sigma}_L^S(t,z(t))\xi_z(t) &= \hat{\mu}_L^S(t,z(t))-\lambda(t)-r(t) \end{cases} \tag{4.52}$$

Since the longevity bond $L(t,S)$ was chosen arbitrarily, the risk premium, $\hat{\mu}_L^S(t,z(t))-\lambda(t)-r(t)$, of any longevity bond, can be written as the linear combination $\tilde{\sigma}_L^S(t,z(t))\xi_z(t)$ of is volatility components, $\xi_z(t)$ being the same for all longevity bonds. Thus equations (4.51) and (4.52) show that the process ξ_z does not depend on the choice of either S or T, and that the process ξ_z is uniquely defined by (4.51).

□

In the next theorem which is the analogue for the longevity bonds $L(t,T)$ of Theorem 2.2.3 for the bonds $B(t,T)$, we give the term structure for the longevity bonds.

Theorem 4.3.5. *Assuming that the support of the process $z(t)$ is entire set $\mathbb{R}_+^2$, $\forall t\in[0,T]$, in an arbitrage free longevity bond market, the function $\hat{L}^T(t,z)$ satisfies the term structure equation*

$$\begin{cases} \hat{L}_t^T(t,z)+\hat{L}_z^{T'}(t,z)\big(\tilde{\mu}^z-\tilde{\Sigma}^z\tilde{\xi}_z\big)(t,z)\\ +\frac{1}{2}tr\left[\left(\tilde{\Sigma}^z\right)^{'}\hat{L}_{zz}^T\tilde{\Sigma}^z\right](t,z)-(r+\lambda)\hat{L}^T(t,z)=0, \qquad (t,z)\in(0,T)\times\mathbb{R}_+^2\\ \qquad\qquad \hat{L}^T(T,z)=1, \qquad z\in\mathbb{R}_+^2 \end{cases} \tag{4.53}$$

where ξ_z is universal, in the sense that ξ_r and ξ_λ do not depend on the specific choice of the maturity T.

Now we generalize Remark 2.2.1 to this setting.

Remark 4.3.1. *Similar to the benchmark $B(t,T)$ in Remark 2.2.1, if for a maturity time T, a longevity bond price process $L(t,T)$ can be observed, then it is called a benchmark of the longevity bond market. If we assume that also $z(t)$ is observable, the obtained results can be interpreted by saying that all bond and longevity bond prices will be determined in terms of $z(t)$ and two different benchmark, $B(t,T)$ and $L(t,T)$. Indeed once the market has determined the dynamics of this benchmarks $B(t,T)$ and $L(t,T)$, then $\mu(t,T)$, $\sigma(t,T)$, $\mu_L(t,T)$, and $\sigma_L(t,T)$ can be considered as known together with $z(t)$, and there the market has implicitly specified ξ_z by equations (4.32) and (4.33). Once ξ_z is thus determined, all other bond and longevity bond prices will be determined by the term structure equation (4.53), respectively.*

Proof of Theorem 4.3.5. Taking into account the notations (4.27), (4.28), we can rewrite (4.33) in terms of $\hat{\mu}_L^T(t,z(t))$ and $\tilde{\sigma}_L^T(t,z(t))$, the latter quantities being given by (4.29) and (4.30) (of Lemma 4.3.3). After some reshuffling, we obtain the equation (4.53), as in the proof of Theorem 2.2.3. Finally, we must also have $\hat{L}^T(T,z)=1$, so we have proved the result. □

Before proceeding any further, we observe that the price dynamics of $\hat{L}(t,T)$ can be expressed by mean of the price market ξ_z. Indeed, by Itô formula and the term structure equations (4.53), we have the price dynamics of the following form

$$\begin{aligned}\frac{d\hat{L}^T(t,z(t))}{\hat{L}^T(t,z(t))} &= \left(r(t)+\lambda(t)+\frac{\hat{L}_z^T}{\hat{L}^T}(t,z(t))\tilde{\Sigma}^z(t,z(t))\hat{\xi}_z(t,z(t))\right)dt+\frac{\hat{L}_r^T}{\hat{L}^T}(t,z(t))\hat{\sigma}^r(t,r(t))dW^r(t)\\ &\quad+\frac{\hat{L}_\lambda^T}{\hat{L}^T}(t,z(t))\hat{\sigma}^\lambda(t,z(t))dW^\lambda(t), \end{aligned}\tag{4.54}$$

or in compact form

$$\begin{aligned}&\frac{d\hat{L}^T(t,z(t))}{\hat{L}^T(t,z(t))}\\ &=\left(r(t)+\lambda(t)+\frac{\hat{L}_z^T}{\hat{L}^T}(t,z(t))\tilde{\Sigma}^z(t,z(t))\hat{\xi}_z(t,z(t))\right)dt+\frac{\hat{L}_z^T}{\hat{L}^T}(t,z(t))\tilde{\Sigma}^z(t,z(t))dW^z(t),\end{aligned}\tag{4.55}$$

The proof of (4.55) follows by observing that

$$d\hat{L}^T(t,z(t))=\hat{L}_t^T(t,z(t))dt+\hat{L}_z^T(t,z(t))dz(t)+\frac{1}{2}tr\left[\left(\tilde{\Sigma}^z\right)'\hat{L}_{zz}^T\tilde{\Sigma}^z\right](t,z(t)),\tag{4.56}$$

and inserting the differential form (4.6), (4.7), (4.8) into (4.56), we obtain

$$\begin{aligned}d\hat{L}^T(t,z(t))&=\hat{L}_t^T(t,z(t))dt+\hat{L}_z^T(t,z(t))dz(t)+\frac{1}{2}tr\left[\left(\tilde{\Sigma}^z\right)'\hat{L}_{zz}^T\tilde{\Sigma}^z\right](t,z(t))\\ &=\left(\hat{L}_t^T(t,z(t))+\hat{L}_z^T(t,z(t))\tilde{\mu}^z(t,z(t))+\frac{1}{2}tr\left[\left(\tilde{\Sigma}^z\right)'\hat{L}_{zz}^T\tilde{\Sigma}^z\right](t,z(t))\right)dt\\ &\quad+\hat{L}_z^T(t,z(t))\tilde{\Sigma}^z(t,z(t))dW^z(t)\\ &=\left(\big(r(t)+\lambda(t)\big)\hat{L}^T+\hat{L}_z^T(t,z(t))\tilde{\Sigma}^z(t,z(t))\hat{\xi}_z(t,z(t))\right)dt+\hat{L}_z^T(t,z(t))\tilde{\Sigma}^z(t,z(t))dW^z(t),\end{aligned}$$

where in the last step we have used the following relation

$$\hat{L}_t^T+\hat{L}_z^T\tilde{\mu}^z+\frac{1}{2}tr\left[\tilde{\Sigma}^{z\prime}\hat{L}_{zz}^T\tilde{\Sigma}^z\right]=\hat{L}_z^T\tilde{\Sigma}^z\hat{\xi}_z+(r(t)+\lambda(t))\hat{L}^T,$$

given by the term structure equation (4.53) when all terms are evaluated at the point $(t,z(t))$.

As in Section 2.2, in the present setting ξ_z is not determined within the model a less to benchmark bonds are specified as shown in Proposition 4.3.4. Alternatively, in order to solve (4.53), we have to specify ξ_z exogenously just as we have to specify μ^z and Σ^z.

Again an application of the Feynman-Kač technique, (see Proposition 2.2.4), gives us a stochastic representation formula. Now we can repeat the same steps as in Section 2.1. Summarizing, we assume that $\xi_z\in\mathcal{L}^2(0,T;\bar{\mathbb{F}}^{W^z})$ and that the measure Q, defined in Lemma 1.4.7, is a probability measure. We observe that if ξ_z satisfies the Novikov condition (1.34), then Q is a probability measure. By Lemma 1.4.7, choosing $\xi(t):=\xi_z(t)$ and assuming (1.39), we have that

- the equivalent martingale measure Q is given by

$$dQ = \exp\left(-\int_0^T \xi_z(t)dW^z(t) - \frac{1}{2}\int_0^T \xi_z'(t)\xi_z(t)dt\right)dP, \tag{4.57}$$

- the process $\bar{W}^z(t)$ defined as

$$\bar{W}^z(t) := \int_0^t \xi_z(s)ds + W^z(t), \tag{4.58}$$

  is a Wiener process with respect to Q.

Under Q, the process z solves the following equation

$$dz(u) = \tilde{\mu}^z(u, z(u)) + \tilde{\Sigma}^z(u, z(u))\left[d\bar{W}^z(u) - \tilde{\xi}_z(u, z(u))du\right]. \tag{4.59}$$

Finally we obtain the following stochastic representation formula.

Proposition 4.3.6. *In an arbitrage free longevity bond market, let us assume that $\xi_z \in \mathcal{L}^2(0, T; \bar{\mathbb{F}}^{W^z})$ and that (4.57) defines a probability measure Q. Then the function $\hat{L}^T$ is given by the formula*

$$\hat{L}(t, z; T) = E^Q_{t,z}\left(e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(s)ds}\right). \tag{4.60}$$

As usual, the subscripts t and z denote that the expectation is taken using the dynamics given by (4.59), with the initial condition $z(t) = z$,i.e.,

$$\begin{cases} dz(s) = & \left[\tilde{\mu}^z(s, z(s)) - \tilde{\Sigma}^z(s, z(s))\tilde{\xi}_z(s, z(s))\right]ds + \tilde{\Sigma}^z(s, z(s))d\bar{W}^z(s), \quad s \in [t, T] \\ z(t) = & z \end{cases} \tag{4.61}$$

where $\bar{W}^z$ is the Wiener process with respect to Q defined in (4.58).

From (4.60), we observe that the longevity bond prices processes are given by[8]

$$L(t, T) = E^Q_{t,z}\left(e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(s)ds}\right)|_{z=z(t)}. \tag{4.63}$$

Proof. We fix t and $z = (r, \lambda)$, set $P(s) = e^{-\int_t^s (r(u)+\lambda(u))du}$, so that

$$dP(s) = -(r(s) + \lambda(s))P(s)ds, \quad s \in [t, T]. \tag{4.64}$$

By Itô's multidimensional formula, we have

$$\begin{aligned} d\hat{L}^T(s, z(s)) =& \hat{L}^T_s ds + \hat{L}^T_z dz + \frac{1}{2}tr\left[(\tilde{\Sigma}^z)'\hat{L}^T_{zz}\tilde{\Sigma}^z\right] \\ =& \left[\hat{L}^T_s + \hat{L}^T_z(\tilde{\mu}^z - \tilde{\Sigma}^z\tilde{\xi}_z) + \frac{1}{2}tr\left[(\tilde{\Sigma}^z)'\hat{L}^T_{zz}\tilde{\Sigma}^z\right]\right]ds + \hat{L}^T_z\tilde{\Sigma}^z d\bar{W}^z(s), \end{aligned} \tag{4.65}$$

where we have used the same shorthand notations (4.43) and (4.44), but considering s instead of t. Now, proceeding exactly as in to proof of Proposition 2.2.4 and fixing (t, z), we define the process $\hat{P}$ as

$$\hat{P}(s) = P(s)\hat{L}^T(s, z(s)), \quad s \in [0, T]. \tag{4.66}$$

Then, by (4.64) and (4.65), we obtain

$$\begin{aligned} & d\hat{P}(s) - \left(P(s)d\hat{L}^T(s, z(s)) + dP(s)\hat{L}^T(s, z(s))\right) \\ & = e^{-\int_t^s (r(u)+\lambda(u))du}d\hat{L}^T(s, z(s)) - (r(s) + \lambda(s))e^{-\int_t^s (r(u)+\lambda(u))du}\hat{L}^T(s, z(s))ds \\ & = e^{-\int_t^s (r(u)+\lambda(u))du}\left\{\left[\hat{L}^T_s - (r(s) + \lambda(s))\hat{L}^T + \hat{L}^T_z(\tilde{\mu}^z - \tilde{\Sigma}^z\tilde{\xi}_z) + \frac{1}{2}tr\left((\tilde{\Sigma}^z)'\hat{L}^T_{zz}\tilde{\Sigma}^z\right)\right]ds \right. \\ & \left. + \hat{L}^T_z\tilde{\Sigma}^z d\bar{W}^z(s)\right\}, \end{aligned}$$

[8] Recall that by Markov property of $z(t)$ with respect to the filtration $\mathcal{H}_t$, where $\mathcal{H}_t = \mathcal{F}^N_t \vee \mathcal{F}^r_t \vee \mathcal{F}^\lambda_t$, we have

$$E^Q\left(e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(s)ds}|\mathcal{H}_t\right) = E^Q_{t,z}\left(e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(s)ds}\right)\Big|_{z=z(t)}. \tag{4.62}$$

.

or equivalently

$$
\begin{aligned}
&\hat{P}(T)-\hat{P}(t)=P(T)\hat{L}^{T}(T,z)-P(t)\hat{L}^{T}(t,z(t))\\
&=\int_t^T e^{-\int_t^s(r(u)+\lambda(u))du}\left[\hat{L}^T_s-(r(s)+\lambda(s))\hat{L}^T+\hat{L}^T_z(\tilde{\mu}^z-\tilde{\Sigma}^z\tilde{\xi}_z)+\frac{1}{2}tr\left((\tilde{\Sigma}^z)'\hat{L}^T_{zz}\tilde{\Sigma}^z\right)\right]ds\\
&+\int_t^T e^{-\int_t^s(r(u)+\lambda(u))du}\hat{L}^T_z\tilde{\Sigma}^z d\bar{W}^z(s).
\end{aligned}
$$

In the above expression the time integral vanishes, since $\hat{L}^T(s,z(s))$ satisfies equation (4.53) evaluated at the point $(s,z(s))$ (see Theorem 4.3.5). Then, taking into account that $\hat{L}^T(T,z) = 1$ and $P(t)=1$, we obtain

$$
e^{-\int_t^T(r(s)+\lambda(s))ds}=\hat{L}^T(t,z)+\int_t^T e^{-\int_t^s(r(u)+\lambda(u))du}\hat{L}^T_z(s,z(s))\tilde{\Sigma}^z(s,z(s))d\bar{W}^z(s). \tag{4.67}
$$

Taking the expectation of (4.67), we have

$$
E^Q_{t,z}\left(e^{-\int_t^T(r(s)+\lambda(s))ds}\right)=\hat{L}(t,z;T),
$$

the expected value of the stochastic integral being equals to zero.

We have proved the announced result.

□

Rewriting formula (4.63) as

$$
L(t,T)=E^Q_{t,z}\left[e^{-\int_t^T(r(s)+\lambda(s))ds}\cdot 1\right]\big|_{z=z(t)}\ , \tag{4.68}
$$

we observe that the value of a T-longevity bond at time t is given as the expected value of one dollar (final payoff), discount to present value at the interest rate given by r, with a cost due to the longevity risk depending on λ. Thus formula (4.68) can be interpreted as the risk-neutral pricing formula for a T-bond at the interest rate given by r with the cost rate given by λ, but in our model we may have different martingale measures for different choices of ξ_z.

4.3.1 A bidimensional CIR model

In this section we take a model where the interest rate $r(t)$ and the stochastic mortality intensity $\lambda(t)$ are dependent, but with uncorrelated driving noises, and we extend Proposition 2.2.5 to this setting, i.e., we want to derive an explicit formula for the price of a zero coupon longevity bond as a function of the interest rate and the stochastic intensity. In particular

We take a CIR model for the interest rate $r(t)$, i.e.,

$$
dr(t)=a_r\left(b_r-r(t)\right)dt+\bar{\sigma}_r\sqrt{r(t)}d\bar{W}^r(t), \tag{4.69}
$$

$$
r(t_0)=r \tag{4.70}
$$

where a_r, b_r, $\bar{\sigma}_r$ and r are strictly positive deterministic constants such that $2\,a_r\,b_r>\bar{\sigma}_r^2$, and $\bar{W}^r(t)$ is a Wiener process under a martingale measure Q (see (2.50) of Section 2.2), while the stochastic intensity is given by (3.51), i.e.,

$$
\begin{cases}
d\lambda^{(c)}(t)=a_\lambda\left(b_\lambda-\lambda^{(c)}(t)+c\,r(t)\right)dt+\bar{\sigma}_\lambda\sqrt{\lambda^{(c)}(t)}d\bar{W}^\lambda(t),\\
\lambda^{(c)}(t_0)=\lambda^{(c)}
\end{cases} \tag{4.71}
$$

where $\bar{W}^\lambda$ is a 1-dimensional Wiener process under a martingale measure Q, independent of $\bar{W}^r$, and a_λ, b_λ, $\bar{\sigma}_\lambda$, c and $\lambda^{(c)}$ are strictly positive deterministic constants such that $2\,a_\lambda\,b_\lambda>\bar{\sigma}_\lambda^2$. We call a term structure of interest rate and mortality intensity model involving (2.50) and (3.51) a bidimensional CIR model.

Recalling that in this setting $z(t)=(r(t),\lambda(t))'$, with $\lambda(t)=\lambda^{(c)}(t)$, a zero coupon longevity bond is a contract that pays off 1 at time T, with price process $L(t,T)=\hat{L}^T(t,z(t))$.

Following the similar approach of Section 2.2.1, we want extend the property (2.51) and the formula (2.43) to this setting as shown in the following proposition.

Proposition 4.3.7. *The term structure for the bidimensional CIR model is given by*

$$\hat{L}^T(t,z) = e^{\psi_z^{(c),0}(T-t)+\psi_r^{(c)}(T-t)r+\psi_\lambda(T-t)\lambda}, \tag{4.72}$$

where

$$\dot{\psi}_\lambda(s) = -a_\lambda \psi_\lambda(s) + \frac{\bar{\sigma}_\lambda^2}{2}\psi_\lambda^2(s) - 1 \tag{4.73}$$

$$\dot{\psi}_r^{(c)}(s) = -a_r \psi_r^{(c)}(s) + \frac{\bar{\sigma}_r^2}{2}\big(\psi_r^{(c)}\big)^2(s) + a_\lambda\, c\, \psi_\lambda(s) - 1 \tag{4.74}$$

$$\dot{\psi}_z^{(c),0}(s) = a_r\, b_r\, \psi_r^{(c)}(s) + a_\lambda\, b_\lambda \psi_\lambda(s) \tag{4.75}$$

with the initial conditions

$$\psi_z^{(c),0}(0) = 0, \quad \psi_r^{(c)}(0) = 0, \quad \psi_\lambda(0) = 0. \tag{4.76}$$

Furthermore ψ_λ and $\psi_r^{(c)}$ are bounded functions, i.e.,

$$-\frac{1}{|\beta_\lambda|} \le \psi_\lambda(s) \le 0 \quad \textit{and} \quad -\frac{1}{|\beta^-|} \le \psi_r^{(c)}(s) \le 0, \tag{4.77}$$

where, setting $h = 1 + \frac{a_\lambda c}{|\beta_\lambda|}$,

$$\beta_\lambda = \frac{\alpha_\lambda - a_\lambda}{2}, \quad \beta^- = \frac{\alpha^- - \frac{a_r}{h}}{2},$$

with

$$\alpha_\lambda = -\sqrt{a_\lambda^2 + 2\bar{\sigma}_\lambda^2}, \quad \alpha^- = -\sqrt{\frac{a_r^2}{h^2} + \frac{2\bar{\sigma}_r^2}{h}}.$$

Remark 4.3.2. *First we observe that by above bounds for ψ_λ and $\psi_r^{(c)}$ and differential equation (4.75), obtain immediately that*

$$-\left(\frac{a_r\, b_r}{|\beta^-|} + \frac{a_\lambda\, b_\lambda}{|\beta_\lambda|}\right) s \le \psi_z^{(c),0}(s) \le 0. \tag{4.78}$$

Then in order to solve the above differential system for ψ_λ, $\psi_r^{(c)}$ and $\psi_z^{(c),0}$, we observe the following. In (4.72) we should have written $\psi_\lambda^{(c)}$ instead of ψ_λ, but as equation (4.73) does not depend on c, the solution $\psi_\lambda(s)$ is the same for both processes $\lambda^{(0)}$ and $\lambda^{(c)}$, i.e., the solution of the equation (4.73) is exactly the expression (3.57), i.e.

$$\psi_\lambda(s) = \frac{1 - e^{\alpha_\lambda s}}{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}},$$

where (see (3.59))

$$\alpha_\lambda = -\sqrt{a_\lambda^2 + 2\bar{\sigma}_\lambda^2}, \quad \beta_\lambda = \frac{\alpha_\lambda - a_\lambda}{2}, \quad \gamma_\lambda = \frac{\alpha_\lambda + a_\lambda}{2}.$$

By results In order to compute $\psi_z^{(c),0}$ we observe that, substituting $\psi_\lambda(s)$ in (4.75) and recalling that (see (3.58))

$$\int_0^s a_\lambda b_\lambda \psi_\lambda(u)du = \psi_\lambda^0(s) = -\frac{2\, a_\lambda b_\lambda}{\bar{\sigma}_\lambda^2} \ln\left(\frac{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}}{\alpha_\lambda}\right) + \frac{a_\lambda b_\lambda}{\beta_\lambda}\, s, \tag{4.79}$$

we obtain

$$\psi_z^{(c),0}(s) = -\frac{2\, a_\lambda b_\lambda}{\bar{\sigma}_\lambda^2} \ln\left(\frac{\beta_\lambda + \gamma_\lambda\, e^{\alpha_\lambda s}}{\alpha_\lambda}\right) + \frac{a_\lambda b_\lambda}{\beta_\lambda}\, s + a_r b_r \int_0^s \psi_r^{(c)}(u)du, \tag{4.80}$$

where the last term is determined using numerical procedures, such as, for example, the standard Euler methods, since it does not seem possible to determine analytically the function $\psi_r^{(c)}$. Nevertheless, as shown in the proof *of Proposition 4.3.7, we can represent* $\psi_r^{(c)}$ *as* $\psi_r^{(c)}(s) = -\frac{2}{\bar{\sigma}_r^2}\frac{x_2(s)}{x_1(s)}$ *(see (4.85)) where,* x_1 *and* x_2 *satisfy a differential system (see (4.84)). Furthermore, as shown in Appendix C,* x_1 *and* x_2 *can be represented as*

$$
\begin{aligned}
x_1(t) &= 1 + \int_0^t x_2(s)ds \\
x_2(t) &= \int_0^t A(s)\,ds + \int_0^t A(s)\Big(\int_s^t C(u_1,t)\,du_1\Big)\,ds \\
&+ \sum_{k=1}^{\infty}\int_0^t A(s)\Big(\int_s^t du_m\Big(\int_{u_m}^t du_{m-1}\ldots\int_{u_3}^t du_2\int_{u_2}^t du_1 C(u_1,t)\Big)\,C(u_2,u_1)\ldots C(u_m,u_{m-1})\Big)\Big)\,ds,
\end{aligned}
$$

with

$$C(u,t) = \Im_A(t) - \Im_A(u) - B, \qquad 0 \le u \le t,$$

where $B = a_r$ *and*

$$\Im_A(t) = \frac{\bar{\sigma}_r^2}{2}\big(a_\lambda\, c\int_0^t \psi_\lambda(s)\,ds + t\big) = \frac{\bar{\sigma}_r^2}{2}\frac{c}{b_\lambda}\,\psi_\lambda^0(t) + \frac{\bar{\sigma}_r^2}{2}\,t.$$

Proof of Proposition 4.3.7. By the term structure equation (4.53), we have that $\hat{L}^T$ is the solution to the following partial differential equation

$$
\begin{cases}
\hat{L}_t^T(t,z) + a_r\,(b_r - r)\,\hat{L}_r^T(t,z) + a_\lambda\,(b_\lambda - \lambda + c\,r)\,\hat{L}_\lambda^T(t,z) \\
\quad + \frac{\bar{\sigma}_r^2}{2}\,r\,\hat{L}_{rr}^T(t,z) + \frac{\bar{\sigma}_\lambda^2}{2}\,\lambda\,\hat{L}_{\lambda\lambda}^T(t,z) = (r+\lambda)\hat{L}^T(t,z), \\
\qquad \hat{L}^T(T,z) = 1.
\end{cases}
\tag{4.81}
$$

Now we consider

$$\hat{L}^T(t,z) = e^{\psi_z^{(c),0}(T-t) + \psi_r^{(c)}(T-t)r + \psi_\lambda(T-t)\lambda}, \tag{4.82}$$

as a guess function. By (4.72) we have

$$
\begin{aligned}
&\hat{L}_t^T(t,r,\lambda) = -\hat{L}^T\left(\dot{\psi}_z^{(c),0}(T-t) + \dot{\psi}_r^{(c)}(T-t)r + \dot{\psi}_\lambda(T-t)\lambda\right), \\
&\hat{L}_r^T(t,r,\lambda) = \hat{L}^T(t,r,\lambda)\psi_r^{(c)}(T-t), \quad \hat{L}_{rr}^T(t,r,\lambda) = \hat{L}^T(t,r,\lambda)\big(\psi_r^{(c)}\big)^2(T-t), \\
&\hat{L}_\lambda^T(t,r,\lambda) = \hat{L}^T(t,r,\lambda)\psi_\lambda(T-t), \quad \hat{L}_{\lambda\lambda}^T(t,r,\lambda) = \hat{L}^T(t,r,\lambda)\big(\psi_\lambda\big)^2(T-t),
\end{aligned}
$$

where $\dot{\psi}_z^{(c),0}(u) = \frac{d\psi_z^{(c),0}}{du}(u)$, $\dot{\psi}_r^{(c)}(u) = \frac{d\psi_r^{(c)}}{du}(u)$, and $\dot{\psi}_\lambda(u) = \frac{d\psi_\lambda}{du}(u)$. Substituting into the partial differential equation (4.81) and dividing each term by the common factor $\hat{L}^T$, we have

$$
\begin{aligned}
&-\left(\dot{\psi}_z^{(c),0}(T-t) + \dot{\psi}_r^{(c)}(T-t)r + \dot{\psi}_\lambda(T-t)\lambda\right) + a_r(b_r - r)\psi_r^{(c)}(T-t) + a_\lambda\,(b_\lambda - \lambda + c\,r)\,\psi_\lambda(T-t) \\
&+ \frac{\bar{\sigma}_r^2}{2}\,r\,\big(\psi_r^{(c)}\big)^2(T-t) + \frac{\bar{\sigma}_\lambda^2}{2}\,\lambda\psi_\lambda^2(T-t) = r + \lambda,
\end{aligned}
$$

so that grouping the terms multiplying r and λ we obtain

$$
\begin{aligned}
&\left(-\dot{\psi}_r^{(c)}(T-t) - a_r\psi_r^{(c)}(T-t) + a_\lambda\,c\,\psi_\lambda(T-t) + \frac{\bar{\sigma}_r^2}{2}\big(\psi_r^{(c)}\big)^2(T-t) - 1\right) r \\
&+ \left(-\dot{\psi}_\lambda(T-t) - a_\lambda\psi_\lambda(T-t) + \frac{\bar{\sigma}_\lambda^2}{2}\psi_\lambda^2(T-t) - 1\right)\lambda \\
&- \dot{\psi}_z^{(c),0}(T-t) + a_r\,b_r\,\psi_r^{(c)}(T-t) + a_\lambda\,b_\lambda\psi_\lambda(T-t) = 0.
\end{aligned}
$$

Since the above equality holds for all $r \ge 0$ and $\lambda \ge 0$, the terms multiplying r and λ are zero. Then setting $T - t = s$ we obtain (4.73), (4.74) and (4.75), with the initial conditions (4.76) from the terminal condition $\hat{L}^T(T,z) = 1$. It follows that if (4.73), (4.74) and (4.75) are solved subject to the boundary conditions (4.76), the function (4.72) provides the price of a zero coupon longevity bond maturing at time T.

To prove that ψ_λ is a bounded continuous function, we recall that, by its explicit expression (3.57) we have that

$$\dot{\psi}_\lambda(s) = \frac{-\alpha_\lambda(\beta_\lambda + \gamma_\lambda)e^{\alpha_\lambda s}}{\big(\beta_\lambda + \gamma_\lambda\,e^{\alpha_\lambda s}\big)^2} = \frac{-\alpha_\lambda^2 e^{\alpha_\lambda s}}{\big(\beta_\lambda + \gamma_\lambda\,e^{\alpha_\lambda s}\big)^2} \le 0,$$

and that (3.60) holds, i.e.,

$$\lim_{s\to\infty}\psi_\lambda(s)=\lim_{s\to\infty}\frac{1-e^{-|\alpha_\lambda|s}}{\beta_\lambda+\gamma_\lambda\,e^{-|\alpha_\lambda|s}}=-\frac{1}{|\beta_\lambda|}.$$

Then, since $\psi_\lambda(0)=0$, the bounds for ψ_λ in (4.77) immediately follow.

Now we turn to equation (4.74). Setting[9]

$$y_r(s)=e^{-\int_0^s\frac{\bar\sigma_r^2}{2}\,\psi_r^{(c)}(u)\,du}, \tag{4.83}$$

and using equation (4.74), we obtain an inhomogeneous second order differential equation for $y_r(s)$

$$\begin{cases}\ddot y_r(s)+a_r\dot y_r(s)-\frac{\bar\sigma_r^2}{2}\big(a_\lambda\,c\,\psi_\lambda(s)+1\big)y_r=0,\\ \dot y_r(0)=0\\ y_r(0)=1.\end{cases}$$

Setting

$$x_1(s)=y_r(s)\quad\text{and}\quad x_2(s)=\dot y_r(s),$$

the above equation becomes

$$\begin{cases}\dot x_1(s)=x_2(s)\\ \dot x_2(s)=A(s)x_1(s)-Bx_2(s)\\ x_1(0)=1;\quad x_2(0)=0\end{cases} \tag{4.84}$$

where

$$A(s)=\frac{\bar\sigma_r^2}{2}\big(a_\lambda\,c\,\psi_\lambda(s)+1\big),\qquad B=a_r,$$

so that

$$\psi_r^{(c)}(s)=-\frac{2}{\bar\sigma_r^2}\frac{x_2(s)}{x_1(s)}. \tag{4.85}$$

The continuity and boundedness of $A(s)$ guarantees existence and uniqueness of the solution of the differential system (4.84) defining $x_1(t)$ and $x_2(t)$.

The boundedness property for A follows by noting that, by (4.77), i.e. $-\frac{1}{|\beta_\lambda|}\le\psi_\lambda(s)\le 0$, immediately implies

$$\frac{\bar\sigma_r^2}{2}\big(1-c\,a_\lambda\frac{1}{|\beta_\lambda|}\big)\le A(s)=\frac{\bar\sigma_r^2}{2}\big(a_\lambda\,c\,\psi_\lambda(s)+1\big)\le\frac{\bar\sigma_r^2}{2},\qquad B=a_r.$$

(Observe that for c sufficiently small, the function $0<A(s)\le\frac{\bar\sigma_r^2}{2}$, for all $s\ge 0$. Observe that even if c is not small, there exists a constant $\overline{A}$ such that $|A(s)|\le\overline{A}$, for all $s\ge 0$.)

We now turn to the proof of the upper and lower bounds of the function $\psi_r^{(c)}$. To this end it is sufficient to show that

$$H_-(\psi_r^{(c)})\le\dot\psi_r^{(c)}(s)\le H_+(\psi_r^{(c)}), \tag{4.86}$$

where

$$H_+(y):=-a_ry+\frac{\bar\sigma_r^2}{2}y^2(s)-1$$

and

$$H_-(y):=H_+(y)-\frac{a_\lambda\,c}{|\beta_\lambda|}=-a_ry+\frac{\bar\sigma_r^2}{2}y^2(s)-h,\quad\text{with } h=1+\frac{a_\lambda\,c}{|\beta_\lambda|}.$$

[9] The transformation is similar to the transformation used to solve a Riccati equation. (See Note 6 in Section 2.2.1).

Indeed, if the bounds for ψ_λ in (4.86) hold, by using Gronwall's inequality we obtain

$$\psi^-_{r,(c)} \le \psi^{(c)}_r(s) \le \psi^+_{r,(c)},$$

where the functions $\psi^\pm_{r,(c)}$ are the solutions

$$\dot\psi^+_{r,(c)} = H_+(\psi^+_{r,(c)}), \qquad \dot\psi^-_{r,(c)} = H_-(\psi^-_{r,(c)}),$$

with $\psi^\pm_{r,(c)}(0) = 0$, i.e.,

$$\psi^+_{r,(c)}(s) = \frac{1 - e^{\alpha^+ s}}{\beta^+ + \gamma^+ e^{\alpha^+ s}}, \qquad \psi^-_{r,(c)}(s) = \frac{1 - e^{\alpha^- s}}{\beta^- + \gamma^- e^{\alpha^- s}} \tag{4.87}$$

where[10]

$$\alpha^+ = -\sqrt{a_r^2 + 2\bar\sigma_r^2}, \quad \beta^+ = \frac{\alpha^+ - a_r}{2}, \quad \gamma^+ = \frac{\alpha^+ + a_r}{2},$$
$$\alpha^- = -\sqrt{\frac{a_r^2}{h^2} + \frac{2\bar\sigma_r^2}{h}}, \quad \beta^- = \frac{\alpha^- - \frac{a_r}{h}}{2}, \quad \gamma^- = \frac{\alpha^- + \frac{a_r}{h}}{2}.$$

Then we immediately get (4.77) for $\psi^{(c)}_r$, since

$$\frac{1 - e^{\alpha^- s}}{\beta^- + \gamma^- e^{\alpha^- s}} \ge -\frac{1}{|\beta^-|} \qquad \text{and} \qquad \frac{1 - e^{\alpha^+ s}}{\beta^+ + \gamma^+ e^{\alpha^+ s}} \le 0. \tag{4.88}$$

It remains to prove (4.86). Using the expression of H_+, equation (4.74) becomes

$$\dot\psi^{(c)}_r(s) = H_+(\psi^{(c)}_r) + a_\lambda\, c\, \psi_\lambda(s).$$

Then, taking into account (4.77), we immediately get (4.86), since

$$H_-(\psi^{(c)}_r) - \frac{a_\lambda\, c}{|\beta_\lambda|} \le H_+(\psi^{(c)}_r) + a_\lambda\, c\, \psi_\lambda(s) \le H_+(\psi^{(c)}_r).$$

□

4.4 Discrete-time Rolling Longevity Bonds

Following the same approach of Section 2.4, in this section the aim is focused on the problem of modelling a discrete-time rolling longevity bond price process.

Exactly as in Section 2.4, we fix a discrete set of times $\mathcal{T} = \{t_k\}_{k\ge0}$ such that $t_k \le t_{k+1}$ and consider a self-financing strategy such that, its total wealth is reinvested at any fixed date $t \in \mathcal{T}$ in discount longevity bonds maturing at time $t + T$ (i.e., no cash component is present). For a fixed T, the price process of this strategy is referred to as the discrete-time rolling longevity bond. As in the case of the rolling bond, we fix $\Delta \in (0, T)$ and take $t_k = k\Delta$, for $k = 0, 1, 2, \ldots$, and we denote $O^\Delta(t, T)$ the corresponding price process. Assume that at time $t \in [t_0, t_1) = [0, \Delta)$ we hold 1 longevity bond, so that $O^\Delta(0, T) = L(0, T)$ and

$$O^\Delta(t, T) = L(t, T) = \hat L^T(t, z(t)) \qquad 0 \le t < \Delta.$$

[10] As it is well known, the general solution of the Riccati equation

$$\dot w(t) = bw^2(t)/2 - aw(t) - 1, \qquad w(0) = 0$$

is given by

$$w(t; a, b) = \frac{1 - e^{\alpha s}}{\beta + \gamma e^{\alpha s}}, \quad \alpha = -\sqrt{a^2 + 2b},\ \beta = \frac{\alpha - a}{2},\ \gamma = \frac{\alpha + a}{2}.$$

Then the general solution of

$$\dot v(t) = bv^2(t)/2 - av(t) - h, \qquad v(0) = 0$$

is $v(t; a, b) = w(ht; a/h, b/h)$, as immediately follows by observing that, setting $\tilde w(s) = w(hs; a', b')$,

$$\dot{\tilde w}(s) = h\dot w(hs) = +ha'w^2(hs) + hb'w(hs) - h.$$

At time $t_1 = \Delta$, the wealth $L(\Delta, T)$ is reinvested in longevity bonds maturing at time $T + \Delta$ and we keep it until time $t_2 = 2\Delta$, so that

$$O^{\Delta}(t,T) = \frac{L(\Delta,T)}{L(\Delta,T+\Delta)}\, L(t,T+\Delta), \qquad \Delta \le t < 2\Delta.$$

and so on for other periods (see Section 2.4). Finally we have that for $t \ge 0$, the price process of the discrete-time rolling longevity bond satisfies

$$O^{\Delta}(t,T) = \prod_{k=1}^{\lfloor t/\Delta \rfloor} \frac{L(k\Delta, T+(k-1)\Delta)}{L(k\Delta, T+k\Delta)}\, L(t, T+\lfloor t/\Delta \rfloor \Delta) = O^{\Delta}(\lfloor t/\Delta \rfloor \Delta)\, \hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}(t, z(t)). \tag{4.89}$$

The last formula leads to the following result.

Proposition 4.4.1. *Let $L(t,T)$ be a zero coupon longevity bond with price processes given by (4.55). For any fixed T, the price process $O^{\Delta}(\cdot, T)$ of the discrete-time rolling longevity bond satisfies*

$$\frac{dO^{\Delta}(t,T)}{O^{\Delta}(t,T)} = \mu_{O^{\Delta}}(t,T)dt + \sigma_{O^{\Delta}}(t,T)dW^{z}(t), \tag{4.90}$$

where

$$\mu_{O^{\Delta}}(t,T) = \hat{\mu}^{T}_{O^{\Delta}}(t,z(t)) = r(t) + \lambda(t) + \frac{\hat{L}_z^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}}(t,z(t))\tilde{\Sigma}^{z}(t,z(t))\hat{\xi}_z(t,z(t)), \tag{4.91}$$

$$\sigma_{O^{\Delta}}(t,T) = \tilde{\sigma}^{T}_{O^{\Delta}}(t,z(t)) = \frac{\hat{L}_z^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}}(t,z(t))\tilde{\Sigma}^{z}(t,z(t)) = \left(\hat{\sigma}^{r} \frac{\hat{L}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}},\ \hat{\sigma}^{\lambda} \frac{\hat{L}_\lambda^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}} \right). \tag{4.92}$$

Proof. Similar to the proof of Proposition 2.4.1, after some reshuffling, we obtain the announced result.

□

In particular, let us consider the CIR bidimensional model introduced in Section 4.3.1. By the explicit formula for longevity bonds (4.72) and the expression (4.89) we obtain the following explicit formula for $t > 0$

$$\begin{aligned} O^{\Delta}(t,T) = {} & \prod_{k=1}^{\lfloor t/\Delta \rfloor} \frac{e^{\psi_z^{(c),0}(T-\Delta) + r(k\Delta)\psi_r^{(c)}(T-\Delta) + \lambda(k\Delta)\psi_\lambda(T-\Delta)}}{e^{\psi_z^{(c),0}(T) + r(k\Delta)\psi_r^{(c)}(T) + \lambda(k\Delta)\psi_\lambda(T)}} \\ & \cdot e^{\psi_z^{(c),0}(T+\lfloor t/\Delta \rfloor \Delta - t) + r(t)\psi_r^{(c)}(T+\lfloor t/\Delta \rfloor \Delta - t) + \lambda(t)\psi_\lambda(T+\lfloor t/\Delta \rfloor \Delta - t)}. \end{aligned} \tag{4.93}$$

Furthermore, in this framework, $\mu_{O^{\Delta}}(t,T)$ and $\sigma_{O^{\Delta}}(t,T)$ given by (4.91) and (4.92) become

$$\begin{aligned} \mu_{O^{\Delta}}(t,T) = \hat{\mu}^{T}_{O^{\Delta}}(t,z(t)) = {} & r(t) + \lambda(t) + \hat{\xi}_r(t,r(t))\hat{\sigma}^{r}(t,r(t))\psi_r^{(c)}(T+\lfloor t/\Delta \rfloor \Delta - t) \\ & + \hat{\xi}_\lambda(t,z(t))\hat{\sigma}^{\lambda}(t,z(t))\psi_\lambda(T+\lfloor t/\Delta \rfloor \Delta - t), \end{aligned} \tag{4.94}$$

$$\sigma_{O^{\Delta}}(t,T) = \tilde{\sigma}^{T}_{O^{\Delta}}(t,z(t)) = \left(\hat{\sigma}^{r}(t,r(t))\psi_r^{(c)}(T+\lfloor t/\Delta \rfloor \Delta - t),\ \hat{\sigma}^{\lambda}(t,z(t))\psi_\lambda(T+\lfloor t/\Delta \rfloor \Delta - t) \right). \tag{4.95}$$

4.5 A new Zero Coupon Longevity Bonds: the term structure equation

In this section we present another type of mortality-linked bonds, i.e., we take into account a zero coupon longevity bond, defined as financial security paying to holder one unit of cash at a fixed date T, if he/she is alive at time T (and zero otherwise). In this setting the payment at the time of maturity, known as the principal value or face value, equals one if holder is alive at time T, else zero, while in the BNP-Paribas longevity bonds the principal value equals always 1.

As in Section 4.3, we denote by τ^{j}, with $j = 1, \ldots, m$, the death time of the jth element of the given population,, m being the size of the population, and by τ the death time of the investor.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space, and we take the following assumptions.

Assumption 4.5.1. *We assume that there exists a strictly positive process λ and a σ-algebra $\mathcal{G}$, with $\mathcal{G} \supset \mathcal{F}^{\lambda}_{\infty} \vee \mathcal{F}^{r}_{\infty}$, such that*

- *τ and τ^j, $j = 1, \ldots, m$, are, conditionally on $\mathcal{G}$, independent and identically distributed random variables on $(\Omega, \mathcal{F}, P)$, so that*
$$\mathcal{L}^P\left(\tau^j \,|\mathcal{G}\right) = \mathcal{L}^P\left(\tau \,|\mathcal{G}\right), \quad j = 1, \ldots, m; \tag{4.96}$$
- *τ is, accordingly to Section 3.2 and 3.4, the first jump time of a doubly stochastic Poisson process $N(t)$ with respect to $\mathcal{G}$ with the intensity function $\lambda(t)$. Analogously for τ^j and $N^j(t)$, all with the same stochastic intensity $\lambda(t)$.*

The aim is focused on the problem of finding an arbitrage-free price process of these new T-zero coupon longevity bonds, and to this aim we will use the approach of Section 4.3. Formally, we have the following definition.

Definition 4.5.1 (New Zero Coupon Longevity Bond)**.** *A zero coupon longevity bond with maturity date T, also called a new T-longevity bond, is a zero coupon bond, which guarantees the holder* 1 *dollar to be paid on the date T if he/she is alive at time T. The (random) price at time t of such longevity bond, with maturity T, is denoted by $L^\tau(t,T)$.*

Let us first extend briefly the Assumptions 4.3.1 and 4.3.2 to this setting.

Assumption 4.5.2. *In addition to the Assumptions 2.2.1 and 2.2.2 on the bond market, and the Assumptions 4.5.1 and 4.2.2, we assume that there exists a market for zero coupon T-longevity bonds for every value of T. We assume furthermore that there exists a process $L(t,T)$ such that*

$$L^\tau(t,T) = 1_{\{\tau>t\}} L(t,T), \tag{4.97}$$

where $L(t,T)$ is $\mathbb{F}$-adapted, with $L(T,T) = 1$.

Assumption 4.5.3. *In addition to Assumptions 2.2.1, 2.2.2, 4.5.1 and 4.2.2, we assume that the market for T-longevity bonds is arbitrage free. We assume furthermore that, for every T, $L(t,T)$ in (4.97) is a deterministic function $\hat{L}$ of t and $z(t)$, where $\hat{L}$ is smooth and strictly positive.*

Note that we are using the same symbol $\hat{L}$ used in the previous Section 4.3, but the function of this section is, in general, different from the one used in the previous section. The latter condition is formally equal to condition (4.23), and therefore (formally) also (4.29) and (4.30) of Lemma 4.3.3 hold.

Obviously, under the above assumptions, the price process of a T-longevity bond has the form

$$L^\tau(t,T) = 1_{\{\tau>t\}} \hat{L}^T(t, z(t)). \tag{4.98}$$

Moreover we have the following boundary condition:

$$L^\tau(T,T) = 1_{\{\tau>T\}} \hat{L}^T(T, z(T)) = 1_{\{\tau>T\}}, \tag{4.99}$$

where in the last equality we have used the condition (4.23) according to the condition $L(T,T) = 1$.

The aim now is to find the price dynamics for T-longevity bonds, $dL^\tau(t,T)$. From (4.97), we have that

$$dL^\tau(t,T) = L(t,T) d\,1_{\{\tau>t\}} + 1_{\{\tau>t\}} dL(t,T), \tag{4.100}$$

then we need to specify $d\,1_{\{\tau>t\}}$ and $dL(t,T)$. First observe that, the process λ being a stochastic intensity, from Proposition 3.3.1 we obtain

$$d\,1_{\{\tau>t\}} = d\left(1 - 1_{\{\tau\le t\}}\right) = -d\,1_{\{\tau\le t\}} = -1_{\{\tau>t\}}\lambda(t)dt - dM^\tau(t), \tag{4.101}$$

where $M^\tau(t)$

$$M^\tau(t) = 1_{\{\tau\le t\}} - \int_0^t 1_{\{\tau>u\}}\lambda(u)du \tag{4.102}$$

is a martingale (see (4.4) with τ instead τ^j).

Secondly, by Assumption 4.5.3 and the multidimensional Itô formula, we have that the dynamics of $L(t,T)$ has the same form as in (4.26), (4.27) and (4.28), and Lemma 4.3.3 holds.

Finally, returning to the dynamics of $L^\tau(t,T)$, substituting (4.101) and (4.26) in (4.100), we obtain

$$\begin{aligned} dL^\tau(t,T) &= \left(-1_{\{\tau>t\}}\lambda(t)dt - dM^\tau(t)\right) L(t,T) + 1_{\{\tau>t\}} dL(t,T) \\ &= L^\tau(t,T)\left[\left(\mu_L(t,T) - \lambda(t)\right)dt + \sigma_L(t,T)dW^z(t)\right] - L(t,T)dM^\tau(t), \\ &= L^\tau(t,T)\left[\mu_L^\tau(t,T)dt + \sigma_L^\tau(t,T)dW^z(t)\right] - L(t,T)dM^\tau(t). \end{aligned} \tag{4.103}$$

where we have set

$$\mu_L^\tau(t,T) = \mu_L(t,T) - \lambda(t), \tag{4.104}$$

$$\sigma_L^\tau(t,T) = \sigma_L(t,T), \tag{4.105}$$

and $\mu_L(t,T)$, $\sigma_L(t,T)$ are deterministic functions of t and $z(t)$ as in (4.27), (4.28) respectively (see also (4.29), (4.30)).

We can now apply the approach of Section 4.3 to this setting. As observed above, the a priori given market consists of the benchmark bond $B(t,T_0)$ and the money market account $G(t)$. Observe that in this market the number M of random sources equals three (the 2-dimensional Wiener process, W^z, and the martingale, M^τ), while the number N of traded assets (besides $G(t)$) equals one. From Corollary 1.4.5, we may thus expect that the market is arbitrage-free, but not complete. Another way of seeing this problem appears if we try to price a certain T-longevity bond, using the technique in Section 4.3, i.e., all the longevity bonds are regarded as derivatives of the underlying process z, in other words a zero coupon longevity bond can be thought of as a derivative on z.

Since on our market there is the 2-dimensional Wiener process W^z and the martingale M^τ, i.e. three random sources, we can specify, besides the benchmark bond $B(t,T_0)$, for a fixed time T_0, the price processes of 2 different benchmark longevity bonds. The price processes of all other longevity bonds will then be uniquely determined by the prices of this benchmarks.

According to Theorem 1.4.6, the following central result extends Proposition 4.3.4 to this setting.

Proposition 4.5.1. *Assume that the bond and longevity bond market is arbitrage free. Fix two benchmarks longevity bonds, $L^\tau(t,T_0)$ and $L^\tau(t,N)$, the price processes of which are given by (4.103), (4.104) and (4.105), with $T = T_0, N$, and $T_0 \neq N$. Assume furthermore that $L^\tau(t,T_0)$ and $L^\tau(t,N)$ are such that*

$$\hat{\sigma}_{L,\lambda}^{T_0}(t,z(t)) \neq \hat{\sigma}_{L,\lambda}^{N}(t,z(t)), \quad \forall t \leq T_0 \wedge N. \tag{4.106}$$

Then there exists a process $\xi_z = (\xi_r, \xi_\lambda)$, and a process $\bar{\lambda}$, such that the so called market price of risk equations

$$\sigma(t,T)\xi_r(t) = \mu(t,T) - r(t), \tag{4.107}$$

$$\sigma_L^\tau(t,T)\xi_z(t) = \mu_L^\tau(t,T) - \bar{\lambda}(t) - r(t) \tag{4.108}$$

hold for all t a.s. and for every choice of maturity time T.

Observe that the condition (4.106) is the mathematical formulation of the requirement that the family of benchmark derivatives is rich enough to span the entire derivative space, as we will see from proof of Proposition 4.5.1. Furthermore observe that the component ξ_r given by (4.107) is the same process computed in Section 2.1.

Considering (4.104) and (4.105), we obtain immediately the following corollary of the previous proposition.

Corollary 4.5.2. *Under same hypotheses of Proposition 4.5.1, we have that the relation*

$$\sigma_L(t,T)\xi_z(t) = \mu_L(t,T) - \big(\lambda(t) + \bar{\lambda}(t)\big) - r(t). \tag{4.109}$$

holds for all t a.s. and for every choice of maturity time T.

Taking into account that the coefficients $\mu(t,T)$, $\sigma(t,T)$, $\mu_L^\tau(t,T)$, $\sigma_L^\tau(t,T)$ are deterministic functions of t and $z(t)$, the same holds for $\xi_z(t)$, namely

$$\xi_z(t) = \tilde{\xi}_z(t,z(t)), \tag{4.110}$$

and for $\bar{\lambda}(t)$: indeed equation (4.109) reads

$$\tilde{\sigma}_L^T(t,z(t))\tilde{\xi}_z(t,z(t)) = \hat{\mu}_L^T(t,z(t)) - \big(\lambda(t) + \bar{\lambda}(t)\big) - r(t). \tag{4.111}$$

Analogously to ξ_r (see Section 2.1), the component ξ_λ has the dimension "risk premium per unit of λ-type volatility", so that ξ_λ is called the market price for the longevity risk due to W^λ. Similarly, we call $\bar{\lambda}$ the market price for the longevity risk due to M^τ, so that if the process $\lambda(t) + \bar{\lambda}(t)$ is a.s. positive, the latter process is characterized as the risk-neutral intensity[11].

[11] The idea is the following. Assume that we can define a measure Q^τ such that the processes

$$\bar{W}^r(t) = W^r(t) + \int_0^t \xi_r(u)\,du, \qquad \bar{W}^\lambda(t) = W^\lambda(t) + \int_0^t \xi_\lambda(u)\,du,$$

Proof of Proposition 4.5.1. We have already proved (4.107) in Section 2.1 (see 2.2.2), then we turn to prove (4.108). By Assumption 4.2.2 and by hypotheses of Proposition 4.5.1, we have one benchmark bond and two different benchmark longevity bonds with price process of the form

$$\begin{aligned}
B(t,T_0) &= \hat{B}^{T_0}(t,r(t)),\\
L^\tau(t,T_0) &= 1_{\{\tau>t\}}\hat{L}^{T_0}(t,z(t)),\\
L^\tau(t,N) &= 1_{\{\tau>t\}}\hat{L}^{N}(t,z(t)),
\end{aligned}$$

where $B(t,T)$ is a zero coupon bond of Section 2.1 (see (2.7) and Lemma 2.2.1). In order to simplify the notation, we will write T instead of T_0.

Considering a zero coupon longevity bond of maturity $S\neq T,N$, we have the corresponding equation for the S-longevity bond

$$dL^\tau(t,S)=L^\tau(t,S)\left[\mu_L^\tau(t,S)dt+\sigma_L^\tau(t,S)dW^z(t)\right]-L(t,S)dM^\tau(t). \tag{4.112}$$

where analogously to (4.104) and (4.105)

$$\mu_L^\tau(t,S)=\mu_L(t,S)-\lambda(t), \tag{4.113}$$
$$\sigma_L^\tau(t,S)=\sigma_L(t,S), \tag{4.114}$$

and analogously to (4.29) and (4.30)

$$\hat{\mu}_L^S(t,z(t)) = \frac{\hat{L}_t^S(t,z(t))+\hat{L}_z^S(t,z(t))\tilde{\mu}^z(t,z(t))+\frac{1}{2}tr\left[\left(\tilde{\Sigma}^z\right)'\hat{L}_{zz}^S\tilde{\Sigma}^z\right](t,z(t))}{\hat{L}^S(t,z(t))} \tag{4.115}$$

$$\tilde{\sigma}_L^S(t,z(t)) = \frac{\hat{L}_z^S(t,z(t))\tilde{\Sigma}^z(t,z(t))}{\hat{L}^S(t,z(t))}=\left(\frac{\hat{L}_r^S\hat{\sigma}^r}{\hat{L}^S},\frac{\hat{L}_\lambda^S\hat{\sigma}^\lambda}{\hat{L}^S}\right)(t,z(t)). \tag{4.116}$$

We now form a portfolio based only on $B(t,T)$, $L^\tau(t,T)$, $L^\tau(t,N)$, and $L^\tau(t,S)$, and as in proof of Proposition 4.3.4, in the present setting nothing will be invested in the bank or loaned by the bank. Thus (see

are Wiener standard processes under Q^τ, and

$$\bar{M}^\tau(t)=M^\tau(t)-\int_0^t 1_{\{\tau>u\}}\bar{\lambda}(u)\,du=1_{\{\tau\le t\}}-\int_0^t 1_{\{\tau>u\}}\big(\lambda(u)+\bar{\lambda}(u)\big)\,du$$

is a martingale under Q^τ,i.e., the mortality intensity under Q^τ is $\lambda(t)+\bar{\lambda}(t)$.

It is important to note that the latter condition is possible if and only if the process $\lambda(t)+\bar{\lambda}(t)$ is a.s. positive. Unfortunately, though the latter condition is intuitive, we were not able to prove this property using only the hypotheses of Proposition 4.5.1.

Then from (4.103), we get

$$\begin{aligned}
dL^\tau(t,T) &= L^\tau(t,T)\left[\mu_L^\tau(t,T)dt+\sigma_L^\tau(t,T)dW^z(t)\right]-L(t,T)dM^\tau(t)\\
&= L^\tau(t,T)\left[\mu_L^\tau(t,T)dt+\sigma_L^\tau(t,T)\big(d\bar{W}^z(t)-\xi_z(t)\,dt\big)\right]\\
&\qquad -L(t,T)\big(d\bar{M}^\tau(t)+1_{\{\tau>t\}}\bar{\lambda}(t)\,dt\big)\\
&= L^\tau(t,T)\left[\big(\mu_L^\tau(t,T)-\sigma_L^\tau(t,T)\xi_z(t)\,dt\big)dt+\sigma_L^\tau(t,T)d\bar{W}^z(t)\right]\\
&\qquad -L(t,T)d\bar{M}^\tau(t)-1_{\{\tau>t\}}\,L(t,T)\,\bar{\lambda}(t)\,dt\\
&= L^\tau(t,T)\left[\big(\mu_L(t,T)-\lambda(t)-\bar{\lambda}(t)-\sigma_L^\tau(t,T)\xi_z(t)\big)dt+\sigma_L^\tau(t,T)d\bar{W}^z(t)\right]-L(t,T)d\bar{M}^\tau(t)
\end{aligned}$$

and condition (4.111) is equivalent to require that, Q^τ is a risk-neutral measure, since, under Q^τ,

$$dL^\tau(t,T)=L^\tau(t,T)\left[r(t)\,dt+\sigma_L^\tau(t,T)d\bar{W}^z(t)\right]-L(t,T)d\bar{M}^\tau(t).$$

The measure Q^τ is then defined as the unique probability measure such that, for all times $t\ge 0$

$$\frac{dQ_t^\tau}{dP_t}=\mathcal{Z}_t\,\mathcal{Z}_t^\tau$$

where Q_t^τ and P_t denote the restrictions of Q^τ and P to $\mathcal{F}_t^N\vee\mathcal{F}_t^r\vee\mathcal{F}_t^\lambda$,

$$\mathcal{Z}_t=\exp\left(-\int_0^t\xi_z(u)dW^z(u)-\frac{1}{2}\int_0^t\xi_z'(u)\xi_z(u)du\right)$$

and

$$\mathcal{Z}_t^\tau=\prod_{\tau\le t}\frac{\lambda(\tau^-)+\bar{\lambda}(\tau^-)}{\lambda(\tau^-)}\exp\left(-\int_0^t 1_{\{\tau>u\}}\bar{\lambda}(u)\,du\right).$$

(Here we have used the convention that $\prod_{i\in\emptyset}a_i=1$.)

Section 1.3) let $h(t) = (h_0(t), h_1(t), h_2(t), h_3(t), h_4(t))$ be the portfolio associated to $X = (X_0, X_1, X_2, X_3)$, where

$$X_0 = G(t), \quad X_1 = B(t,T), \quad X_2 = L^\tau(t,T), \quad X_3 = L^\tau(t,N), \quad X_4 = L^\tau(t,S), \tag{4.117}$$

and

$$h_0(t) = h^G(t) = 0, \quad (h_1(t), h_2(t), h_3(t), h_4(t)) = (h_T(t), h^L_T(t), h^L_N(t), h^L_S(t)). \tag{4.118}$$

Similar to Section 4.3, instead of specifying the absolute number of shares held of a certain asset, i.e. $h(t)$, it may be convenient to consider the corresponding relative portfolio $(U_T(t)U^L_T(t), U^L_N(t), U^L_S(t))$. Setting $u(t) = (u_T(t), u^L_T(t), u^L_N(t), u^L_S(t))'$, by (1.16) and (1.17), we have

$$\begin{aligned}
U_T(t) &= 1_{\{B(t,T)>0\}} u_T(t) = u_T(t) \\
U^L_T(t) &= 1_{\{L^\tau(t,T)>0\}} u^L_T(t) = 1_{\{\tau>t\}} u^L_T(t) \\
U^L_N(t) &= 1_{\{L^\tau(t,N)>0\}} u^L_N(t) = 1_{\{\tau>t\}} u^L_N(t) \\
U^L_S(t) &= 1_{\{L^\tau(t,S)>0\}} u^L_S(t) = 1_{\{\tau>t\}} u^L_S(t),
\end{aligned}$$

for the relative portfolio corresponding to $B(t,T)$, $L^\tau(t,T)$, $L^\tau(t,N)$ and $L^\tau(t,S)$, with (4.39) holds.

The dynamics of the value process for the corresponding self-financing portfolio (see (1.18) and (2.26)) are given by

$$\begin{aligned}
&\frac{dV(t)}{V(t)} \\
&= u_T(t)\frac{dB(t,T)}{B(t,T)} + 1_{\{\tau>t\}}\left(u^L_T(t)\frac{dL^\tau(t,T)}{L^\tau(t,T)} + u^L_N(t)\frac{dL^\tau(t,N)}{L^\tau(t,N)} + u^L_S(t)\frac{dL^\tau(t,S)}{L^\tau(t,S)}\right).
\end{aligned} \tag{4.119}$$

The price processes for T-bond, (see (2.9), (2.10) and (2.11)), with respect to dW^z are given by (4.42).

Then, inserting in (4.119) the dynamics (4.103) and (4.42) of the price processes involved, by (4.104), (4.105), (4.113) and (4.114), we get

$$\begin{aligned}
&\frac{dV(t)}{V(t)} \\
&= \Big[u_T(t)\hat{\mu}^T + u^L_T(t)\Big(\hat{\mu}^T_L - \lambda(t)\Big) + u^L_N(t)\Big(\hat{\mu}^N_L - \lambda(t)\Big) + u^L_S(t)\Big(\hat{\mu}^S_L - \lambda(t)\Big)\Big]dt \\
&+ \Big[u_T(t)\tilde{\sigma}^T + u^L_T(t)\tilde{\sigma}^T_L + u^L_N(t)\tilde{\sigma}^N_L + u^L_S(t)\tilde{\sigma}^S_L\Big]dW^z(t) \\
&- \Big[u^L_T(t) + u^L_N(t) + u^L_S(t)\Big]dM^\tau(t),
\end{aligned}$$

where for the notational convenience, the arguments $(t, r(t))$ and $(t, z(t))$ "have been suppressed", so that we have used the same shorthand notations (4.43) and (4.44) for the process $B(t,T)$, $L(t,T)$ and similarly for the process $L(t,S)$ and $L(t,N)$. Here, when it is convenient, we will use the above notations (4.43) and (4.44). We try to choose $u_T(t)$, $u^L_T(t)$, $u^L_S(t)$ and $u^L_N(t)$, so that the market is arbitrage-free. Now, we can extend Proposition 1.4.1 to our case, setting $I = (0, \tau)$. Indeed we have

$$P(|I| > 0) = P(\tau > 0) = 1, \tag{4.120}$$

thus the proof of Proposition 1.4.1 is valid even if I is a stochastic interval. Then, the portfolio rate of return and the short rate of interest must be equal, namely[12]

$$u_T(t)\hat{\mu}^T + u^L_T(t)\Big(\hat{\mu}^T_L - \lambda(t)\Big) + u^L_N(t)\Big(\hat{\mu}^N_L - \lambda(t)\Big) + u^L_S(t)\Big(\hat{\mu}^S_L - \lambda(t)\Big) = r(t), \tag{4.121}$$

necessarily holds for all t, with probability 1, and then, using (4.39), we obtain for all t

$$\begin{aligned}
&u_T(t)\Big(\hat{\mu}^T(t, r(t)) - r(t)\Big) + u^L_T(t)\Big(\hat{\mu}^T_L(t, z(t)) - \lambda(t) - r(t)\Big) \\
&+ u^L_N(t)\Big(\hat{\mu}^N_L(t, z(t)) - \lambda(t) - r(t)\Big) + u^L_S(t)\Big(\hat{\mu}^S_L(t, z(t)) - \lambda(t) - r(t)\Big) = 0.
\end{aligned} \tag{4.122}$$

[12]For the notational convenience we are using the notations (4.43) and (4.44).

Moreover we look for a portfolio minimizing the risk associated to the derivative, i.e., such that the corresponding value process has no driving Wiener process, W^z, and no martingale M^τ. This means that we want to solve the equations

$$u_T(t)\tilde{\sigma}^T(t,r(t)) + u_T^L(t)\tilde{\sigma}_L^T(t,z(t)) + u_N^L(t)\tilde{\sigma}_L^N(t,z(t)) + u_S^L(t)\tilde{\sigma}_L^S(t,z(t)) = 0, \tag{4.123}$$

$$u_T^L(t,z(t)) + u_N^L(t,z(t)) + u_S^L(t,z(t)) = 0. \tag{4.124}$$

Observe that the equation (4.124) together with (4.39) implies that $u_T(t) = 1$, i.e., we invest in the benchmark bond $B(t,T)$, choosing h_T^L, h_N^L, and h_S^L such that $h_T^L(t)L(t,T) + h_N^L(t)L(t,N) + h_S^L(t)L(t,S) = 0$. In order to see some structure, let H be the following matrix

$$H(t,z) = H(t,r,\lambda) = \begin{pmatrix} \hat{\mu}^T - r & \hat{\mu}_L^T - \lambda - r & \hat{\mu}_L^N - \lambda - r & \hat{\mu}_L^S - \lambda - r \\ \hat{\sigma}^T & \hat{\sigma}_{L,r}^T & \hat{\sigma}_{L,r}^N & \hat{\sigma}_{L,r}^S \\ 0 & \hat{\sigma}_{L,\lambda}^T & \hat{\sigma}_{L,\lambda}^N & \hat{\sigma}_{L,\lambda}^S \\ 0 & 1 & 1 & 1 \end{pmatrix} \tag{4.125}$$

so that we now write (4.122)–(4.124) in matrix form as

$$H(t,z(t))u(t) = H(t,r(t),\lambda(t))u(t) = 0, \tag{4.126}$$

where we have used the notations (4.43) and (4.44). If H were invertible, then the system (4.126) would have a unique solution, i.e., the null solution, but this solution does not satisfy the condition (4.39), then H must be singular. For readability reasons, we study $H^{'}$, the transpose of H, i.e.,

$$H'(t,z) = H'(t,r,\lambda) = \begin{pmatrix} \hat{\mu}^T - r & \hat{\sigma}^T & 0 & 0 \\ \hat{\mu}_L^T - \lambda - r & \hat{\sigma}_{L,r}^T & \hat{\sigma}_{L,\lambda}^T & 1 \\ \hat{\mu}_L^N - \lambda - r & \hat{\sigma}_{L,r}^N & \hat{\sigma}_{L,\lambda}^N & 1 \\ \hat{\mu}_L^S - \lambda - r & \hat{\sigma}_{L,r}^S & \hat{\sigma}_{L,\lambda}^S & 1 \end{pmatrix}. \tag{4.127}$$

The matrix H' being singular, the columns are linearly dependent. Since under the condition (4.106), i.e., $\hat{\sigma}_L^{\lambda,T}(t,z(t)) \neq \hat{\sigma}_L^{\lambda,N}(t,z(t))$, the matrix

$$\sigma = \begin{pmatrix} \hat{\sigma}^T(t,r(t)) & 0 & 0 \\ \hat{\sigma}_{L,r}^T(t,z(t)) & \hat{\sigma}_{L,\lambda}^T(t,z(t)) & 1 \\ \hat{\sigma}_{L,r}^N(t,z(t)) & \hat{\sigma}_{L,\lambda}^N(t,z(t)) & 1 \end{pmatrix}.$$

is invertible (with probability 1 for each t), the first column of $H^{'}$ can be written as a linear combination of the other columns. We thus deduce the existence of the 3-dimensional process $\xi = \big(\xi_r, \xi_\lambda, \bar{\lambda}\big)'$ such that setting $\mathbf{1} = (1,1,1)'$ and $\mathbf{1}_\lambda = (0,1,1)'$,

$$\sigma\xi = \mu - \lambda\mathbf{1}_\lambda - r\mathbf{1}, \qquad \text{i.e.,} \qquad \begin{cases} \big(\tilde{\sigma}^T(t,r(t)),0\big)\xi(t) &= \hat{\mu}^T(t,r(t)) - r(t) \\ \big(\tilde{\sigma}_L^T(t,z(t)),1\big)\xi(t) &= \hat{\mu}_L^T(t,z(t)) - \lambda(t) - r(t) \\ \big(\tilde{\sigma}_L^N(t,z(t)),1\big)\xi(t) &= \hat{\mu}_L^N(t,z(t)) - \lambda(t) - r(t) \end{cases} \tag{4.128}$$

and therefore

$$\big(\tilde{\sigma}_L^S(t,z(t)),1\big)\xi(t) = \hat{\mu}_L^S(t,z(t)) - \lambda(t) - r(t), \tag{4.129}$$

or equivalently

$$\tilde{\sigma}_L^S(t,z(t))\xi_z(t) = \hat{\mu}_L^S(t,z(t)) - (\lambda(t) + \bar{\lambda}(t)) - r(t). \tag{4.130}$$

Since the longevity bond $L(t,S)$ was chosen arbitrarily, the risk premium, $\hat{\mu}_L^S(t,z(t)) - (\lambda(t)+\bar{\lambda}(t)) - r(t)$, of any longevity bond, can be written as a linear combination of its volatility components, $\tilde{\sigma}_L^S(t,z(t))$, $\xi_z(t)$ being the same for all longevity bonds. Thus equations (4.128) and (4.129) show that the process ξ does not depend on the choice of either S or T, and that the process ξ is uniquely defined by (4.128).

□

Observe that under suitable conditions we can find the same kind of results of Section 4.3. To this end we make the following assumption.

Assumption 4.5.4. *We assume that*

$$\bar{\lambda} = 0.$$

By Assumption 4.5.4 and by (4.109) we obtain that the market price $\xi_z(t)$ is given exactly by (4.32) and (4.33). Therefore the risk neutral measure Q^τ is formally defined as the measure Q defined in (4.57) and the doubly stochastic Poisson process $N(t)$ defining τ has the intensity $\lambda(t)$ also under Q^τ. Therefore from now on we will write Q instead of Q^τ

Taking into account (4.97) and (4.98), we observe that our aim is to determine the function $\hat{L}^\tau(t,z)$, then by Theorem 4.3.5 and Proposition 4.3.6 (see (4.53), (4.60)) we obtain that the longevity bond price processes $L^\tau(t,T)$ are given by[13]

$$L^\tau(t,T) = 1_{\tau>t} L(t,T) = 1_{\tau>t} E^Q_{t,z}\left(e^{-\int_t^T r(s)ds} e^{-\int_t^T \lambda(s)ds}\right)|_{z=z(t)} \, , \tag{4.131}$$

where $L(t,T)$ is price process obtain in the previous Section 4.3.

Rewriting formula (4.131) as

$$L^\tau(t,T) = 1_{\{\tau>t\}} E^Q_{t,z}\left[e^{-\int_t^T (r(s)+\lambda(s))ds} \cdot 1\right]|_{z=z(t)} \, , \tag{4.132}$$

we observe that, if the holder is alive at time T, the value of a T-longevity bond at time t is given as the expected value of one dollar (final payoff), discount to present value at the interest rate given by $\lambda + r$. Thus formula (4.132) can be interpreted as the risk-neutral pricing formula for a T-bond at the interest rate given by $r + \lambda$.

Remark 4.5.1. *Observe that if we consider a financial market consisting of both assets $L(t,T)$ of Section 4.3 and $L^\tau(t,T)$ with $\bar{\lambda} = 0$ (with $\tau > t$), then we have an arbitrage on the financial market. Indeed, if we buy a T-longevity bond $L(t,T)$ and we sell a new T-longevity bond $L^\tau(t,T)$, then the net investment at time t is zero, whereas our wealth at any time $s > t$ will be positive with positive probability. Therefore the two longevity bonds cannot be traded on the same market with the prices $L^\tau(t,T)$ determined by $\bar{\lambda} = 0$.*

[13]Recall that by Markov property of $z(t)$ with respect to the filtration $\mathcal{H}_t$, where $\mathcal{H}_t = \mathcal{F}^N_t \vee \mathcal{F}^r_t \vee \mathcal{F}^\lambda_t$, we have

$$E^Q\left(e^{-\int_t^T r(s)ds} e^{-\int_t^T \lambda(s)ds} |\mathcal{H}_t\right). = E^Q_{t,z}\left(e^{-\int_t^T r(s)ds} e^{-\int_t^T \lambda(s)ds}\right)\Big|_{z=z(t)}.$$

Chapter 5

The Optimal Portfolio

5.1 Introduction

So far we have described a bonds market model, a bond-stock market model, and a bond-longevity bond market model in Section 2.2, 2.3, and 4.3, respectively.

In many concrete applications, it is natural to consider an optimal control problem. In particular we study the optimal consumption and asset allocation problem for an agent with a stochastic time horizon coinciding with her/his death.

The object of the agent is to maximize the expected utility of her/his consumption in a market model with a riskless asset, a stock, a T-bond, and a T-longevity bond, where T is a suitable deterministic time such that, on the basis of demographic considerations, at time T the agent will be dead (for example, for an agent that at time 0 is 65 years old, the time T should be taken greater than or equal to 35). Since in the real market a bond and a longevity bond with such a maturity T do not exist, then we introduce a market model more realistic than the previous model, introducing a rolling bond and a rolling longevity bond on the market.

In this chapter we focus on solving these optimal problems following the (stochastic) dynamic programming approach via the so-called Hamilton-Jacobi-Bellman equation, which is a second order (in the stochastic case) partial differential equation, and the verification technique. Note that this approach actually gives solutions to the whole family of problems (with different initial times and states), and in particular, the original problem.

We refer to Fleming and Soner [12] for the optimal portfolio and (stochastic) dynamic programming theory, to Menoncin [18] for the case of a market with longevity bonds, and to Rutkowski [20] for the rolling bond.

5.2 Financial Market with Longevity Bond (BLS market)

In this section we present a financial model on which we will work: we consider a market model which, besides the money account $G(t)$ and the risk asset with price process $S(t)$, contains a (zero coupon) T-bond and a (zero coupon) T-longevity bond, with price processes $B(t,T)$ and $L(t,T)$, where T is a suitable deterministic time such that, on the basis of demographic considerations, at time T the agent will be dead. The latter bonds are introduced in Sections 2.2 and 4.3, respectively. In the sequel we will shortly refer to this market as the BLS market model.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let τ be the death time of the investor, and the vector process $z(t) = (r(t), \lambda(t))$ be the state variables vector where the processes $r(t)$ and $\lambda(t)$ are referred to as the riskless interest rate, and the stochastic mortality intensity of the investor, respectively. We will discuss later on the conditions of the filtration.

Summarizing, we assume that

$$P(\tau \leq T) = 1 \tag{5.1}$$

and that the market is described by two structures, i.e., the so called state variables described by the vector process $z(t) = (r(t), \lambda(t))$, and the financial assets traded on the market. In details, using the notations introduced in the previous chapters, the vector process $z(t)$ evolves as follows

$$dz(t) = \mu^z(t)dt + \Sigma^z(t)dW^z(t), \tag{5.2}$$

(see (4.9), (4.10) and (4.11)), the money market account $G(t)$ is given by (1.3), and the financial assets are

1. A zero coupon bond, with maturity T, with price process $B(t,T)$;

2. A zero coupon longevity bond, with maturity T, with price process $L(t,T)$;
3. A risk asset with price process $S(t)$.

Furthermore, since (5.1) holds, then (see Lemma 3.3.7) the process $\lambda(t)$ satisfies the following condition

$$P\left(\int_{t_0}^{T} \lambda(u)du = \infty\right) = 1, \quad \forall t_0 \geq 0.$$

By the results obtained in Chapters 2 and 4, let ξ_S be the market price for the stock given by (2.70), and let $\xi_z = (\xi_r, \xi_\lambda)'$ be the market price for the riskless interest rate and the longevity risk given by (4.32) and (4.33). In the sequel we denote by ξ the market price, where

$$\xi(t) = \tilde{\xi}(t, z(t), S(t)) = \left(\hat{\xi}_r(t, r(t)), \hat{\xi}_\lambda(t, z(t)), \hat{\xi}_S(t, r(t), S(t))\right)'. \tag{5.3}$$

Then the processes $B(t,T)$, $L(t,T)$ and $S(t)$ can be described by the differential equations (2.37), (4.54) and (2.69), so that we can summarize the BLS market structures in the follow matrix form

$$dz(t) = \mu^z(t)dt + \Pi(t)dW(t), \tag{5.4}$$
$$dA(t) = diag\big[A(t)\big]\left(\mu^A(t)dt + \Sigma^A(t)dW(t)\right), \tag{5.5}$$

where

$$\mu^z(t) = \tilde{\mu}^z(t, z(t)) = \begin{pmatrix} \hat{\mu}^r(t, r(t)) \\ \hat{\mu}^\lambda(t, z(t)) \end{pmatrix}, \qquad W(t) = \begin{pmatrix} W^r(t) \\ W^\lambda(t) \\ W^S(t) \end{pmatrix}, \tag{5.6}$$

$$\Pi(t) = \tilde{\Pi}(t, z(t)) = \begin{pmatrix} \hat{\sigma}^r(t, r(t)) & 0 & 0 \\ 0 & \hat{\sigma}^\lambda(t, z(t)) & 0 \end{pmatrix}, \tag{5.7}$$

and

$$A(t) = \tilde{A}(t, z(t), S(t)) = \begin{pmatrix} \hat{B}^T(t, r(t)) \\ \hat{L}^T(t, z(t)) \\ S(t) \end{pmatrix}, \tag{5.8}$$

$$diag\big[A(t)\big] = diag\big[\tilde{A}(t, z(t), S(t))\big] = \begin{pmatrix} \hat{B}^T(t, r(t)) & 0 & 0 \\ 0 & \hat{L}^T(t, z(t)) & 0 \\ 0 & 0 & S(t) \end{pmatrix}, \tag{5.9}$$

$$\begin{aligned} \mu^A(t) &= \tilde{\mu}^A(t, z(t), S(t)) \\ &= \begin{pmatrix} r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\frac{\hat{B}^T_r}{\hat{B}^T}(t, r(t)) \\ r(t) + \lambda(t) + \frac{\hat{L}^T_z}{\hat{L}^T}(t, z(t))\tilde{\Sigma}^z(t, z(t))\hat{\xi}_z(t, z(t)) \\ r(t) + \hat{\sigma}^S_r(t, r(t), S(t))\hat{\xi}_r(t, r(t)) + \hat{\sigma}^S_S(t, r(t), S(t))\hat{\xi}_S(t, r(t), S(t)) \end{pmatrix}, \end{aligned} \tag{5.10}$$

$$\begin{aligned} \Sigma^A(t) &= \tilde{\Sigma}^A(t, z(t), S(t)) \\ &= \begin{pmatrix} \hat{\sigma}^r(t, r(t))\frac{\hat{B}^T_r}{\hat{B}^T}(t, r(t)) & 0 & 0 \\ \frac{\hat{L}^T_r}{\hat{L}^T}(t, z(t))\hat{\sigma}^r(t, r(t)) & \frac{\hat{L}^T_\lambda}{\hat{L}^T}(t, z(t))\hat{\sigma}^\lambda(t, z(t)) & 0 \\ \hat{\sigma}^S_r(t, r(t), S(t)) & 0 & \hat{\sigma}^S_S(t, r(t), S(t)) \end{pmatrix}. \end{aligned} \tag{5.11}$$

Remark 5.2.1. *Observe that $\mu^A(t)$ and $\Sigma^A(t)$ are deterministic functions of t, $z(t)$, and $S(t)$ since on the one hand the drift and diffusion coefficients of $z(t)$ are deterministic functions of t and $z(t)$, and on the other hand the drift and diffusion coefficients of $S(t)$ are deterministic functions of t, $r(t)$ and $S(t)$ (see the condition 2. of Remark 2.3.1), i.e, $\mu^S(t) = \hat{\mu}^S(t, r(t), S(t))$, $\sigma^S_r(t) = \hat{\sigma}^S_r(t, r(t), S(t))$ and $\sigma^S_S(t) = \hat{\sigma}^S_S(t, r(t), S(t))$.*

In the sequel we assume the following standing conditions.

Assumption 5.2.1. *We assume that the matrix $\Sigma^A(t)$ is invertible, i.e., the financial market is complete (see Corollary 1.4.5).*

Remark 5.2.2. *Let $\Sigma^A(t)$ be given by (5.11). Then the BLS market is complete whenever*

$$\hat{\sigma}^r(t,r(t))\hat{B}_r^T(t,r(t)) > 0, \quad \hat{\sigma}^\lambda(t,z(t))\hat{L}_\lambda^T(t,z(t)) > 0, \quad \hat{\sigma}_S^S(t,r(t),S(t)) > 0 \qquad \forall(t,\omega). \tag{5.12}$$

Indeed, since the matrix $\Sigma^A(t)$ is lower triangular, the functions $\hat{B}^T(t,r(t))$ and $\hat{L}^T(t,z(t))$ are strictly positive (see Assumptions 2.2.1 and 4.3.2), the previous conditions implies that $\Sigma^A(t)$ is invertible and market completeness follows by Corollary 1.4.5.

In order to model the evolution of the stochastic mortality intensity, $\lambda(t)$, we assume that $N(t)$ is a doubly stochastic Poisson process respect to $\mathcal{G}$ as defined in (3.33), i.e.,

$$N(t) = \hat{N}\left(\int_0^t \lambda(u)du\right),$$

where the standard Poisson process $\hat{N}(t)$ is independent of the intensity process $\lambda(t)$, with respect to a suitable filtration. Before specifying the filtration we introduce a further process, the investor wealth process $V(t)$, and consider the multidimensional process $(z(t), S(t), V(t))$ (see the next Section 5.2.1). Section 5.2.2 is devoted to the assumptions on the filtration.

5.2.1 The investor's wealth in BLS market

We now form a portfolio (see Section 1.3) associated to $G(t)$, $B(t,T)$, $L(t,T)$, and $S(t)$, i.e., let $h(t) = (h_0(t), h_1(t), h_2(t), h_3(t))$ be the portfolio associated to $X = (X_0, X_1, X_2, X_3)$, where

$$X_0 = G(t), \quad X_1 = B(t,T), \quad X_2 = L(t,T), \quad X_3 = S(t),$$

and

$$h_0(t) = h^G(t), \quad (h_1(t), h_2(t), h_3(t)) = (h_T^B(t), h_T^L(t), h^S(t)) = h^A(t).$$

Denoting the consumption rate by the process $C(t)$, we assume that (h, C) is a self-financing portfolio-consumption pair. Similarly to Section 1.3, instead of specifying $h(t)$, the absolute number of shares held of a certain asset, it may be convenient to consider $(U^G(t), U_T^B(t), U_T^L(t), U^S(t))$, the corresponding relative portfolio. By (1.16) and (1.17) we have

$$\begin{aligned} U^G(t) &= 1_{\{G(t)>0\}}\, u^G(t) &&= u^G(t) \\ U_T^B(t) &= 1_{\{B(t,T)>0\}} u_T^B(t) &&= u_T^B(t) \\ U_T^L(t) &= 1_{\{L(t,T)>0\}} u_T^L(t) &&= u_T^L(t) \\ U^S(t) &= 1_{\{S(t)>0\}}\, u^S(t) &&= u^S(t) \end{aligned}$$

for the relative portfolio corresponding to $G(t)$, $B(t,T)$, $L(t,T)$, and $S(t)$, with

$$u^G(t) + u_T^B(t) + u_T^L(t) + u^S(t) = 1. \tag{5.13}$$

Since here T is fixed, from now on we will drop the subscript T in $u_T^B(t)$ and $u_T^L(t)$, and so we write $u^B(t)$ and $u^L(t)$, respectively. Let $u^A(t) = (u^B(t), u^L(t), u^S(t))$ be the relative portfolio corresponding to $h^A(t)$. The dynamics of the value process for the self-financing portfolio-consumption pair (see (1.21)) are given by

$$\begin{cases} dV(t) = V(t)\left[u^G(t)\frac{dG(t)}{G(t)} + u^B(t)\frac{dB(t,T)}{B(t,T)} + u^L(t)\frac{dL(t,T)-dD(t,T)}{L(t,T)} + u^S(t)\frac{dS(t)}{S(t)}\right] - C(t)dt, \\ V(t_0) = V, \end{cases}$$

or in the compact form

$$\begin{cases} dV(t) = V(t)\left[u^G(t)\frac{dG(t)}{G(t)} + u^A(t)diag^{-1}[A(t)]dA(t) - u^L(t)\frac{dD(t,T)}{L(t,T)}\right] - C(t)dt, \\ V(t_0) = V, \end{cases}$$

where

$$u^A(t) = \left(u^B(t), u^L(t), u^S(t)\right). \tag{5.14}$$

After substituting the expression for u^G taken from (5.13), i.e.,

$$u^G(t) = 1 - (u_T^B(t) + u_T^L(t) + u^S(t)) = 1 - u^A(t)\mathbf{1}, \tag{5.15}$$

where $\mathbf{1} = (1,1,1)'$, the dynamics of the process $V(t)$ can be written as

$$dV(t) = V(t)\left[(1 - u^A(t)\mathbf{1})\,\frac{dG(t)}{G(t)} + u^A(t)diag^{-1}\big[A(t)\big]dA(t) - u^L(t)\frac{dD(t,T)}{L(t,T)}\right] - C(t)dt,$$

so that, by the expression (5.5) for the differential form dA, and (4.24) for dD after some simplifications, we obtain

$$\begin{aligned} dV(t) &= V(t)\Big[\big(1 - u^A(t)\mathbf{1}\big)r(t)dt + u^A(t)\big(\mu^A(t)dt + \Sigma^A(t)dW(t)\big) - u^L(t)\frac{dD(t,T)}{L(t,T)}\Big] - C(t)dt \\ &= \Big[V(t)r(t) + V(t)u^A(t)\big(\mu^A(t) - r(t)\mathbf{1} - \lambda(t)\mathbf{1}_\lambda\big) - C(t)\Big]dt + V(t)u^A(t)\Sigma^A(t)dW(t), \end{aligned} \tag{5.16}$$

where $\mathbf{1}_\lambda = (0,\,1,\,0)'$.

Let us consider the agent at time t_0 with a stochastic time horizon τ, coinciding with her/his death time, i.e., she/he will act in the time interval $[t_0, \tau)$. At time t_0 the agent has the initial wealth V, and her/his problem is how to allocate investments and consumption over the time horizon. Since the admissible strategies involve consumption, and we restrict the investment-consumption pair to be self-financing, the second fundamental asset pricing theorem (see Theorem 1.4.3) is not valid. Then the objective of the agent is to choose a portfolio-consumption strategy to maximizing her/his preferences. Formally we are considering a stochastic optimal control problem. In Appendix D.1 we focus on some necessary mathematical tools for studying a general class of optimal control problems.

5.2.2 Assumptions on the filtration

Now we extend Assumption 4.2.1 and condition (4.15) on the σ-algebra $\mathcal{G}$ to this setting so that we have

$$\begin{aligned} &\mathcal{F}_t^\lambda \vee \mathcal{F}_t^r \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V \vee \mathcal{F}_t^N \subseteq \mathcal{F}_t, \quad \forall t \in [0,T]\,, \\ &\mathcal{G} \supset \mathcal{F}_\infty^r \vee \mathcal{F}_\infty^\lambda \vee \mathcal{F}_\infty^S \vee \mathcal{F}_\infty^V. \end{aligned}$$

As we will see below, in this setting it is necessary to distinguish $\mathcal{F}_t^N$ from all other filtrations. To this end we introduce a filtration $\mathbb{G}$ containing $\mathcal{F}_t^\lambda \vee \mathcal{F}_t^r \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V$. Recallind that in Section 2.3, by Assumptions 1.2.1 and 1.2.2, we have considered the augmented filtration associated to the process W^S, i.e., $\bar{\mathbb{F}}^{W^S}$, in the sequel, according to (4.3), we assume that

$$\bar{\mathbb{F}}^W \subset \mathbb{G}.$$

Summarizing we formalize the above assumptions as follows.

Assumption 5.2.2. *We assume that on $(\Omega, \mathcal{F}, P)$ there exists a σ-algebra $\mathcal{G}$ and a filtration $\mathbb{G}$ such that $\forall t \in [0,T]$*

$$\bar{\mathcal{F}}_t^W \subset \mathcal{G}_t, \tag{5.17}$$

$$\mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V \subseteq \mathcal{G}_t, \tag{5.18}$$

and

$$\mathcal{F}_\infty^r \vee \mathcal{F}_\infty^\lambda \vee \mathcal{F}_\infty^S \vee \mathcal{F}_\infty^V \subseteq \mathcal{G}. \tag{5.19}$$

As already discussed, the crucial point is the filtration with respect to which the process λ is a stochastic mortality intensity. In particular we recall that, usually, the stochastic intensity is considered with respect to a filtration $\mathbb{H}$ satisfying the usual conditions and such that

$$\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t \subseteq \mathcal{G} \vee \mathcal{F}_t^N, \quad \forall t \in [0,T]\,.$$

Furthermore, by Proposition 3.3.4 we know that the $\mathbb{H}$-stochastic intensity is still λ. In particular we can take

$$\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_t^N. \tag{5.20}$$

5.3 Optimal control problem with a Longevity Bond

Consider the BLS market introduced in Section 5.2, where besides the money market account $G(t)$, the financial assets traded are a bond, a longevity bond and a stock with price process $A(t) = (B(t,T), L(t,T), S(t))'$ (see (5.4)-(5.98)).

As in Section 5.2.1 we denote the agent's relative portfolio weights at time t by

$$(u^G(t), u^B(t), u^L(t), u^S(t)) = \big(u^G(t), u^A(t)\big),$$

for the riskless asset $G(t)$ and the risk assets $A(t)$, while her/his consumption rate at time t is denoted by $C(t)$. Depending upon the situation at hand, it may be convenient to introduce the relative consumption rate $c(t)$, i.e.,

$$C(t) = c(t)V(t), \quad \forall t \geq t_0, \tag{5.21}$$

for a suitable process $c(t)$.

Now let us assume the consumer-investor's preferences can be represented through a utility function on the consumption, $U(C)$. The utility function $U : (0,\infty) \to (0,\infty)$ is taken to be twice differentiable on $(0,\infty)$, strictly increasing and concave in its argument, and the concavity represents an investor who is risk averse. This means that she/he is willing to pay in order to avoid a risk. Furthermore the function $U(C)$ is assumed to satisfy one of the standard assumptions of economic growth theory, Inada conditions[1]. Thus we have that $U(C)$ is such that

$$\dot{U}(C) > 0, \quad \ddot{U}(C) < 0 \text{ and } \lim_{C\to 0} \dot{U}(C) = \infty, \quad \lim_{C\to+\infty} \dot{U}(C) = 0. \tag{5.22}$$

The objective of the agent is to choose a portfolio-consumption strategy $(u^A(t), c(t))$ in such a way as to maximize her/his expected utility over $[t_0, \tau)$ (see Section 5.2.1), where τ is a stochastic time horizon coinciding with her/his death time. According to Sections 3.2 and 3.4, and definition (3.33), let τ be the first jump time of the process $N(t) = \hat{N}\big(\int_0^t \lambda(s)ds\big)$, where $\hat{N}(t)$ is a standard Poisson process, independent of $\mathcal{G} = \mathcal{G}_\infty$. The objective of the agent is to choose a portfolio-consumption strategy $(u^A(t), c(t))$ based on the information available until time t, for $t < \tau$, i.e., until the agent is still alive. The information available is represented by $\mathcal{H}_t$, and we have assumed in (5.152) that $\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_t^N$. Therefore it is natural[2] to consider strategies $(u^A(t), c(t))$ that are $\mathbb{G}$-adapted and virtually defined for all times t, i.e., without the restriction $t < \tau$. Finally we assume that the agent's expected utility is[3]

$$E^P\left[\int_{t_0}^{\infty} 1_{\{\tau>t\}} e^{-\rho t} U(c(t)V(t))dt \Big| \mathcal{H}_{t_0}\right], \tag{5.23}$$

where the constant parameter ρ, that measures the subjective discount factor[4], is assumed to be strictly positive ($\rho \in (0,\infty)$).

Summarizing, the aim is to maximize, over a suitable set of admissible strategies, the functional (5.23), i.e.,

$$\sup_{c(\cdot),u^A(\cdot)\in Adm} E^P\left[\int_{t_0}^{\infty} 1_{\{\tau>t\}} e^{-\rho t} U\big(c(t)V(t)\big)dt \Big| \mathcal{H}_{t_0}\right], \tag{5.24}$$

[1] We say that a strictly concave increasing function $f : (0,\infty) \to (0,\infty)$ that is differentiable on $(0,\infty)$ satisfies Inada conditions, named after the economist Ken-Ichi Inada, if

(i) the limit of the derivative towards 0 is positive infinity,

(ii) the limit of the derivative towards positive infinity is 0.

[2] Observe that until $t < \tau$ the information coming from $\mathcal{F}_t^N$ is simply given by the event $\{t < \tau\}$.

[3] Equivalently, we could consider

$$E^P\left[\int_{t_0}^{\tau} e^{-\rho t} U(c(t)V(t))dt \Big| \mathcal{H}_{t_0}\right],$$

but then we should add the condition on $\tau > t_0$, while the latter condition is not necessary if we write the agent's expected utility as in (5.23).
Furthermore, since we have assumed that τ is bounded above by T, (see condition (5.1)) it would be natural to write

$$E^P\left[\int_{t_0}^{T} 1_{\{\tau>t\}} e^{-\rho t} U(c(t)V(t))dt \Big| \mathcal{H}_{t_0}\right]$$

instead of (5.23). Nevertheless we prefer the formulation (5.23) since it is possible to consider also the case of random times τ that do not satisfy condition (5.1), as we will do in the subsequent Section 5.5.

[4] The positive subjective discount factor $\rho \geq 0$ means that the consumer takes less satisfaction from delayed consumption. In some economic models the subjective discount factor may be chosen negative, meaning that future consumption is evaluated more than present one.

where Adm has to be specified.

A natural constraint is that the consumption rate $C(t)$ is a positive process, and it may be reasonable to require the consumer's wealth $V(t)$ never becomes negative or null, so that a constraint on the relative consumption rate is given by

$$c(t) \geq 0, \quad \forall t \geq t_0. \tag{5.25}$$

Furthermore to avoid arbitrage the admissible strategies are subject to the so-called budget constraint

$$V = V(t_0) = E^Q\left[\int_{t_0}^{T'\wedge\tau} \frac{c(t)V(t)}{G(t)}\,dt + \frac{V(T\wedge\tau)}{G(T\wedge\tau)}\Big|\mathcal{H}_{t_0}\right], \quad \forall\, T' > 0 \tag{5.26}$$

where the measure Q is defined on $\mathcal{H}_\infty$ as the unique measure such that, for each $T' > 0$, $Q_{T'} = Q|_{\mathcal{H}_{T'}}$ is the risk-neutral measure on $\mathcal{H}_{T'}$, i.e.,

$$\frac{dQ_{T'}}{dP_{T'}} = \exp\Big\{-\frac{1}{2}\int_0^{T'} |\xi(s)|^2\,ds + \int_0^{T'} \xi(s)\,dW_s\Big\}, \tag{5.27}$$

with $P_{T'} = P|_{\mathcal{H}_{T'}}$. Since the consumption rate is given by $c(t)V(t)$,

$$\int_{t_0}^{T'\wedge\tau} \frac{c(t)V(t)}{G(t)}\,dt$$

represents the discounted[5] consumption up to time $T'\wedge\tau$, while

$$\frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}$$

represents the discounted wealth at time $T'\wedge\tau$. Letting T' go to ∞, the above quantities converge[6] to the total discounted consumption $\int_{t_0}^{\tau} \frac{c(t)V(t)}{G(t)}\,dt$ and the discounted heredity $\frac{V(\tau)}{G(\tau)}$. Then the budget constraint has the satisfying interpretation that the sum of the expected total discounted consumption plus the expected discounted heredity equals the initial endowment V and each of them cannot exceed V. In other words, the agent does not want to leave debts to her/his heirs.

Since we restrict the consumer's investment-consumption strategies to be self-financing, the dynamics of the corresponding value process are given by (5.16) with $V(t_0) = V$. Now inserting (5.21) into (5.16) we obtain

$$\begin{cases} dV(t) = V(t)\Big[\Big(r(t) + u^A(t)\big(\mu^A(t) - r(t)\mathbf{1} - \lambda(t)\mathbf{1}_\lambda\big) - c(t)\Big)dt + u^A(t)\Sigma^A(t)dW(t)\Big] \\ V(t_0) = V, \end{cases} \tag{5.28}$$

so that we have an explicit solution given by

$$\begin{aligned} V(t) =& V\exp\left\{\int_{t_0}^{t}\Big(r(s) + u^A(s)\big(\mu^A(s) - r(s)\mathbf{1} - \lambda(s)\mathbf{1}_\lambda\big) - c(s) - \frac{1}{2}\left|u^A(t)\Sigma^A(t)\right|^2\Big)ds \right.\\ &\left. + \int_{t_0}^{t} u^A(s)\Sigma^A(s)dW(s)\right\}. \end{aligned} \tag{5.29}$$

The above expression will be essential to prove that the above optimal control problem with budget constraint is equivalent to a problem with simpler constraints, as shown in the following theorem.

Theorem 5.3.1. *Let $\mathcal{P}^\tau$ be the optimization problem (5.24) over the set of admissible strategies, with constraints (5.25) and (5.26). Define $\mathcal{Z}^A(t)$ as follows*

$$\mathcal{Z}^A(t) = \exp\Big\{-\frac{1}{2}\int_0^{t} |\xi^A(s)|^2\,ds - \int_0^{t} \xi^A(s)\,dW_s\Big\}, \tag{5.30}$$

where

$$\xi^A(s) := \xi(s) - u^A(s)\Sigma^A(s). \tag{5.31}$$

[5] In this setting the discounting is accomplished by the price process G (see Section 1.1).

[6] In the case $\tau \leq T$ with probability 1, then the above quantities are equal to $\int_{t_0}^{\tau} \frac{c(t)V(t)}{G(t)}\,dt$ and $\frac{V(\tau)}{G(\tau)}$, for all $T' \geq T$.

Then $\mathcal{P}^\tau$ is equivalent to the following problem[7]

$$1_{\{\tau>t_0\}} \sup_{c(\cdot),u^A(\cdot)\in Adm} E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big)dt \Big| \mathcal{G}_{t_0}\right], \tag{5.32}$$

where[8]

$$Adm = \{c(\cdot), u^A(\cdot) : c \text{ and } u^A \text{ are } \mathbb{G}\text{-adapted, } c \geq 0 \text{ and } u^A \text{ such that (5.30) is a } \mathbb{G}\text{-martingale}\}.$$

Before proceeding with the proof, we need the following preliminary results.

Lemma 5.3.2. *For any $\mathbb{G}$-adapted consumption-investment strategy (c, u^A) such that (5.25) hold, we have that*

$$E^P\left[\int_{t_0}^{\infty} 1_{\{\tau>t\}} e^{-\rho t} U(c(t)V(t))dt \Big| \mathcal{H}_{t_0}\right] = 1_{\{\tau>t_0\}} E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big)dt \Big| \mathcal{G}_{t_0}\right], \tag{5.33}$$

Proof of Lemma 5.3.2. To show (5.33), we start by recalling that, besides being a stochastic intensity for the doubly stochastic process $N(t)$ with respect to the filtration $\mathcal{H}_t = \mathcal{F}_t^N \vee \mathcal{G}_t$, the process $\lambda(t)$ is a stochastic intensity for $N(t)$ also with respect to the larger filtration $\mathcal{F}_t^N \vee \mathcal{G}$. Then, by the iterated conditional expectations property and (3.40) with $T_1 = \tau$, the conditional expectation in (5.23) can be written as

$$\begin{aligned}
&1_{\{\tau>t_0\}} E^P\left[\int_{t_0}^{\infty} 1_{\{\tau>t\}} e^{-\rho t} U\big(c(t)V(t)\big)dt \Big| \mathcal{H}_{t_0}\right] \\
&= 1_{\{\tau>t_0\}} \int_{t_0}^{\infty} E^P\left[1_{\{\tau>t\}} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{H}_{t_0}\right] dt \\
&= 1_{\{\tau>t_0\}} \int_{t_0}^{\infty} E^P\left[E^P\left[1_{\{\tau>t\}} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{F}_{t_0}^N \vee \mathcal{G}\right] \Big| \mathcal{H}_{t_0}\right] dt \\
&= 1_{\{\tau>t_0\}} \int_{t_0}^{\infty} E^P\left[E^P\left[1_{\{\tau>t\}} \Big| \mathcal{F}_{t_0}^N \vee \mathcal{G}\right] e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{H}_{t_0}\right] dt \\
&= 1_{\{\tau>t_0\}} \int_{t_0}^{\infty} E^P\left[P\left(\tau > t \Big| \mathcal{F}_{t_0}^N \vee \mathcal{G}\right) e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{H}_{t_0}\right] dt \\
&= 1_{\{\tau>t_0\}} \int_{t_0}^{\infty} E^P\left[e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{H}_{t_0}\right] dt
\end{aligned} \tag{5.34}$$

where we have used Lemma A.1.2 of Appendix A, the inclusion $\mathcal{H}_{t_0} \subseteq \mathcal{F}_{t_0}^N \vee \mathcal{G}$, and, for $t > t_0$, $\mathcal{F}_{t_0}^N \vee \mathcal{G}$ measurability of $U\big(c(t)V(t)\big)$ (the latter property follows since $c(t)$ and $u^A(t)$ are $\mathbb{G}$-adapted). To obtain the announced result, we finally observe that $\mathcal{H}_{t_0} \subseteq \mathcal{G}_{t_0} \vee \mathcal{F}_\infty^{\hat{N}}$, $\mathcal{G}_{t_0} \subseteq \mathcal{G}_t \subseteq \mathcal{G}$, the sigma-algebras $\mathcal{G}$ and $\mathcal{F}_\infty^{\hat{N}}$ are independent (under P), the random variables $e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big)$ are $\mathcal{G}_t$-measurable[9] (and therefore independent of $\mathcal{F}_\infty^{\hat{N}}$). Then the redundant conditioning property implies that

$$\begin{aligned}
&\int_{t_0}^{\infty} E^P\left[e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{H}_{t_0}\right] dt \\
&= \int_{t_0}^{\infty} E^P\left[E^P\left[e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{F}_\infty^{\hat{N}} \vee \mathcal{G}_{t_0}\right] \Big| \mathcal{H}_{t_0}\right] dt \\
&= \int_{t_0}^{\infty} E^P\left[E^P\left[e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{G}_{t_0}\right] \Big| \mathcal{H}_{t_0}\right] dt \\
&= \int_{t_0}^{\infty} E^P\left[e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) \Big| \mathcal{G}_{t_0}\right] dt \\
&= E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big) dt \Big| \mathcal{G}_{t_0}\right].
\end{aligned}$$

□

[7] Since we have assumed that τ is bounded above by T, (see condition (5.1)) it would be natural to write

$$1_{\{\tau>t_0\}} \sup_{c(\cdot),u^A(\cdot)\in Adm} E^P\left[\int_{t_0}^{T} e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big)dt \Big| \mathcal{G}_{t_0}\right]$$

instead of (5.32). Nevertheless we prefer the formulation (5.32) since it is possible to consider also the case of random times τ that do not satisfy condition (5.1), as we will do in the subsequent Section 5.5.

[8] We recall that the other constraint (5.13) is automatically satisfied when $u^G(t)$ is given by (5.15).

[9] Since we consider only $\mathbb{G}$-adapted strategies $(c(t), u^A(t))$, also $V(t)$ is a $\mathbb{G}$-adapted process and the random variables $e^{-\int_{t_0}^t \lambda(u)du} e^{-\rho t} U\big(c(t)V(t)\big)$ are $\mathcal{G}_t$-measurable, which is essential to apply Lemma A.1.2.

Proposition 5.3.3. *Let u^A be a $\mathbb{H}$-adapted strategy. Assume that*

$$\mathcal{Z}^A(t) = \exp\Big\{ -\frac{1}{2}\int_0^t |\xi^A(s)|^2\,ds - \int_0^t \xi^A(s)\,dW_s\Big\}, \tag{5.35}$$

where

$$\xi^A(s) := \xi(s) - u^A(s)\Sigma^A(s), \tag{5.36}$$

is a $\mathbb{H}$-martingale under the measure P. Then for each positive valued $\mathbb{H}$-adapted strategy c the constraint (5.26) holds, i.e.,

$$V = V(t_0) = E^Q\left[\int_{t_0}^{T'\wedge\tau} \frac{c(t)V(t)}{G(t)}\,dt + \frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}\Big|\mathcal{H}_{t_0}\right], \quad \forall\, T' > 0,$$

where the measure Q is given by (5.27).

Before giving the proof of the above proposition we note that the definitions (5.30) and (5.31) differ from the definitions (5.35) and (5.36) only in that the control $u^A(t)$ is $\mathbb{G}$-adapted.

Proof of Proposition 5.3.3. Since the measure Q is given by (5.27), the constraint (5.26) becomes

$$V = E^P\left[\int_0^{T'} \frac{dQ_t}{dP_t}\mathbf{1}_{\{\tau>t\}}\frac{c(t)V(t)}{G(t)}\,dt + \frac{dQ_{T'\wedge\tau}}{dP_{T'\wedge\tau}}\frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}\Big|\mathcal{H}_{t_0}\right] \tag{5.37}$$

Letting $c(t)$ be a positive strategy and taking into account $G(t)$, $V(t)$ and $\frac{dQ_t}{dP_t}$ given by (1.3), (5.28) and (5.27), we obtain

$$\begin{aligned}
\log\left(\frac{dQ_t}{dP_t}\frac{V(t)}{V\,G(t)}\right) &= -\frac{1}{2}\int_0^t |\xi(s)|^2\,ds - \int_0^t \xi(s)\,dW_s \\
&\quad + \int_0^t \Big(r(s) - c(s) + u^A(s)\big(\mu^A(s) - r(s)\mathbf{1} - \lambda(s)\mathbf{1}_\lambda\big)\Big)ds \\
&\quad - \int_0^t \frac{1}{2}|u^A(s)\Sigma^A(s)|^2\,ds + \int_0^t u^A(s)\Sigma^A(s)dW(s) - \int_0^t r(s)\,ds,
\end{aligned}$$

so that recalling that

$$\mu^A(s) - r(s)\mathbf{1} - \lambda(s)\mathbf{1}_\lambda = \Sigma^A(s)\xi(s)$$

we obtain

$$\begin{aligned}
\log\left(\frac{dQ_t}{dP_t}\frac{V(t)}{V_0\,G(t)}\right) &= -\int_0^t c(s)\,ds - \frac{1}{2}\int_0^t |\xi(s)|^2\,ds - \int_0^t \xi(s)\,dW_s + \int_0^t u^A(s)\,\Sigma^A(s)\xi(s)ds \\
&\quad - \int_0^t \frac{1}{2}|u^A(s)\Sigma^A(s)|^2\,ds + \int_0^t u^A(s)\Sigma^A(s)dW(s) \\
&= -\int_0^t c(s)\,ds - \int_0^t \frac{1}{2}|\xi(s) - u^A(s)\Sigma^A(s)|^2\,ds - \int_0^t \big(\xi(s) - u^A(s)\Sigma^A(s)\big)\,dW(s) \\
&= -\int_0^t c(s)\,ds - \frac{1}{2}\int_0^t |\xi^A(s)|^2\,ds - \int_0^t \xi^A(s)\,dW_s.
\end{aligned}$$

Thanks to the assumption that (5.35) is a martingale, we can define Q^A as the unique measure such that, for each $t > 0$, $Q^A|_{\mathcal{H}_t} := Q_t^A$, where

$$\frac{dQ_t^A}{dP_t} = \exp\Big\{ -\frac{1}{2}\int_0^t |\xi^A(s)|^2\,ds - \int_0^t \xi^A(s)\,dW_s\Big\}. \tag{5.38}$$

Then we obtain that

$$\begin{aligned}
E^P\left[\int_{t_0}^{T'} \frac{dQ_t}{dP_t}\mathbf{1}_{\{\tau>t\}}\frac{c(t)V(t)}{G(t)}\,dt\Big|\mathcal{H}_{t_0}\right] &= E^P\left[\int_0^{T'} \frac{dQ_t^A}{dP_t}\mathbf{1}_{\{\tau>t\}}V\,c(t)\,e^{-\int_0^t c(s)\,ds}\,dt\Big|\mathcal{H}_{t_0}\right] \\
&= V\,E^{Q_{T'}^A}\left[\int_0^{T'} \mathbf{1}_{\{\tau>t\}}\,c(t)\,e^{-\int_0^t c(s)\,ds}\,dt\Big|\mathcal{H}_{t_0}\right] \\
&= V\,E^{Q_{T'}^A}\left[\int_0^{T'\wedge\tau} c(t)\,e^{-\int_0^t c(s)\,ds}\,dt\Big|\mathcal{H}_{t_0}\right] \\
&= V\,E^{Q_{T'}^A}\left[1 - e^{-\int_0^{T'\wedge\tau} c(s)\,ds}\Big|\mathcal{H}_{t_0}\right],
\end{aligned} \tag{5.39}$$

and that

$$E^P\left[\frac{dQ_{T'\wedge\tau}}{dP_{T'\wedge\tau}}\frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}\Big|\mathcal{H}_{t_0}\right] = E^P\left[\frac{dQ^A_{T'\wedge\tau}}{dP_{T'\wedge\tau}}V\,e^{-\int_0^{T'\wedge\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right] = V\,E^{Q^A_{T'}}\left[e^{-\int_0^{T'\wedge\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right]. \tag{5.40}$$

Then by (5.39) and (5.40) we obtain the announced result.

□

Remark 5.3.1. *Letting T' go to ∞ in the constraint (5.26), we obtain also a convergence result, i.e.,*

$$V = \lim_{T'\to\infty} E^P\left[\int_0^{T'}\frac{dQ_t}{dP_t}\mathbf{1}_{\{\tau>t\}}\frac{c(t)V(t)}{G(t)}\,dt + \frac{dQ_{T'\wedge\tau}}{dP_{T'\wedge\tau}}\frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}\Big|\mathcal{H}_{t_0}\right]. \tag{5.41}$$

Indeed, by (5.39) and (5.40) in the proof of Proposition 5.3.3, letting T' go to ∞, we obtain that

$$\begin{aligned} E^P\left[\int_{t_0}^{T'}\frac{dQ_t}{dP_t}\mathbf{1}_{\{\tau>t\}}\frac{c(t)V(t)}{G(t)}\,dt\Big|\mathcal{H}_{t_0}\right] &= V\,E^{Q^A_{T'}}\left[1-e^{-\int_0^{T'\wedge\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right] \\ &\underset{T'\to\infty}{\longrightarrow} V\,E^{Q^A}\left[1-e^{-\int_0^{\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right], \end{aligned}$$

and that

$$\begin{aligned} E^P\left[\frac{dQ_{T'\wedge\tau}}{dP_{T'\wedge\tau}}\frac{V(T'\wedge\tau)}{G(T'\wedge\tau)}\Big|\mathcal{H}_{t_0}\right] &= V\,E^{Q^A_{T'}}\left[e^{-\int_0^{T'\wedge\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right] \\ &\underset{T'\to\infty}{\longrightarrow} V\,E^{Q^A}\left[e^{-\int_{t_0}^{\tau}c(s)\,ds}\Big|\mathcal{H}_{t_0}\right], \end{aligned}$$

respectively.

Proof of Theorem 5.3.1. The proof follows by Lemma 5.3.2 and Proposition 5.3.3 with $\mathbb{G}$-adapted consumption-investment strategy (c, u^A). Indeed, by Lemma 5.3.2 we have that for any $\mathbb{G}$-adapted consumption-investment strategy (c, u^A) (and without requiring that (5.30) is a martingale) the agent's expected utility (5.23) can be written in terms of the conditional expectation with respect to $\mathcal{G}_{t_0}$ instead of $\mathcal{H}_{t_0}$.

Furthermore, by Proposition 5.3.3, we can replace the budget constraint with the requirement that the strategy u^A has the property that (5.30) is a $\mathbb{H}$-martingale. The process (5.30) being $\mathbb{G}$-adapted, and the σ-algebras $\mathcal{F}^{\hat{N}}_\infty$ and $\mathcal{G}_\infty$ being independent under P, the property that $\mathcal{G}_t \subseteq \mathcal{H}_t \subseteq \mathcal{G}_t \vee \mathcal{F}^{\hat{N}}_\infty$ implies that the process (5.30) is a $\mathbb{G}$-martingale if and only if is a $\mathbb{H}$-martingale. □

Finally, in order to solve the problem (5.24) with the budget constraint (5.26), we proceed as follows. We start by solving the problem (5.32) disregarding the budget constraint, and find the optimal (unconstrained) controls. Then the (uncostrained) optimal controls are also optimal for the problem with the budget constraint, if the corresponding process (5.30) is a martingale. Next section is devoted to the unconstrained optimization problem.

5.3.1 Optimal Markov control problem without the budged constraint

In this section the aim is focused on the problem (5.32) without considering $1_{\{\tau>t_0\}}$, i.e.,

$$\sup_{c(\cdot),u^A(\cdot)\in Adm} E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^{t}\lambda(u)du}e^{-\rho t}U\big(c(t)V(t)\big)dt\Big|\mathcal{G}_{t_0}\right]. \tag{5.42}$$

Since in most concrete case it is natural to require that the control processes are adapted to the state processes, the purpose of this section is to study an optimal Markov control problem. The idea is that if we restrict to admissible consumption-portfolio strategies $(c(t), u^A(t))$ which are deterministic functions of $(z(t), S(t), V(t))'$, then the latter process is Markovian with respect to $\mathcal{G}_t$, so that

$$\begin{aligned} &E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^{t}\lambda(u)du}e^{-\rho t}U\big(c(t)V(t)\big)dt\Big|\mathcal{G}_{t_0}\right] \\ &\quad= E^P\left[\int_{t_0}^{\infty} e^{-\int_{t_0}^{t}\lambda(u)du}e^{-\rho t}U\big(c(t)V(t)\big)dt\Big|z(t_0), S(t_0), V(t_0)\right]. \end{aligned}$$

These consumption-portfolio strategies can be considered as Markov control policies and, denoting by $\mathfrak{U}_{ad}$ the set of such strategies, instead of the problem (5.42), we consider the optimal control problem

$$\begin{aligned} & J(t_0, z, S, V) \\ & = \sup_{c(\cdot), u^A(\cdot) \in \mathfrak{U}_{ad}} E^P \left[\int_{t_0}^{\infty} e^{-\int_{t_0}^{t} \lambda(u) du} e^{-\rho t} U\big(c(t)V(t)\big) dt \Big| z(t_0) = z, S(t_0) = S, V(t_0) = V \right]. \end{aligned} \tag{5.43}$$

Due to the integrand $e^{-\int_{t_0}^{t} \lambda(u)du}$ the above control problem cannot be considered in the framework of optimal Markov control problems.
Nevertheless, if we set $z_0(t) = z_0 e^{-\int_{t_0}^{t} \lambda(u)du}$, or equivalently

$$dz_0(t) = -\lambda(t) z_0(t) dt, \tag{5.44}$$

and

$$\mathcal{G}_t^0 = \mathcal{F}_t^{z_0} \vee \mathcal{G}_t,$$

then the above problem can be formulated in a optimal Markov control problem setting.

In this section we consider control processes that, besides depending on $z(t)$, $S(t)$, and $V(t)$, may depend also on $z_0(t)$, i.e.,

$$u^A(t) = \bar{u}^A(t, z_0(t), z(t), S(t), V(t)), \tag{5.45}$$
$$c(t) = \bar{c}(t, z_0(t), z(t), S(t), V(t)), \tag{5.46}$$

for some measurable deterministic function $\bar{u}^A$, and $\bar{c}$, where $\bar{u}^A, \bar{c} \in \bar{\mathcal{U}}_{ad}$, with value in $U_{ad} = \mathbb{R}^3 \times [0, \infty)$. We will denote the class of admissible Markov control processes as

$$\bar{\mathfrak{U}}_{ad} = \{u^A(\cdot), c(\cdot) : \ u^A(t) = \bar{u}^A(t, z_0(t), z(t), S(t), V(t)), \ c(t) = \bar{c}(t, z_0(t), z(t), S(t), V(t))\}.$$

Note that if $(c(t), u^A(t)) \in \bar{\mathfrak{U}}_{ad}$, then (**i**) the constraint on consumption

$$c(t) \geq 0. \tag{5.47}$$

is automatically satisfied; (**ii**) there is no constraint on the portfolio $u^A(t)$, except Markovianity. When considering such strategies, $(z_0(t), z(t), S(t), V(t))'$ is a Markovian process with respect to $\mathcal{G}_t^0$, as shown in the following proposition.

Proposition 5.3.4. *Under Assumption 5.2.2, for all strategies in $\bar{\mathfrak{U}}_{ad}$ the process $(z_0(t), z(t), S(t), V(t))'$, defined in (5.44), (5.4), (5.28), and (5.52), respectively, is Markovian with respect to $\mathcal{G}_t^0$.*

Proof. Observe that $(z_0(t), z(t), S(t))'$ is a Markovian process with respect to $\mathcal{F}_t^{z_0} \vee \mathcal{F}_t^z \vee \mathcal{F}_t^S$, while $V(t)$ is not a Markovian process with respect to $\mathcal{F}_t^V$, since in (5.28), for example, we have the dependence on the process $z(t)$. Since (5.45) and (5.46) hold, the process $(z_0(t), z(t), S(t), V(t))$ is a Markovian process with respect to $\mathcal{F}_t^{z_0} \vee \mathcal{F}_t^z \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V$. Due to the independence of $\mathcal{F}_t^{z_0} \vee \mathcal{F}_t^z \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V$ and $\mathcal{F}_\infty^{\hat{N}}$ the process $(z_0(t), z(t), S(t), V(t))$ is a Markovian process also with respect to $\mathcal{F}_t^{z_0} \vee \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^S \vee \mathcal{F}_\infty^{\hat{N}}$. On the one hand, by (3.34) we have that

$$\mathcal{G}_t^0 \subseteq \mathcal{F}_t^{z_0} \vee \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V \vee \mathcal{F}_t^N \subseteq \mathcal{F}_t^{z_0} \vee \mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^S \vee \mathcal{F}_\infty^{\hat{N}}, \quad \forall t, \tag{5.48}$$

on the other hand $\mathcal{F}_t^{z_0} \vee \mathcal{F}_t^z \vee \mathcal{F}_t^S \vee \mathcal{F}_t^V \subseteq \mathcal{G}_t^0$, for all t, and, by Lemma A.1.3, the result follows. □

Now we consider the following optimal Markov control problem

$$\bar{J}(t_0, z_0, z, S, V) = \sup_{u^A, c \in \bar{\mathfrak{U}}_{ad}} E^P \left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t) U\big(c(t)V(t)\big) dt \Big| z_0, z, S, V \right], \tag{5.49}$$

with dynamics of state processes given by

$$dz_0(t) = -\lambda(t) z_0(t) dt \tag{5.50}$$
$$dz(t) = \mu^z(t) dt + \Pi(t) dW(t) \tag{5.51}$$
$$dS(t) = S(t) \left[(r(t) + \sigma_r^S(t)\xi_r(t) + \sigma_S^S(t)\xi_S(t)) dt + \sigma^S(t) dW(t) \right], \tag{5.52}$$
$$dV(t) = V(t) \left[(r(t) + u^A(t) (\mu^A(t) - r(t)\mathbf{1} - \lambda(t)\mathbf{1}_\lambda) - c(t)) dt + u^A(t)\Sigma^A(t) dW(t) \right], \tag{5.53}$$

where $\Sigma^A(t)$ and $\mu^A(t)$ are given by (5.10) and (5.11) respectively, and $\sigma^S(t) = \big(\sigma^S_r(t), 0, \sigma^S_S(t)\big)$ is given by

$$\sigma^S(t) = \tilde{\sigma}^S(t, r(t), S(t)) = \Big(\hat{\sigma}^S_r(t, r(t), S(t)),\, 0\,,\, \hat{\sigma}^S_S(t, r(t), S(t))\Big), \tag{5.54}$$

with initial conditions

$$z_0(t_0) = z_0, \quad z(t_0) = z, \quad S(t_0) = S, \quad V(t_0) = V, \tag{5.55}$$

under the constraint that the strategies $(c(t), u^A(t)) \in \bar{\mathfrak{U}}_{ad}$.

In the sequel we denote the above optimal Markov control problem as $(\bar{\mathbf{P}})$.

Finally, we turn to the relations between the problem (5.43), with value function $J(t_0, z, S, V)$, and problem $(\bar{\mathbf{P}})$. As we will see in the sequel (see (5.91)) the value function for the problem (5.43) is given by

$$J(t_0, z, S, V) = \bar{J}(t_0, 1, z, S, V), \tag{5.56}$$

and, if there exists an optimal control policy

$$\begin{aligned} &\bar{c}_{sup}\big(t, z_0(t), r(t), \lambda(t), S(t), V(t)\big),\\ &\bar{u}^A_{sup}\big(t, z_0(t), r(t), \lambda(t), S(t), V(t)\big), \end{aligned}$$

for $(\bar{\mathbf{P}})$, then the corresponding optimal control $(c_{sup}(t), u^A_{sup}(t))$ is given by

$$\begin{aligned} c_{sup}(t) &= \bar{c}_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t)\big),\\ u^A_{sup}(t) &= \bar{u}^A_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t)\big), \end{aligned}$$

provided that the process

$$\mathcal{Z}^A_{sup}(t) = \exp\Big\{ -\frac{1}{2}\int_0^t |\xi^A_{sup}(s)|^2\, ds + \int_0^t \xi^A_{sup}(s)\, dW_s\Big\}, \tag{5.57}$$

where

$$\xi^A_{sup}(s) := u^A_{sup}(s)\Sigma^A(s) + \xi(s), \tag{5.58}$$

is a $\mathbb{G}$-martingale under the measure P.

The rest of this section is devoted to problem $(\bar{\mathbf{P}})$.

In the sequel, for the notational convenience, we use the following notation

$$Z(t) = \begin{pmatrix} z_0(t)\\ z_1(t)\\ z_2(t)\\ z_3(t)\\ z_4(t) \end{pmatrix} = \begin{pmatrix} z_0(t)\\ r(t)\\ \lambda(t)\\ S(t)\\ V(t) \end{pmatrix}, \tag{5.59}$$

with the dynamics given by

$$\begin{cases} dZ(t) = \mu(t)dt + \Gamma(t)dW(t),\\ Z(t_0) = Z, \end{cases} \tag{5.60}$$

where $Z = (z_0,\, r,\, \lambda,\, V,\, S)'$, and

$$\mu(t) = \hat{\mu}(t, u^A(t), c(t), Z(t)) \tag{5.61}$$

$$\Gamma(t) = \hat{\Gamma}(t, u^A(t), c(t), Z(t)), \tag{5.62}$$

with, as already recalled, $u^A(t)$ and $c(t)$ given by (5.45) and (5.46), and $\hat{\mu}(t, u^A, c, Z)$ and $\hat{\Gamma}(t, u^A, c, Z)$ given by

$$\hat{\mu} = \begin{pmatrix} -\lambda z_0 \\ \hat{\mu}^r \\ \hat{\mu}^\lambda \\ S\left(\hat{\sigma}_r^S \hat{\xi}_r + \hat{\sigma}_S^S \hat{\xi}_S\right) \\ V\left(r + u^A\left(\tilde{\mu}^A - r\mathbf{1} - \lambda \mathbf{1}_\lambda\right) - c\right) \end{pmatrix}, \tag{5.63}$$

$$\hat{\Gamma} = \begin{pmatrix} 0 & 0 & 0 \\ \hat{\sigma}^r & 0 & 0 \\ 0 & \hat{\sigma}^\lambda & 0 \\ S\hat{\sigma}_r^S & 0 & S\hat{\sigma}_S^S \\ V\left(u^B \hat{\sigma}^r \frac{\hat{B}_r^T}{\hat{B}^T} + u^L \hat{\sigma}^r \frac{\hat{L}_r^T}{\hat{L}^T} + u^S \hat{\sigma}_r^S\right) & V u^L \hat{\sigma}^\lambda \frac{\hat{L}_\lambda^T}{\hat{L}^T} & V u^S \hat{\sigma}_S^{r,S} \end{pmatrix}. \tag{5.64}$$

Denoting the conditional expectation given $Z(t_0) = Z$ as $E^P_{t_0,Z}$, the value function for $(\bar{\mathbf{P}})$ can be rewritten as

$$\bar{J}(t_0, Z) = \sup_{u^A(\cdot), c(\cdot) \in \bar{\mathfrak{U}}_{ad}} E^P_{t_0,Z}\left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t) U\big(c(t)V(t)\big)\, dt\right]. \tag{5.65}$$

By using the stochastic dynamic programming technique (see Appendix D.1), and recalling that $c(t)$ and $u^A(t)$ are given by (5.45) and (5.46), the corresponding Hamilton-Jacobi-Bellman equation for the problem is given by[10]

$$\frac{\partial J}{\partial t_0}(t_0, Z) + \sup_{u^A, c \in U_{ad}} \left\{ e^{-\rho t_0} z_0 U\big(Vc\big) + \mathcal{A}^{u^A, c} J(t_0, Z) \right\} = 0, \tag{5.66}$$

where $U_{ad} = \mathbb{R}^3 \times [0, \infty)$, and

$$\begin{aligned}
\mathcal{A}^{u^A,c} J(t_0, Z) &= \sum_{i=0}^{4} \hat{\mu}_i(t_0, u^A, c, Z) \frac{\partial J}{\partial z_i}(t_0, Z) + \frac{1}{2} \sum_{i,j=1}^{4} \left[\left(\hat{\Gamma}\hat{\Gamma}'\right)(t_0, u^A, c, Z)\right]_{i,j} \frac{\partial^2}{\partial z_i \partial z_j}(t_0, Z), \\
F(t_0, u^A, c, Z) &= e^{-\rho t_0} z_0 U\big(Vc\big).
\end{aligned}$$

Following the scheme at the end of Appendix D.1 (see points $1-4$), now we consider point 1. In the sequel for the notational convenience[11] we denote t_0 by t.

Fixed an arbitrary point $(t, Z) \in (0, \infty) \times (0, 1] \times (0, \infty)^4$ and any function $J(t, Z)$ sufficiently smooth, we now have to solve the optimization problem

$$\sup_{u^A, c \in U_{ad}} \left\{ e^{-\rho t} z_0 U\big(Vc\big) + \mathcal{A}^{u^A, c} J(t, Z) \right\}. \tag{5.67}$$

We remember that u^A and c are the only variables, whereas t and Z are considered to be fixed parameters (see (D.17) at point 2 with $v = (u^A, c)$).

As usual we use the following notations

$$J_{z_0} = \frac{\partial J}{\partial z_0}, \quad J_r = \frac{\partial J}{\partial r}, \quad J_\lambda = \frac{\partial J}{\partial \lambda}, \quad J_S = \frac{\partial J}{\partial S}, \quad J_V = \frac{\partial J}{\partial V} \tag{5.68}$$

$$J_{rr} = \frac{\partial^2 J}{\partial r^2},\ J_{\lambda\lambda} = \frac{\partial^2 J}{\partial \lambda^2},\ J_{r\lambda} = \frac{\partial J}{\partial r \partial \lambda},\ J_{rS} = \frac{\partial J}{\partial r \partial S},\ J_{\lambda S} = \frac{\partial J}{\partial \lambda \partial S},\ J_{rV} = \frac{\partial J}{\partial r \partial V},\ J_{\lambda V} = \frac{\partial J}{\partial \lambda \partial V}, \tag{5.69}$$

$$J_{SS} = \frac{\partial^2 J}{\partial S^2},\ J_{VS} = \frac{\partial J}{\partial V \partial S},\ J_{VV} = \frac{\partial J}{\partial V \partial V}, \tag{5.70}$$

[10]Note that, with a little abuse of notation in the sequel we will denote by J a generic function of (t, Z). The reader should not be confused with the value function of the Markov optimal control problem (5.43) since the function J used here depends also on z_0, while the value function does not.

[11]Here we denote by J the function denoted by H in Appendix D.1.

and also

$$J_z = \begin{pmatrix} J_r \\ J_\lambda \end{pmatrix}, \quad J_{zV} = \begin{pmatrix} J_{rV} \\ J_{\lambda V} \end{pmatrix}, \quad J_{zz} = \begin{pmatrix} J_{rr} & J_{r\lambda} \\ J_{\lambda r} & J_{\lambda\lambda} \end{pmatrix} \quad J_{ZV} = \begin{pmatrix} J_{z_0 V} \\ J_{rV} \\ J_{\lambda V} \\ J_{SV} \\ J_{VV} \end{pmatrix}. \tag{5.71}$$

Then we have the following proposition.

Proposition 5.3.5. *Under the completeness Assumption 5.2.1, let $J(t,Z)$ be a regular function*[12]*, i.e., $J \in C^{1,2}$, and let $\bar{c}_{sup}(t,Z;J)$ and $\bar{u}^A_{sup}(t,Z;J)$ be the functions such that for each fixed choice of (t,Z) and any function $J \in C^{1,2}$ are the solutions of the optimization problem (5.67). Then*[13]

$$\bar{c}_{sup}(t,Z;J) = \frac{1}{V}\dot{U}^{-1}\left(\frac{J_V}{e^{-\rho t}z_0}\right), \tag{5.72}$$

$$\bar{u}^B_{sup}(t,Z;J) = \frac{1}{\frac{\hat{B}^T_r}{\hat{B}^T}}\frac{1}{V}\left(-\frac{J_V}{J_{VV}}\frac{\hat{\xi}_r\hat{\sigma}^S_S - \hat{\xi}_S\hat{\sigma}^S_r}{\hat{\sigma}^r\,\hat{\sigma}^S_S} - \frac{J_{rV}}{J_{VV}}\right) - \frac{\frac{\hat{L}^T_r}{\hat{L}^T}}{\frac{\hat{B}^T_r}{\hat{B}^T}}\frac{1}{V}\left(-\frac{J_V}{J_{VV}}\frac{\hat{\xi}_\lambda}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}\hat{\sigma}^\lambda} - \frac{1}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}}\frac{J_{V\lambda}}{J_{VV}}\right) \tag{5.73}$$

$$\bar{u}^L_{sup}(t,Z;J) = \frac{1}{V}\left(-\frac{J_V}{J_{VV}}\frac{\hat{\xi}_\lambda}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}\hat{\sigma}^\lambda} - \frac{1}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}}\frac{J_{V\lambda}}{J_{VV}}\right) \tag{5.74}$$

$$\bar{u}^S_{sup}(t,Z;J) = \frac{1}{V}\left(-\frac{J_V}{J_{VV}}\frac{\hat{\xi}_S}{\hat{\sigma}^S_S} - S\frac{J_{VS}}{J_{VV}}\right), \tag{5.75}$$

where the argument (t,Z) "has been suppressed" for the notational convenience.

Theorem 5.3.6. *Let us consider the following HJB equation for $(\bar{\mathbf{P}})$*

$$\frac{\partial H}{\partial t}(t,Z) + e^{-\rho t}z_0 U\left(\dot{U}^{-1}\left(\frac{\frac{\partial H}{\partial V}(t,Z)}{e^{-\rho t}z_0}\right)\right) + \mathcal{L}^{\bar{u}^A_{sup}(\cdot;H),\bar{c}_{sup}(\cdot;H)}H(t,Z) = 0, \tag{5.76}$$

with the boundary condition given by

$$\lim_{t\to\infty} H(t,z_0,z,S,V) = 0, \tag{5.77}$$

and assume that (5.76) admits a unique classical solution J. Then J coincides with the value function $\bar{J}$ for problem $(\bar{\mathbf{P}})$ (see e.g. (5.65)) and the optimal controls are

$$c_{sup}(t) = \bar{c}_{sup}\big(t, z_0(t), r(t), \lambda(t), V(t), S(t); J\big), \tag{5.78}$$

$$u^B_{sup}(t) = \bar{u}^B_{sup}\big(t, z_0(t), r(t), \lambda(t), V(t), S(t); J\big), \tag{5.79}$$

$$u^L_{sup}(t) = \bar{u}^L_{sup}\big(t, z_0(t), r(t), \lambda(t), V(t), S(t); J\big), \tag{5.80}$$

$$u^S_{sup}(t) = \bar{u}^S_{sup}\big(t, z_0(t), r(t), \lambda(t), V(t), S(t); J\big), \tag{5.81}$$

where $\bar{c}_{sup}$, $\bar{u}^B_{sup}$, $\bar{u}^L_{sup}$, and $\bar{u}^S_{sup}$ are given by (5.72)-(5.75).

Proof of Proposition 5.3.5. We now have to solve the optimization problem (5.67). To this end we need to write the explicit form of $\mathcal{A}^{u^A,c}J(t,Z)$, i.e.,

$$\begin{aligned}
\mathcal{A}^{u^A,c}J =& -z_0\,\lambda\,J_{z_0} + \hat{\mu}^r J_r + \hat{\mu}^\lambda J_\lambda + S\left(\hat{\sigma}^S_r\hat{\xi}_r + \hat{\sigma}^S_S\hat{\xi}_S\right)J_S + V\Big(r + u^A\left(\hat{\mu}^A - r\mathbf{1} - \lambda\mathbf{1}_\lambda\right) - c\Big)J_V \\
&+ \frac{1}{2}\Big[(\hat{\sigma}^r)^2 J_{rr} + (\hat{\sigma}^\lambda)^2 J_{\lambda\lambda} + S^2\left|\tilde{\sigma}^S\right|^2 J_{SS} + V^2\left|u^A\tilde{\Sigma}^A\right|^2 J_{VV}\Big] + \hat{\sigma}^\lambda\big(u^L\hat{\sigma}^\lambda\hat{L}^T_\lambda\big)J_{\lambda V} \\
&+ V\hat{\sigma}^r\left(u^B\hat{\sigma}^r\frac{\hat{B}^T_r}{\hat{B}^T} + u^L\hat{\sigma}^r\frac{\hat{L}^T_r}{\hat{L}^T} + u^S\hat{\sigma}^S_r\right)J_{rV} + S\hat{\sigma}^S_r\hat{\sigma}^r J_{rS} + SVu^A\tilde{\Sigma}^A\tilde{\sigma}^{S,\prime}J_{VS},
\end{aligned} \tag{5.82}$$

[12] We recall that here J is not necessarily a value function.

[13] Observe that $\bar{c}_{sup}$ and $\bar{u}^A_{sup}$ are the corresponding function u^*_J of point 2 in the scheme at the end of Appendix D.1.

or in compact form with $z(t) = (r(t), \lambda(t))'$,

$$\begin{aligned}\mathcal{A}^{u^A,c}J &= -z_0\,\lambda\,J_{z_0} + \hat{\mu}^{z\prime}J_z + S\left(\hat{\sigma}^S_r\hat{\xi}_r + \hat{\sigma}^S_S\hat{\xi}_S\right)J_S + V\Big(r + u^A M - c\Big)J_V + \frac{1}{2}tr[\tilde{\Pi}\tilde{\Pi}'J_{zz}] \\ &+ Vu^A\tilde{\Sigma}^A\tilde{\Pi}'J_{zV} + \frac{1}{2}S^2\,|\tilde{\sigma}^S|^2\,J_{SS} + S\hat{\sigma}^S_r\hat{\sigma}^r J_{rS} + SVu^A\tilde{\Sigma}^A\tilde{\sigma}^{S\prime}J_{VS} + \frac{1}{2}V^2\left|u^A\tilde{\Sigma}^A\right|^2 J_{VV}.\end{aligned} \tag{5.83}$$

Let us decompose $\mathcal{A}^{u^A,c}J$ as follows

$$\mathcal{A}^{u^A,c}J = A_1^cJ + A_2^{u^A}J + A_3J,$$

where, setting $Q = \tilde{\Sigma}^A\tilde{\Sigma}^{A\prime}$, $M = \hat{\mu}^A - r\mathbf{1} - \lambda\mathbf{1}_\lambda$,

$$\begin{aligned}A_1^cJ &= -cVJ_V, \\ A_2^{u^A}J &= Vu^AMJ_V + Vu^A\tilde{\Sigma}^A\tilde{\Pi}'J_{zV} + VSu^A\tilde{\Sigma}^A\tilde{\sigma}^{S\prime}J_{VS} + \frac{1}{2}V^2u^AQu^{A\prime}J_{VV},\end{aligned}$$

and A_3J is implicitely defined. Note that A_3J does not depend neither on c nor on u^A. Then, solve the maximization problem (5.67), is equivalent to

$$\sup_{u^A,c\in U_{ad}}\left\{e^{-\rho t}z_0U\big(Vc\big) + A_1^cJ + A_2^{u^A}J\right\} + A_3J. \tag{5.84}$$

The first order condition on consumption is given by

$$e^{-\rho\,t}z_0V\dot{U}(cV) - VJ_V = 0, \tag{5.85}$$

thus (5.72) follows. Indeed we get $\bar{c}_{sup}$ by setting the partial derivative of (5.84) with respect to c equal to zero. Analogously, we have the first order condition on the portfolio composition. Indeed we get $\bar{u}^A_{sup}$ by setting the partial derivative of (5.84) with respect to u^A equal to zero, but in this case the calculation of the derivative is not immediate. Considering in (5.84) only the terms depending on u^A, i.e., $A_2^{u^A}J$, and setting

$$w = u^{A\prime}, \quad m = V\left(MJ_V + \tilde{\Sigma}^A\tilde{\Pi}'J_{zV} + VS\tilde{\Sigma}^A\tilde{\sigma}^{S\prime}J_{VS}\right), \quad d = V^2J_{VV}, \tag{5.86}$$

we can rewrite $A_2^{u^A}J$ as

$$f(w) = w'\,m + \frac{1}{2}d\,w'Qw = \sum_{\ell=1}^{3}w_\ell m_\ell + \frac{1}{2}d\sum_{\ell=1}^{3}w_\ell\sum_{k=1}^{3}Q_{\ell k}w_k.$$

Then taking into account that $Q_{ik} = Q_{ki}$, the partial derivative of f with respect to w_i is given by[14]

$$\frac{\partial f}{\partial w_i} = m_i + d\,(Qw)_i \tag{5.87}$$

where $(v)_i$ denotes the ith component of the vector v. Finally we obtain $\bar{u}^A_{sup}$ by setting (5.87) equal to zero, thus giving us the equation

$$(Qw)_i = -\frac{m_i}{d},$$

i.e.,

$$Qu^{A\prime} = -\frac{MJ_V + \tilde{\Sigma}^A\tilde{\Pi}'J_{zV} + S\tilde{\Sigma}^A\tilde{\sigma}^{S\prime}J_{VS}}{VJ_{VV}}, \tag{5.88}$$

[14] Observe that

$$f(w) = \sum_{k=1}^{3}\sum_{\ell=1}^{3}w_\ell Q_{\ell k}w_k = w_i^2Q_{ii} + \sum_{k\neq i}w_iQ_{ik}w_k + \sum_{\ell\neq i}w_\ell Q_{\ell i}w_i = w_i^2Q_{ii} + 2\sum_{k\neq i}w_iQ_{ik}w_k,$$

then

$$\frac{\partial f}{\partial w_i} = 2\,w_iQ_{ii} + 2\sum_{k\neq i}Q_{ik}w_k = 2(Qw)_i.$$

which gives us the compact form of (5.72)-(5.75):

$$\bar{u}^{A\prime}_{sup}(t,Z) = -\frac{\frac{\partial J}{\partial V}(t,Z)}{V\frac{\partial^2 J}{\partial V^2}(t,Z)}Q^{-1}M - \frac{1}{V\frac{\partial^2 J}{\partial V^2}(t,Z)}Q^{-1}\tilde{\Sigma}^A\tilde{\Pi}'\frac{\partial^2 J}{\partial z\partial V}(t,Z) - \frac{S\frac{\partial^2 J}{\partial V\partial S}(t,Z)}{V\frac{\partial^2 J}{\partial V^2}(t,Z)}Q^{-1}\tilde{\Sigma}^A\tilde{\sigma}^{S\prime}. \tag{5.89}$$

Indeed the matrix Q is invertible, since $\tilde{\Sigma}^A$ is invertible (see Assumption 5.2.1),

$$Q^{-1} = \left(\tilde{\Sigma}^A\tilde{\Sigma}^{A\prime}\right)^{-1} = \left(\tilde{\Sigma}^{A\prime}\right)^{-1}\left(\tilde{\Sigma}^A\right)^{-1} = \left(\left(\tilde{\Sigma}^A\right)^{-1}\right)'\left(\tilde{\Sigma}^A\right)^{-1}, \tag{5.90}$$

and

$$((\tilde{\Sigma}^A)^{-1})' = \begin{pmatrix} \frac{1}{\frac{\hat{B}_r^T}{\hat{B}^T}\hat{\sigma}^r} & -\frac{\frac{\hat{L}_r^T}{\hat{L}^T}}{\frac{\hat{B}_r^T}{\hat{B}^T}\frac{\hat{L}_\lambda^T}{\hat{L}^T}\hat{\sigma}^\lambda} & -\frac{1}{\frac{\hat{B}_r^T}{\hat{B}^T}\hat{\sigma}^r}\frac{\hat{\sigma}_r^S}{\hat{\sigma}_S^S} \\ 0 & \frac{1}{\frac{\hat{L}_\lambda^T}{\hat{L}^T}\hat{\sigma}^\lambda} & 0 \\ 0 & 0 & \frac{1}{\hat{\sigma}_S^S} \end{pmatrix},$$

and finally, taking into account $\tilde{\Sigma}^A\tilde{\xi} = M$, after some reshuffling, we have the announced result[15].

□

Proof of Theorem 5.3.6. By (5.72) and (5.75) we see clearly (compare point 3 in the scheme in the end of Appendix D.1) that $\bar{c}_{sup}$ and $\bar{u}^A_{sup} = \left(\bar{u}^B_{sup}, \bar{u}^L_{sup}, \bar{u}^S_{sup}\right)$ will of course depend on our choice of t and Z, but it will also depend on the function J and its partial derivatives. As already discussed in same scheme (see point 4.), inserting (5.72)-(5.75) into the partial differential equation (5.66), we get the HJB equation (5.76) with boundary condition (5.77). Then by the Verification Theorem (see Theorem D.1.2) we obtain the announced results. □

In particular turning to the problem (5.43) and taking into account the results of Theorem 5.3.6, the corresponding value function J is given by the value function $\bar{J}$ with $z_0 = 1$, i.e.,

$$J(t_0, z, S, V) = \bar{J}(t_0, 1, z, S, V), \tag{5.91}$$

[15] To obtain the announced results we need to compute

$$Q^{-1}M = ((\tilde{\Sigma}^A)^{-1})'\,\tilde{\xi} = \begin{pmatrix} \frac{\hat{\xi}_r\hat{\sigma}_S^S - \hat{\xi}_S\hat{\sigma}_r^S}{\frac{\hat{B}_r^T}{\hat{B}^T}\hat{\sigma}^r\,\hat{\sigma}_S^S} - \frac{\frac{\hat{L}_r^T}{\hat{L}^T}}{\frac{\hat{B}_r^T}{\hat{B}^T}}\frac{\hat{\xi}_\lambda}{\frac{\hat{L}_\lambda^T}{\hat{L}^T}\hat{\sigma}^\lambda} \\ \frac{\hat{\xi}_\lambda}{\frac{\hat{L}_\lambda^T}{\hat{L}^T}\hat{\sigma}^\lambda} \\ \frac{\hat{\xi}_S}{\hat{\sigma}_S^S} \end{pmatrix},$$

$$Q^{-1}\tilde{\Sigma}^A\tilde{\Pi}'J_{zV} = ((\tilde{\Sigma}^A)^{-1})'\tilde{\Pi}'J_{zV}, = \begin{pmatrix} J_{rV}\frac{1}{\frac{\hat{B}_r^T}{\hat{B}^T}} - J_{\lambda V}\frac{\frac{\hat{L}_r^T}{\hat{L}^T}}{\frac{\hat{B}_r^T}{\hat{B}^T}\frac{\hat{L}_\lambda^T}{\hat{L}^T}} \\ J_{\lambda V}\frac{1}{\frac{\hat{L}_\lambda^T}{\hat{L}^T}} \\ 0 \end{pmatrix},$$

$$Q^{-1}\tilde{\Sigma}^A\tilde{\sigma}^{S\prime} = ((\tilde{\Sigma}^A)^{-1})'\tilde{\sigma}^{S\prime} = \begin{pmatrix} \frac{1}{\frac{\hat{B}_r^T}{\hat{B}^T}}\left(\frac{\hat{\sigma}_r^S}{\hat{\sigma}^r} - \frac{\hat{\sigma}_r^S}{\hat{\sigma}^r}\right) \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

and substitute the above expressions in (5.89).

and the optimal controls are

$$c_{sup}(t) = \bar{c}_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t); \bar{J}\big), \tag{5.92}$$

$$u^{B}_{sup}(t) = \bar{u}^{B}_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t); \bar{J}\big), \tag{5.93}$$

$$u^{L}_{sup}(t) = \bar{u}^{L}_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t); \bar{J}\big), \tag{5.94}$$

$$u^{S}_{sup}(t) = \bar{u}^{S}_{sup}\big(t, 1, r(t), \lambda(t), S(t), V(t); \bar{J}\big), \tag{5.95}$$

where $\bar{c}_{sup}$, $\bar{u}^{B}_{sup}$, $\bar{u}^{L}_{sup}$, and $\bar{u}^{S}_{sup}$ are given by (5.72)-(5.75).

Remark 5.3.2. *First of all observe that in a market without longevity bonds, the results are similar, but instead of the functions $\bar{c}_{sup}$, $\bar{u}^{B}_{sup}$, $\bar{u}^{L}_{sup}$, $\bar{u}^{S}_{sup}$, and $\bar{u}^{G}_{sup} = 1 - \bar{u}^{B}_{sup} - \bar{u}^{L}_{sup} - \bar{u}^{S}_{sup}$, one would get functions $\bar{c}^{0}_{sup}$, $\bar{u}^{B,0}_{sup}$, $\bar{u}^{S,0}_{sup}$, and $\bar{u}^{G,0}_{sup} = 1 - \bar{u}^{B,0}_{sup} - \bar{u}^{S,0}_{sup}$, which do not depend on λ. Namely, setting*

$$\bar{\bar{u}}^{B}_{sup}(t, Z; H) = \frac{1}{\frac{\hat{B}^T_r}{\hat{B}^T}}\,\frac{1}{V}\left(-\frac{H_V}{H_{VV}}\,\frac{\hat{\xi}_r\hat{\sigma}^S_S - \hat{\xi}_S\hat{\sigma}^S_r}{\hat{\sigma}^r\,\hat{\sigma}^S_S} - \frac{H_{rV}}{H_{VV}}\right),$$

then, for any function $H(t, r, S, V)$, which does not depend on λ, one gets

$$\begin{aligned}
\bar{c}^{0}_{sup}(t, r, S, V; H) &= \bar{c}_{sup}(t, 1, r, \lambda, S, V; H)\;(= \bar{c}_{sup}(t, 1, r, 0, S, V; H))\\
\bar{u}^{B,0}_{sup}(t, r, S, V; H) &= \bar{\bar{u}}^{B}_{sup}(t, 1, r, \lambda, S, V; H)\;\big(= \bar{\bar{u}}^{B}_{sup}(t, 1, r, 0, S, V; H)\big)\\
\bar{u}^{S,0}_{sup}(t, r, S, V; H) &= \bar{u}^{S}_{sup}(t, 1, r, \lambda, S, V; H)\;\big(= \bar{u}^{S}_{sup}(t, 1, r, 0, S, V; H)\big),
\end{aligned}$$

and therefore

$$\begin{aligned}
\bar{u}^{G,0}_{sup}(t, r, S, V; H) &= 1 - \bar{\bar{u}}^{B}_{sup}(t, 1, r, \lambda, S, V; H) - \bar{u}^{S}_{sup}(t, 1, r, \lambda, S, V; H)\\
&\quad\big(= 1 - \bar{\bar{u}}^{B}_{sup}(t, 1, r, 0, S, V; H) - \bar{u}^{S}_{sup}(t, 1, r, 0, S, V; H)\big).
\end{aligned}$$

Furthermore, observe that

$$\bar{u}^{B}_{sup}(t, 1, r, \lambda, S, V; \bar{J}) = \bar{\bar{u}}^{B}_{sup}(t, 1, r, \lambda, S, V; \bar{J}) - \frac{\frac{\hat{L}^T_r}{\hat{L}^T}}{\frac{\hat{B}^T_r}{\hat{B}^T}}\,\bar{u}^{L}_{sup}(t, 1, r, \lambda, S, V; \bar{J})$$

and therefore

$$\begin{aligned}
\bar{u}^{G}_{sup}(t, 1, r, \lambda, S, V; \bar{J}) &= 1 - \bar{\bar{u}}^{B}_{sup}(t, 1, r, \lambda, S, V; \bar{J}) - \bar{u}^{S}_{sup}(t, 1, r, \lambda, S, V; \bar{J})\\
&\quad - \left(1 - \frac{\frac{\hat{L}^T_r}{\hat{L}^T}}{\frac{\hat{B}^T_r}{\hat{B}^T}}\right)\bar{u}^{L}_{sup}(t, 1, r, \lambda, S, V; \bar{J}).
\end{aligned}$$

Finally, as remarked in Menoncin [18], we observe that, the previous expression may be interpreted as follows: in a certain sense, $u^{L}_{sup}(t)V(t)$, the optimal amount of money to be invested in the longevity bond, is taken from both the amounts of money $\bar{\bar{u}}^{B}_{sup}(t, Z(t); \bar{J})V(t)$ and $\bar{u}^{G}_{sup}(t, Z(t); \bar{J})V(t)$, that one would invest in the ordinary bond and the riskless asset (i.e., the liquidity) proportionally to $\frac{L^T_r}{L^T}/\frac{B^T_r}{B^T}$ and $1 - \frac{L^T_r}{L^T}/\frac{B^T_r}{B^T}$, respectively.

5.3.2 Optimal Markov control problem with the risk asset given by a Geometric Brownian motion

The model considered in Section 5.2 is slightly more general than the original model considered by Menoncin in [18], the difference being that the risk asset $S(t)$ is a Geometric Brownian motion, i.e., the drift and diffusion coefficients are constants: $\mu_S(t) = \mu_S$ and $\sigma_S(t) = \sigma_S$. In this section we specialize to this setting and find again the same results of Menoncin [18], but using a slightly different approach. Moreover (see Remark 5.3.3) we observe that this approach can be easily generalized to the case when the drift and diffusion coefficients $\mu_S(t)$ and $\sigma_S(t)$ are deterministic functions of time and the interest rate, but do not depend on the risk asset.

As in Section 5.2, let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let τ be the death time of the investor, and the vector process $z(t) = (r(t), \lambda(t))$ be the state variables vector. Taking into account that in this setting the process $S(t)$ is a geometric Brownian motion (see (2.63)), let $\xi_S(t)$ be the market price for the stock given by (2.66), then the expression (5.3) for the market price $\xi(t)$ becomes

$$\xi(t) = \tilde{\xi}(t, z(t)) = \Big(\hat{\xi}_r(t, r(t)), \hat{\xi}_\lambda(t, z(t)), \hat{\xi}_S(t, r(t))\Big)'. \tag{5.96}$$

In this setting we have that the market structure is given by (5.4) and (5.5), i.e.,

$$\begin{aligned} dz(t) &= \mu^z(t)dt + \Pi(t)dW(t), \\ dA(t) &= diag\big[A(t)\big]\left(\mu^A(t)dt + \Sigma^A(t)dW(t)\right), \end{aligned}$$

where $\mu^z(t)$, $W(t)$, $\Pi(t)$, $A(t)$ and $diag\big[A(t)\big]$ are given by (5.6), (5.7), (5.8) and (5.9), while

$$\begin{aligned} \mu^A(t) &= \tilde{\mu}^A(t, z(t)) \\ &= \begin{pmatrix} r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\frac{\hat{B}^T_r}{\hat{B}^T}(t, r(t)) \\ r(t) + \lambda(t) + \frac{\hat{L}^T_r}{\hat{L}^T}(t, z(t))\hat{\sigma}^r(t, r(t))dW^r(t) + \frac{\hat{L}^T_\lambda}{\hat{L}^T}(t, z(t))\hat{\sigma}^\lambda(t, z(t))dW^\lambda(t) \\ r(t) + \sigma^S_r\hat{\xi}_r(t, r(t)) + \sigma^S_S\hat{\xi}_S(t, r(t)) \end{pmatrix}, \end{aligned} \tag{5.97}$$

$$\begin{aligned} \Sigma^A(t) &= \tilde{\Sigma}^A(t, z(t)) \\ &= \begin{pmatrix} \hat{\sigma}^r(t, r(t))\frac{\hat{B}^T_r}{\hat{B}^T}(t, r(t)) & 0 & 0 \\ \hat{\sigma}^r(t, r(t))\frac{\hat{L}^T_r}{\hat{L}^T}(t, z(t)) & \hat{\sigma}^\lambda(t, z(t))\frac{\hat{L}^T_\lambda}{\hat{L}^T}(t, z(t)) & 0 \\ \sigma^S_r & 0 & \sigma^S_S \end{pmatrix}, \end{aligned} \tag{5.98}$$

and σ^S_r, σ^S_S are deterministic constants.

Remark 5.3.3. *We observe that σ^S_r and σ^S_S are constants (see (5.97) and (5.98)), then $\mu^A(t)$ and $\Sigma^A(t)$ are deterministic functions of t and $z(t)$. As we will see below, this property of $\mu^A(t)$ and $\Sigma^A(t)$ is crucial for the results that we get at the end of this section. This property holds also under the condition* 1. *of Remark 2.3.1 that the drift and diffusion coefficients of $S(t)$ depend on t, $r(t)$, i.e., $\mu^S(t) = \hat{\mu}^S(t, r(t))$, $\sigma^S_r(t) = \hat{\sigma}^S_r(t, r(t))$ and $\sigma^S_S(t) = \hat{\sigma}^S_S(t, r(t))$. As observed in in Remark 2.3.1, then also $\xi_S(t)$ is also a deterministic function of t and $r(t)$.*

The purpose of this section is to specialize the optimal Markov control problem of Section 5.3.1 to this setting. To this end we consider control processes that depend on $z_0(t)$, $z(t)$, and $V(t)$, (see (5.44)-(5.46))

$$u^A(t) = \hat{u}^A(t, z_0(t), z(t), V(t)) \tag{5.99}$$

$$c(t) = \hat{c}(t, z_0(t), z(t), V(t)), \tag{5.100}$$

for some measurable deterministic function $\hat{u}^A$ and $\hat{c}$, where $\hat{u}^A, \hat{c} \in \hat{\mathcal{U}}_{ad}$ with value in $U_{ad} = \mathbb{R}^3 \times [0, \infty)$. We will denote the class of admissible Markov control processes as

$$\hat{\mathfrak{U}}_{ad} = \{u^A(\cdot), c(\cdot) : u^A(t) = \hat{u}^A(t, z_0(t), z(t), V(t)),\ c(t) = \hat{c}(t, z_0(t), z(t), V(t))\}.$$

Then we consider the following optimal Markov control problem

$$\bar{J}(t_0, z_0, z, V) = \sup_{u^A, c \in \hat{\mathfrak{U}}_{ad}} E^P\left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t) U\big(c(t)V(t)\big)dt \Big| z_0, z, V\right], \tag{5.101}$$

with dynamics of state processes given by

$$dz_0(t) = -\lambda(t)z_0(t)dt \tag{5.102}$$

$$dz(t) = \mu^z(t)dt + \Pi(t)dW(t) \tag{5.103}$$

$$dV(t) = V(t)\left[\left(r(t) + u^A(t)\left(\mu^A(t) - r(t)\mathbf{1} - \lambda\mathbf{1}_\lambda\right) - c(t)\right)dt + u^A(t)\Sigma^A(t)dW(t)\right], \tag{5.104}$$

where $\Sigma^A(t)$ and $\mu^A(t)$ are given by (5.97) and (5.98) respectively, and initial conditions

$$z_0(t_0) = z_0, \quad z(t_0) = z, \quad V(t_0) = V, \tag{5.105}$$

under the constraint

$$c(t) \geq 0. \tag{5.106}$$

In the sequel we denote the above optimal Markov control problem as $(\hat{\mathbf{P}})$.

Observe that the difference between the problem $(\bar{\mathbf{P}})$ and the problem $(\hat{\mathbf{P}})$ is that in $(\bar{\mathbf{P}})$ the process $S(t)$ is the state process, while here is not, and $\mu^A(t)$ and $\Sigma^A(t)$ in $(\bar{\mathbf{P}})$ depend on t, $z(t)$, and $S(t)$, while here depend only on t and $z(t)$, but do not depend on $S(t)$.

In the sequel, for the notational convenience, we use the following notation

$$\check{Z}(t) = \begin{pmatrix} \check{z}_0(t) \\ \check{z}_1(t) \\ \check{z}_2(t) \\ \check{z}_3(t) \end{pmatrix} = \begin{pmatrix} z_0(t) \\ r(t) \\ \lambda(t) \\ V(t) \end{pmatrix}, \tag{5.107}$$

with the dynamics given by

$$\begin{cases} d\check{Z}(t) = \mu(t)dt + \Gamma(t)dW(t), \\ \check{Z}(t_0) = \check{Z}, \end{cases} \tag{5.108}$$

where $\check{Z} = (z_0,\, r,\, \lambda,\, V)'$, and

$$\mu(t) = \hat{\mu}(t, u^A(t), c(t), Z(t)) \tag{5.109}$$

$$\Gamma(t) = \hat{\Gamma}(t, u^A(t), c(t), Z(t)), \tag{5.110}$$

with, as already recalled, $u^A(t)$ and $c(t)$ given by (5.99) and (5.100), and $\hat{\mu}(t, u^A, c, Z)$ and $\hat{\Gamma}(t, u^A, c, Z)$ given by

$$\hat{\mu} = \begin{pmatrix} -\lambda z_0 \\ \hat{\mu}^r \\ \hat{\mu}^\lambda \\ V\left(r + u^A\left(\tilde{\mu}^A - r\mathbf{1} - \lambda \mathbf{1}_\lambda\right) - c\right) \end{pmatrix}, \tag{5.111}$$

$$\hat{\Gamma} = \begin{pmatrix} 0 & 0 & 0 \\ \hat{\sigma}^r & 0 & 0 \\ 0 & \hat{\sigma}^\lambda & 0 \\ V\left(u^B \hat{\sigma}^r \frac{\hat{B}_r^T}{\hat{B}^T} + u^L \hat{\sigma}^r \frac{\hat{L}_r^T}{\hat{L}^T} + u^S \sigma_r^S\right) & V u^L \hat{\sigma}^\lambda \frac{\hat{L}_\lambda^T}{\hat{L}^T} & V u^S \sigma_S^S \end{pmatrix}. \tag{5.112}$$

Observe that the difference between (5.111)-(5.112) and (5.63)-(5.64) is that $\tilde{\mu}^A$ in (5.63) may depend also on $S(t)$, while here $\tilde{\mu}^A$ depends only on t and $z(t)$, and $\sigma_r^S(t)$ and $\sigma_S^S(t)$ in (5.64) are functions depending on t, $S(t)$ and $r(t)$, while here are constants[16], i.e., are equal to σ_r^S and σ_S^S.

Denoting the conditional expectation given $\check{Z}(t_0) = \check{Z}$ as $E^P_{t_0,\check{Z}}$, the value function for $(\hat{\mathbf{P}})$ can be rewritten as

$$\hat{J}(t_0, \check{Z}) = \sup_{u^A(\cdot), c(\cdot) \in \hat{\mathfrak{U}}_{ad}} E^P_{t_0,\check{Z}}\left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t) U\big(c(t)V(t)\big)dt\right]. \tag{5.113}$$

By using the stochastic dynamic programming technique (see Appendix D.1), and recalling that $c(t)$ and $u^A(t)$ are given by (5.99) and (5.100), the corresponding Hamilton-Jacobi-Bellman equation is given by

$$\frac{\partial J}{\partial t_0}(t_0, \check{Z}) + \sup_{u^A, c \in U_{ad}} \left\{e^{-\rho t_0} z_0 U\big(Vc\big) + \mathcal{A}^{u^A,c} J(t_0, \check{Z})\right\}, \tag{5.114}$$

where

$$U_{ad} = \mathbb{R}^3 \times [0, \infty),$$

and

$$\mathcal{A}^{u^A,c} J(t_0, \check{Z}) = \sum_{i=0}^{3} \hat{\mu}_i(t_0, u^A, c, \check{Z}) \frac{\partial J}{\partial z_i}(t_0, \check{Z}) + \frac{1}{2} \sum_{i,j=1}^{3} \left[\left(\hat{\Gamma}\hat{\Gamma}'\right)(t_0, u^A, c, \check{Z})\right]_{i,j} \frac{\partial^2 J}{\partial z_i \partial z_j}(t_0, \check{Z}),$$

$$F(t_0, u^A, c, \check{Z}) = e^{-\rho t_0} z_0 U\big(Vc\big).$$

[16] Similar considerations hold when we are in the case considered in Remark 5.3.3.

Following the scheme at the end of Appendix D.1 (see points $1-5$), now we consider point 1. In the sequel for the notational convenience we denote $t_0 = t$ and by a little abuse of notation we denoted the function H of Appendix D.1 with J.

Fixed an arbitrary point $(t, \check{Z}) \in (0, \infty) \times (0, 1] \times (0, \infty)^3$ and any function $J(t, Z)$ sufficiently smooth, we now have to solve the optimization problem

$$\sup_{u^A, c \in \hat{U}_{ad}} \left\{ e^{-\rho t} z_0 U\big(Vc\big) + \mathcal{A}^{u^A, c} J(t, \check{Z}) \right\}. \tag{5.115}$$

We remember that here u^A and c are the only variables, whereas t and $\check{Z}$ are considered to be fixed parameters (see (D.17) at point 2 with $v = (u^A, c)$).

Using the notations (5.68)-(5.70), we have the following proposition.

Proposition 5.3.7. *Under the completeness Assumption 5.2.1, let $J(t, \check{Z})$ be a regular function*[17]*, i.e., $J \in C^{1,2}$, and let $\hat{c}_{sup}(t, \check{Z}; J)$ and $\hat{u}^A_{sup}(t, \check{Z}; J)$ be the functions such that for each fixed choice of $(t, \check{Z})$ and any function $J \in C^{1,2}$ are the solutions of the optimization problem (5.115). Then*[18]

$$\hat{c}_{sup}(t, \check{Z}; J) = \frac{1}{V} \dot{U}^{-1} \left(\frac{J_V}{e^{-\rho\, t} z_0} \right), \tag{5.116}$$

$$\hat{u}^B_{sup}(t, \check{Z}; J) = \frac{1}{\frac{\hat{B}^T_r}{\hat{B}^T}} \frac{1}{V} \left(- \frac{J_V}{J_{VV}} \frac{\hat{\xi}_r \bar{\sigma}^S_S - \xi_S \bar{\sigma}^S_r}{\hat{\sigma}^r\, \bar{\sigma}^S_S} - \frac{J_{rV}}{J_{VV}} \right) - \frac{\frac{\hat{L}^T_r}{\hat{L}^T}}{\frac{\hat{B}^T_r}{\hat{B}^T}} \frac{1}{V} \left(- \frac{J_V}{J_{VV}} \frac{\hat{\xi}_\lambda}{\frac{\hat{L}^T_\lambda}{\hat{L}^T} \hat{\sigma}^\lambda} - \frac{1}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}} \frac{J_{V\lambda}}{J_{VV}} \right) \tag{5.117}$$

$$\hat{u}^L_{sup}(t, \check{Z}; J) = \frac{1}{V} \left(- \frac{J_V}{J_{VV}} \frac{\hat{\xi}_\lambda}{\frac{\hat{L}^T_\lambda}{\hat{L}^T} \hat{\sigma}^\lambda} - \frac{1}{\frac{\hat{L}^T_\lambda}{\hat{L}^T}} \frac{J_{V\lambda}}{J_{VV}} \right) \tag{5.118}$$

$$\hat{u}^S_{sup}(t, \check{Z}; J) = -\frac{1}{V} \frac{J_V}{J_{VV}} \frac{\xi_S}{\bar{\sigma}^S_S}, \tag{5.119}$$

where the argument $(t, \check{Z})$ "has been suppressed" for the notational convenience.

Theorem 5.3.8. *Let us consider the following HJB problem for $(\hat{\mathbf{P}})$*

$$\frac{\partial H}{\partial t}(t, \check{Z}) + e^{-\rho t} z_0 U \left(\dot{U}^{-1} \left(\frac{\frac{\partial H}{\partial V}(t, \check{Z})}{e^{-\rho t} z_0} \right) \right) + \mathcal{L}^{\hat{u}^A_{sup}(\cdot; H), \hat{c}_{sup}(\cdot; H)} H(t, \check{Z}), \tag{5.120}$$

with the boundary condition given by

$$\lim_{t \to \infty} H(t, z_0, z, V, S) = 0, \tag{5.121}$$

and assume that (5.120) admits a unique classical solution J. Then J coincides with the value function $\hat{J}$ for problem $(\hat{\mathbf{P}})$ (see e.g. (5.113)) and the optimal controls are

$$c_{sup}(t) = \hat{c}_{sup}\big(t, \check{Z}(t); J\big), \tag{5.122}$$

$$u^B_{sup}(t) = \hat{u}^B_{sup}\big(t, \check{Z}(t); J\big), \tag{5.123}$$

$$u^L_{sup}(t) = \hat{u}^L_{sup}\big(t, \check{Z}(t); J\big), \tag{5.124}$$

$$u^S_{sup}(t) = \hat{u}^S_{sup}\big(t, \check{Z}(t); J\big), \tag{5.125}$$

where $\hat{c}_{sup}$, $\hat{u}^B_{sup}$, $\hat{u}^L_{sup}$, and $\hat{u}^S_{sup}$ are given by (5.116)-(5.119).

Obviously also in this setting Remark 5.3.2 holds with suitable adjustments such as, for example, $\hat{c}_{sup}$ and $\hat{u}^A_{sup}$ instead of $\bar{c}_{sup}$ and $\bar{u}^A_{sup}$, respectively.

Proof of Proposition 5.3.7. The first order condition on consumption is exactly as in the proof of Proposition 5.3.5. For the first order condition on the portfolio composition we proceed as in the proof of Proposition 5.3.5 without the terms depending $S(t)$, and by using the same notations we have

$$A^{u^A}_2 J = V u^A M J_V + \frac{1}{2} V^2 u^A Q u^{A\prime} J_{VV} + V u^A \tilde{\Sigma}^A \tilde{\Pi}' J_{zV}, \tag{5.126}$$

[17] We recall that here J is not necessarily a value function.

[18] Observe that $\hat{c}_{sup}$ and $\hat{u}^A_{sup}$ are the corresponding function u^*_J of point 2 in the scheme in the end of Appendix D.1.

where $Q = \tilde{\Sigma}^A \tilde{\Sigma}^{A\prime}$, and $M(t) = (\mu^A(t) - r(t)\mathbf{1} - \lambda \mathbf{1}_\lambda)$. Setting

$$w = u^{A\prime}, \quad m = V\left(MJ_V + \tilde{\Sigma}^A \tilde{\Pi}' J_{zV}\right), \quad d = V^2 J_{VV}, \tag{5.127}$$

we obtain

$$(Qw)_i = -\frac{m_i}{d}, \tag{5.128}$$

i.e.,

$$Qu^{A\prime} = -\frac{MJ_V + \tilde{\Sigma}^A \tilde{\Pi}' J_{zV}}{VJ_{VV}}, \tag{5.129}$$

which gives us the compact form of (5.116)-(5.119):

$$\hat{u}^{A\prime}_{sup}(t, \check{Z}) = -\frac{\frac{\partial J}{\partial V}(t, \check{Z})}{V\frac{\partial^2 J}{\partial V^2}(t, \check{Z})} Q^{-1} M - \frac{1}{V\frac{\partial^2 J}{\partial V^2}(t, \check{Z})} Q^{-1} \tilde{\Sigma}^A \tilde{\Pi}' \frac{\partial^2 J}{\partial z \partial V}(t, \check{Z}). \tag{5.130}$$

Finally, after some reshuffling, (see the proof of Proposition 5.3.5 with $\hat{u}^A_{sup}(t)$ instead of $\bar{u}^A_{sup}(t)$), we obtain the announced result.

□

Proof of Theorem 5.3.8. By (5.116) and (5.130) we see clearly (compare point 3 in the scheme in the end of Appendix D.1) that $\hat{c}_{sup}$ and $\hat{u}^A_{sup} = \left(\hat{u}^B_{sup}, \hat{u}^L_{sup}, \hat{u}^S_{sup}\right)$ will of course depend on our choice of t and $\check{Z}$, but it will also depend on the function J, and its partial derivatives. As already discussed in the same scheme (see point 4.), inserting (5.116)-(5.130) into the partial differential equation (5.114), we get the HJB equation (5.120) with boundary condition (5.121). Then by the Verification Theorem (see Theorem D.1.2) we obtain the announced results.

□

In particular turning to the problem considered by Menoncin in [18], and taking into account the results of Theorem 5.3.6, we find the same results of Menoncin [18]. Indeed, the corresponding value function the problem considered by Menoncin, denoted by J, is given by the value function $\hat{J}$ with $z_0 = 1$, i.e.,

$$J(t_0, z, V) = \hat{J}(t_0, 1, z, V),$$

and the optimal controls are

$$c_{sup}(t) = \hat{c}_{sup}\big(t, 1, r(t), \lambda(t), V(t); \hat{J}\big), \tag{5.131}$$
$$u^B_{sup}(t) = \bar{u}^B_{sup}\big(t, 1, r(t), \lambda(t), V(t); \hat{J}\big), \tag{5.132}$$
$$u^L_{sup}(t) = \bar{u}^L_{sup}\big(t, 1, r(t), \lambda(t), V(t); \hat{J}\big), \tag{5.133}$$
$$u^S_{sup}(t) = \bar{u}^S_{sup}\big(t, 1, r(t), \lambda(t), V(t); \hat{J}\big), \tag{5.134}$$

where $\hat{c}_{sup}$, $\hat{u}^B_{sup}$, $\hat{u}^L_{sup}$, and $\hat{u}^S_{sup}$ are given by (5.116)-(5.130).

5.4 Financial Market with Rolling (UOS market)

In this section we present another financial model on which we will work: we consider a market model which, besides the money account $G(t)$ and the risk asset with price process $S(t)$, contains a (discrete-time) rolling bond and a (discrete-time) rolling longevity bond, with price processes $U^\Delta(t,T)$ and $O^\Delta(t,T)$, where T is a fixed maturity time. The latter (zero coupon) bonds are introduced in Sections 2.4 and 4.4, respectively. In the sequel we will shortly refer to this market as the UOS market model.

We will extend the results of Section 5.2 to this setting and, to keep this section self-contained, we will repeat also some of the notations and assumptions already introduced in Section 5.2.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let τ be the death time of the investor, and the vector process $z(t) = (r(t), \lambda(t))$ be the state variables vector where the processes $r(t)$ and $\lambda(t)$ are referred to as the riskless interest rate, and the stochastic mortality intensity of the investor, respectively.

Summarizing, we assume that

$$P(\tau < \infty) = 1 \tag{5.135}$$

and that the market is described by two structures, i.e., the so called state variables described by the vector process $z(t) = (r(t), \lambda(t))$, and the financial assets traded on the market. In details, using the notations introduced in the previous chapters, the vector process $z(t)$ evolves as follows

$$dz(t) = \mu^z(t)dt + \Sigma^z(t)dW^z(t), \tag{5.136}$$

(see (4.9), (4.10) and (4.11)), the money market account $G(t)$ is given by (1.3), and the financial assets are

1. A discrete-time rolling bond, with maturity T, with price process $U^\Delta(t,T)$;
2. A discrete-time longevity rolling bond, with maturity T, with price process $O^\Delta(t,T)$;
3. A risk asset with price process $S(t)$.

Furthermore, since (5.135) holds, then (see Lemma 3.3.7) the process $\lambda(t)$ satisfies the following condition

$$P\left(\int_{t_0}^{\infty} \lambda(u)du = \infty\right) = 1, \quad \forall t_0 \geq 0.$$

By the results obtained in Chapters 2 and 4, let the process $\xi_S(t)$ be the market price for the stock given by (2.70), and let $\xi_z(t) = (\xi_r(t), \xi_\lambda(t))'$ be the market price for the riskless interest rate and the longevity risk given by (4.32) and (4.33). In the sequel we denote by $\xi(t)$ the market price, where

$$\xi(t) = \tilde{\xi}(t, z(t), S(t)) = \left(\hat{\xi}_r(t, r(t)), \hat{\xi}_\lambda(t, z(t)), \hat{\xi}_S(t, r(t), S(t))\right)'. \tag{5.137}$$

Then the processes $U^\Delta(t,T)$, $O^\Delta(t,T)$ and $S(t)$ can be described by the differential equations (2.74), (4.90), and (2.71), respectively, so that we can summarize the UOS market structures in the follow matrix form

$$dz(t) = \mu^z(t)dt + \Pi(t)dW(t), \tag{5.138}$$
$$dA^\Delta(t) = diag\big[A^\Delta(t)\big]\left(\mu^{A,\Delta}(t)dt + \Sigma^{A,\Delta}(t)dW(t)\right), \tag{5.139}$$

where

$$\mu^z(t) = \tilde{\mu}^z(t, z(t)) = \begin{pmatrix} \hat{\mu}^r(t, r(t)) \\ \hat{\mu}^\lambda(t, z(t)) \end{pmatrix}, \qquad W(t) = \begin{pmatrix} W^r(t) \\ W^\lambda(t) \\ W^S(t) \end{pmatrix}, \tag{5.140}$$

$$\Pi(t) = \tilde{\Pi}(t, z(t)) = \begin{pmatrix} \hat{\sigma}^r(t, r(t)) & 0 & 0 \\ 0 & \hat{\sigma}^\lambda(t, z(t)) & 0 \end{pmatrix}, \tag{5.141}$$

and

$$A^\Delta(t) = \begin{pmatrix} U^\Delta(t,T) \\ O^\Delta(t,T) \\ S(t) \end{pmatrix}, \quad diag\big[A^\Delta(t)\big] = \begin{pmatrix} U^\Delta(t,T) & 0 & 0 \\ 0 & O^\Delta(t,T) & 0 \\ 0 & 0 & S(t) \end{pmatrix},$$

$$\begin{aligned}\mu^{A,\Delta}(t) &= \tilde{\mu}^{A,\Delta}(t, z(t), S(t)) \\ &= \begin{pmatrix} r(t) + \hat{\xi}_r(t, r(t))\hat{\sigma}^r(t, r(t))\frac{\hat{B}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{B}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, r(t)) \\ r(t) + \lambda(t) + \frac{\hat{L}_z^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, z(t))\tilde{\Sigma}^z(t, z(t))\hat{\xi}_z(t, z(t)) \\ r(t) + \hat{\sigma}_r^S(t, r(t), S(t))\hat{\xi}_r(t, r(t)) + \hat{\sigma}_S^S(t, r(t), S(t))\hat{\xi}_S(t, r(t), S(t)) \end{pmatrix},\end{aligned} \tag{5.142}$$

$$\begin{aligned}\Sigma^{A,\Delta}(t) &= \tilde{\Sigma}^{A,\Delta}(t, z(t), S(t)) \\ &= \begin{pmatrix} \hat{\sigma}^r(t, r(t))\frac{\hat{B}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{B}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, r(t)) & 0 & 0 \\ \hat{\sigma}^r(t, r(t))\frac{\hat{L}_r^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, z(t)) & \hat{\sigma}^\lambda(t, r(t))\frac{\hat{L}_\lambda^{T+\lfloor t/\Delta \rfloor \Delta}}{\hat{L}^{T+\lfloor t/\Delta \rfloor \Delta}}(t, z(t)) & 0 \\ \hat{\sigma}_r^S(t, r(t), S(t)) & 0 & \hat{\sigma}_S^S(t, r(t), S(t)) \end{pmatrix},\end{aligned} \tag{5.143}$$

Remark 5.4.1. *Observe that $\mu^{A,\Delta}(t)$ and $\Sigma^{A,\Delta}(t)$ are deterministic functions of t, $z(t)$, and $S(t)$ since on the one hand the drift and diffusion coefficients of $z(t)$ are deterministic functions of t and $z(t)$, and on the other hand the drift and diffusion coefficients of $S(t)$ are deterministic functions of t, $r(t)$ and $S(t)$ (see the condition* 2. *of Remark 2.3.1), i.e, $\mu^S(t) = \hat{\mu}^S(t, r(t), S(t))$, $\sigma_r^S(t) = \hat{\sigma}_r^S(t, r(t), S(t))$ and $\sigma_S^S(t) = \hat{\sigma}_S^S(t, r(t), S(t))$.*

In the sequel we assume the following standing condition.

Assumption 5.4.1. *We assume that the matrix $\Sigma^{A,\Delta}$ is invertible, i.e., the financial market is complete (see Corollary 1.4.5).*

Remark 5.4.2. *Let $\Sigma^{A,\Delta}(t)$ be given by (5.143). Then the UOS market is complete whenever*

$$\hat{\sigma}^r(t,r(t))\hat{B}_r^{T+\lfloor t/\Delta\rfloor\Delta} > 0, \quad \hat{\sigma}^\lambda(t,r(t))\hat{L}_\lambda^{T+\lfloor t/\Delta\rfloor\Delta} > 0, \quad \hat{\sigma}_S^S(t,r(t),S(t)) > 0 \qquad \forall(t,\omega).$$

Indeed, since the matrix $\Sigma^{A,\Delta}(t)$ is lower triangular, the functions $\hat{B}^{T+\lfloor t/\Delta\rfloor\Delta}(t,r(t))$ and $\hat{L}^{T+\lfloor t/\Delta\rfloor\Delta}(t,z(t))$ are striclty positive (see Assumptions 2.2.1 and 4.3.2), the previous conditions implies that $\Sigma^{A,\Delta}(t)$ is invertible and market completeness follows by Corollary 1.4.5.

In order to model the evolution of the stochastic mortality intensity, $\lambda(t)$, we assume that $N(t)$ is a doubly stochastic Poisson process respect to $\mathcal{G}$ as defined in (3.33), i.e.,

$$N(t) = \hat{N}\left(\int_0^t \lambda(u)du\right),$$

where the standard Poisson process $\hat{N}(t)$ is independent of the intensity process $\lambda(t)$, with respect to a suitable filtration. Before specifying the filtration we introduce a further process, the investor wealth process $V(t)$, and consider the multidimensional process $(z(t), S(t), V(t))$ (see the next Section 5.4.1). Section 5.4.2 is devoted to the assumptions on the filtration.

5.4.1 The investor's wealth in UOS market

We now form a portfolio (see Section 1.3) associated to $G(t)$, $U^\Delta(t,T)$, $O^\Delta(t,T)$ and $S(t)$, i.e., let $h(t) = (h_0(t), h_1(t), h_2(t), h_3(t))$ be the portfolio associated to $X = (X_0, X_1, X_2, X_3)$, where

$$X_0 = G(t), \quad X_1 = U^\Delta(t,T), \quad X_2 = O^\Delta(t,T), \quad X_3 = S(t),$$

and

$$h_0(t) = h^G(t), \quad (h_1(t), h_2(t), h_3(t)) = (h_T^{U,\Delta}(t), h_T^{O,\Delta}(t), h^S(t)) = h^{A,\Delta}(t).$$

Denoting the consumption rate by the process $C(t)$, we assume that (h, C) is a self-financing portfolio-consumption pair. Similarly to Section 1.3, instead of specifying $h(t)$, the absolute number of shares held of a certain asset, it may be convenient to consider $(U^G(t), U_T^{U,\Delta}(t), U_T^{O,\Delta}(t), U^S(t))$, the corresponding relative portfolio. By (1.16) and (1.17) we have

$$\begin{aligned} U^G(t) &= 1_{\{G(t)>0\}}\, u^G(t) \quad = u^G(t) \\ U_T^{U,\Delta}(t) &= 1_{\{U^\Delta(t,T)>0\}} u_T^{U,\Delta}(t) = u_T^{U,\Delta}(t) \\ U_T^{O,\Delta}(t) &= 1_{\{O^\Delta(t,T)>0\}} u_T^{O,\Delta}(t) = u_T^{O,\Delta}(t) \\ U^S(t) &= 1_{\{S(t)>0\}}\, u^S(t) \quad = u^S(t) \end{aligned}$$

for the relative portfolio corresponding to $G(t)$, $U^\Delta(t,T)$, $O^\Delta(t,T)$, and $S(t)$, with

$$u^G(t) + u_T^{U,\Delta}(t) + u_T^{O,\Delta}(t) + u^S(t) = 1. \tag{5.144}$$

Since here T is fixed, from now on we will drop the subscript T in $u_T^{U,\Delta}(t)$ and $u_T^{O,\Delta}(t)$, and so we write $u^{U,\Delta}(t)$ and $u^{O,\Delta}(t)$, respectively. Let $u^{A,\Delta}(t) = (u^{U,\Delta}(t), u^{O,\Delta}(t), u^S(t))$ be the relative portfolio corresponding to $h^{A,\Delta}(t)$. The dynamics of the value process for the self-financing portfolio-consumption pair (see (1.21)) are given by

$$\begin{cases} dV(t) = V(t)\left[u^G(t)\frac{dG(t)}{G(t)} + u^{U,\Delta}(t)\frac{dU^\Delta(t,T)}{U^\Delta(t,T)} + u^{O,\Delta}(t)\frac{dO^\Delta(t,T)-dD(t,T)}{O^\Delta(t,T)} + u^S(t)\frac{dS(t)}{S(t)}\right] - C(t)dt, \\ V(t_0) = V, \end{cases}$$

or in the compact form

$$\begin{cases} dV(t) = V(t)\left[u^G(t)\frac{dG(t)}{G(t)} + u^{A,\Delta}(t) diag^{-1}[A^\Delta(t)]dA^\Delta(t) - u^{O,\Delta}(t)\frac{dD(t,T)}{O(t,T)}\right] - C(t)dt, \\ V(t_0) = V, \end{cases}$$

where

$$u^{A,\Delta}(t) = (u^{U,\Delta}(t), u^{O,\Delta}(t), u^S(t)). \tag{5.145}$$

After substituting the expression for u^G taken from (5.13), i.e.,

$$u^G(t) = 1 - (u_T^{U,\Delta}(t) + u_T^{O,\Delta}(t) + u^S(t)) = 1 - u^{A,\Delta}(t)\mathbf{1}, \tag{5.146}$$

where $\mathbf{1} = (1,1,1)'$, the dynamics of the process $V(t)$ can be written as

$$dV(t) = V(t)\left[(1 - u^{A,\Delta}(t)\mathbf{1})\frac{dG(t)}{G(t)} + u^{A,\Delta}(t)diag^{-1}\big[A^\Delta(t)\big]dA^\Delta(t) - u^{O,\Delta}(t)\frac{dD(t,T)}{L(t,T)}\right] - C(t)dt,$$

so that, by the expression (5.139) for the differential form dA^Δ, and (4.24) for dD after some simplifications, we obtain

$$\begin{aligned} dV(t) &= V(t)\Big[\big(1 - u^{A,\Delta}(t)\mathbf{1}\big)r(t)dt + u^{A,\Delta}(t)\big(\mu^{A,\Delta}(t)dt + \Sigma^{A,\Delta}(t)dW(t)\big) - u^{O,\Delta}(t)\frac{dD(t,T)}{O(t,T)}\Big] - C(t)dt \\ &\Big[V(t)r(t) + V(t)u^{A,\Delta}(t)\big(\mu^{A,\Delta}(t) - r(t)\mathbf{1} - \lambda(t)\mathbf{1}_\lambda\big) - C(t)\Big]dt + V(t)u^{A,\Delta}(t)\Sigma^{A,\Delta}(t)dW(t), \end{aligned} \tag{5.147}$$

where $\mathbf{1}_\lambda = (0,\,1,\,0)'$.

Let us consider the agent at time t_0 with a stochastic time horizon τ, coinciding with her/his death time, i.e., she/he will act in the time interval $[t_0, \tau)$. At time t_0 the agent has the initial wealth V, and her/his problem is how to allocate investments and consumption over the time horizon. Since the admissible strategies involve consumption, and we restrict the investment-consumption pair to be self-financing, the second fundamental asset pricing theorem (see Theorem 1.4.3) is not valid. Then the objective of the agent is to choose a portfolio-consumption strategy to maximizing her/his preferences. Formally we are considering a stochastic optimal control problem. In Appendix D.1 we focus on some necessary mathematical tools for studying a general class of optimal control problems.

5.4.2 Assumptions on the filtration

Now we extend Assumption 4.2.1 and condition (4.15) on the σ-algebra $\mathcal{G}$ to this setting so that we have

$$\begin{aligned} &\mathcal{F}_t^\lambda \vee \mathcal{F}_t^r \vee \mathcal{F}_t^S \vee \mathcal{F}_t^N \vee \sigma(V) \subseteq \mathcal{F}_t, \quad \forall t \in [0,T], \\ &\mathcal{G} \supset \mathcal{F}_\infty^r \vee \mathcal{F}_\infty^\lambda \vee \mathcal{F}_\infty^S \vee \sigma(V). \end{aligned}$$

As we will see below, in this setting it is necessary to discern $\mathcal{F}_t^N$ from all other filtrations. To this end we introduce another filtration $\mathbb{G}$ such that, according to the above, the filtration $\mathbb{G}$ contains $\mathcal{F}_t^\lambda \vee \mathcal{F}_t^r \vee \mathcal{F}_t^S \vee \sigma(V)$. However, in Section 2.3, by Assumptions 1.2.1 and 1.2.2, we have considered the augmented filtration associated to the process W^S, i.e., $\bar{\mathbb{F}}^{W^S}$. Then, in the sequel, according to (4.3), it is necessary to assume that

$$\bar{\mathbb{F}}^W \subset \mathbb{G}.$$

Let us formulate this as a formalized assumption.

Assumption 5.4.2. *We assume that on $(\Omega, \mathcal{F}, P)$ there exists a σ-algebra $\mathcal{G}$ and a filtration $\mathbb{G}$ such that $\forall t \in [0,T]$*

$$\bar{\mathcal{F}}_t^W \subset \mathcal{G}_t, \tag{5.148}$$

$$\mathcal{F}_t^r \vee \mathcal{F}_t^\lambda \vee \mathcal{F}_t^S \vee \sigma(V) \subseteq \mathcal{G}_t, \tag{5.149}$$

and

$$\mathcal{F}_\infty^r \vee \mathcal{F}_\infty^\lambda \vee \mathcal{F}_\infty^S \vee \sigma(V) \subseteq \mathcal{G}. \tag{5.150}$$

As already discussed, the crucial point is the filtration with respect to which the process λ is a stochastic mortality intensity. In particular we recall that, usually, the stochastic intensity is considered with respect to a filtration $\mathbb{H}$ satisfying the usual conditions and such that

$$\mathcal{F}_t^\lambda \vee \mathcal{F}_t^N \subseteq \mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_t^N \subseteq \mathcal{G} \vee \mathcal{F}_t^N, \quad \forall t \in [0,T]. \tag{5.151}$$

Furthermore, by Proposition 3.3.4 we know that the $\mathbb{H}$-stochastic intensity is still λ. In particular we can take

$$\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_t^N. \tag{5.152}$$

5.5 Optimal control problem with Rolling

5.5.1 Optimal Markov control problem without the budged constraint

Chapter 6

The Optimal Portfolio: the CRRA utility case

6.1 Introduction

In Section 5.3, we have considered two different problem respect to the classes of admissible control processes, i.e., control processes such that depend on the processes z_0, z, V and S, (see ($\bar{\mathbf{P}}$)) or only through the processes z_0, z and V (see ($\hat{\mathbf{P}}$)).

In this chapter we take into account the case of a complete market with a CRRA investor, i.e., we consider the CRRA (Constant Relative Risk Aversion) utility function, and we solve the control problems ($\bar{\mathbf{P}}$) and ($\hat{\mathbf{P}}$) taking into account a particular factorization for the corresponding value function.

Finally, in the last section we present a specific model which allows us to compute the exact amount of wealth that must be allocated to the financial assets. In particular we present a model where the stochastic mortality $\lambda(t)$ dependent on the interest rate $r(t)$, and we take as assets traded on the market a rolling bond, a T-zero coupon longevity and a stock, where T is a suitable deterministic time such that, on the basis of demographic considerations, at time T the investor will be dead.

We refer to Fleming and Soner [12] for the optimal portfolio and (stochastic) dynamic programming theory, and to Rutkowski [20] for the case of a market with rolling bonds.

6.2 The Optimal Consumption and Portfolio for ($\bar{\mathbf{P}}$) in BLS market

Let (5.65) be the value function for the control problem ($\bar{\mathbf{P}}$), i.e.,

$$J(t_0, Z) = \sup_{u^A, c\, \in \bar{\mathfrak{U}}_{ad}} E^P \left[\int_{t_0}^{\infty} e^{-\rho\, t} z_0(t) U\big(V(t)c(t)\big) dt \Big| Z(t_0) = Z \right], \tag{6.1}$$

where $Z(t) = (z_0(t), r(t), \lambda(t), V(t), S(t))'$ with dynamics given by (5.60)-(5.64). According to the results of Section 5.3.1, now we solve the partial differential equation (5.76). Let the utility function be the CRRA function, i.e.,

$$U(C) = \frac{1}{1-\delta} C^{1-\delta}, \tag{6.2}$$

with $\delta > 1$, so that

$$\dot{U}^{-1}(y) = y^{-\frac{1}{\delta}} \quad \text{and} \quad U\left(\dot{U}^{-1}(y)\right) = \frac{1}{1-\delta} y^{1-\frac{1}{\delta}}, \tag{6.3}$$

then the partial differential equation (5.76) becomes

$$\frac{\partial H}{\partial t}(t, Z) + \frac{\delta}{1-\delta} e^{-\frac{\rho}{\delta} t} z_0^{\frac{1}{\delta}} \left(\frac{\partial H}{\partial V}(t, Z) \right)^{1-\frac{1}{\delta}} + \mathcal{L}^{\bar{u}^A_{sup}(\cdot;H), \bar{c}_{sup}(\cdot;H)} H(t, Z). \tag{6.4}$$

On the other hand, substituting the CRRA function given by (6.2) in (6.1) we obtain

$$J(t_0, Z) = \frac{1}{1-\delta} \sup_{u^A, c\, \in \bar{\mathfrak{U}}_{ad}} E^P \left[\int_{t_0}^{\infty} e^{-\rho\, t} z_0(t) \big(V(t)c(t)\big)^{1-\delta} dt \Big| Z(t_0) = Z \right], \tag{6.5}$$

and using the explicit form of $V(t)$ given by (5.29), we have that

$$J(t_0, z_0, z, V, S) = \frac{V^{1-\delta}}{1-\delta} F^{\delta}(t_0, z_0, z, V, S) \tag{6.6}$$

where

$$\begin{aligned} &F^{\delta}(t_0, z_0, z, V, S) \\ &= \sup_{u^A, c \in \bar{\mathfrak{U}}_{ad}} E^P_{z_0,z,S,V} \left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t) c(t)^{1-\delta} e^{(1-\delta)\int_{t_0}^t \left(r(s)+u^A(s)M(s)-c(s)-\frac{1}{2}|u^A(s)\Sigma^A(s)|^2 ds\right) ds} \right. \\ &\qquad \left. e^{(1-\delta)\int_{t_0}^t u^A(s)\Sigma^A(s) dW(s)} dt \right] \end{aligned}$$

with $M = \mu^A - r\mathbf{1} - \lambda \mathbf{1}_\lambda$, and $c(t)$, $u^A(t)$, $\mu^A(t)$ and $\Sigma^A(t)$ given by (5.45), (5.46), (5.10) and (5.11), respectively.

Now we may proceed to solve directly the above problem, but we observe that the function $F(t_0, z_0, z, V, S)$ has to be determined in order that $\bar{J}(t_0, z_0, z, V, S)$ satisfies the Hamilton-Jacobi-Bellman equation (6.4). To this end we need to calculate the following derivatives[1]

$$J_t = J\delta \frac{F_t}{F}, \tag{6.7}$$

$$J_V = J\left(\frac{1-\delta}{V} + \delta\frac{F_V}{F}\right), \quad J_{VV} = J\left[\left(\frac{1-\delta}{V} + \delta\frac{F_V}{F}\right)^2 - \frac{1-\delta}{V^2} + \delta\frac{F_{VV}F - F_V^2}{F^2}\right], \tag{6.8}$$

and, denoting $(z_0, r, \lambda, V, S)') = (z_0, z_1, z_2, z_3, z_4)'$ (similarly to the notations (5.59)), for $i \neq 3$ (i.e.$z_i \neq V$),

$$J_{z_i} = J\delta\frac{F_{z_i}}{F}, \quad J_{z_i z_j} = J\left(\delta^2\frac{F_{z_i}F_{z_j}}{F^2} + \delta\frac{F_{z_i z_j} - F_{z_i}F_{z_j}}{F^2}\right) \tag{6.9}$$

$$J_{V z_i} = J\left[\delta\frac{F_{z_i}}{F}\left(\frac{1-\delta}{V} + \delta\frac{F_V}{F}\right) + \delta\frac{F_{V z_i}F - F_V F_{z_i}}{F^2}\right]. \tag{6.10}$$

Substituting the value function $J(t_0, z_0, z, V, S)$ and its partial derivatives in Hamilton-Jacobi-Bellman equation (6.4), we obtain that the function $F(t_0, z_0, z, V, S)$ also solves a partial differential equation, but in general we do not know how to solve it explicitly. Nevertheless we can find a solution independent of V, i.e.,

$$F(t_0, z_0, z, V, S) = \check{F}(t_0, z_0, z, S). \tag{6.11}$$

With a little abuse of notation in the sequel we will continue to denote $\check{F}$ as F.

If (6.11) holds, then

$$J(t_0, z_0, z, V, S) = \frac{V^{1-\delta}}{1-\delta} F^{\delta}(t_0, z_0, z, S) \tag{6.12}$$

is our candidate value function for the optimal Markov control problem.

Furthermore, since the market is complete, (see Assumption 5.2.1) the function $F(t_0, z_0, z, S)$ can be represented through the Feynman-Kač theorem as shown in the following proposition.

Proposition 6.2.1. *If the market is arbitrage free and complete, then the function F in (6.12) has the representation*

$$F(t_0, z_0, z, S) = E^{\bar{Q}}_{t_0,z_0,z,S}\left(\int_{t_0}^{\infty} e^{-\frac{\rho}{\delta}s} z_0^{\frac{1}{\delta}}(s) e^{-\frac{\delta-1}{\delta}\int_{t_0}^s \left(r(u)+\frac{1}{2\delta}\xi'(u)\xi(u)\right)du} ds\right), \tag{6.13}$$

where the measure $\bar{Q}$ and the subscripts t_0, z_0, z and S denote that the expectation are taken using the following dynamics

$$dz_0(s) = -\lambda(s) z_0(s) ds \tag{6.14}$$

$$dz(s) = \left(\mu^z(s) + \frac{1-\delta}{\delta}\Pi(s)\xi(s)\right) ds + \Pi(s) dW^{\bar{Q}}(s) \tag{6.15}$$

$$dS(s) = S(s)\left[\left(r(s) + \sigma^S_r(s)\xi_r(s) + \sigma^S_S(s)\xi_S(s) + \frac{1-\delta}{\delta}\sigma^S(s)\xi(s)\right) ds + \sigma^S(s) dW^{\bar{Q}}(s)\right] \tag{6.16}$$

with the initial conditions given by $z_0(t_0) = z_0$, $z(t_0) = z$, and $S(t_0) = S$.

[1] For the notational convenience, the argument (t_0, z_0, z, V, S) “have been suppressed“ so that we have used the shorthand notation of the form

$$J = J(t_0, z_0, z, V, S), \quad F = F(t_0, z_0, z, V, S)$$

and similarly for the partial derivatives terms.

Proof. As already discussed, the function $F(t_0, z_0, z, S)$ must solve the partial differential equation (6.4). Considering the value function $J(t_0, z_0, z, V, S)$ given by (6.12), we have that the partial derivative in (6.8) and (6.10) becomes

$$J_V = J\frac{1-\delta}{V}, \quad J_{VV} = J\delta\frac{\delta-1}{V^2}, \tag{6.17}$$

and for $z_i \neq V$

$$J_{Vz_i} = J\frac{\delta(1-\delta)}{V}\frac{F_{z_i}}{F}. \tag{6.18}$$

For the notational convenience, in the sequel we denote t_0 by t. After substituting (6.7), (6.9), (6.17) and (6.18) into (6.4) we obtain

$$\begin{aligned}
&F_t + e^{-\frac{\rho}{\delta}t} z_0^{\frac{1}{\delta}} - z_0\lambda F_{z_0} + (\mu_z' + \frac{1-\delta}{\delta}M'Q^{-1}\Sigma^A\Pi')F_z + \frac{1-\delta}{\delta}\left(r + \frac{1}{2\delta}M'Q^{-1}M\right)F \\
&+ \frac{1}{2}tr[\Pi\,\Pi' F_{zz}] + \frac{1-\delta}{2}\frac{1}{F}F_z'\Pi(\Sigma^{A'}Q^{-1}\Sigma^A - I)\Pi' F_z + \frac{1-\delta}{2}\frac{1}{F}S^2\,F_S^2\sigma^S(\Sigma^{A'}Q^{-1}\Sigma^A - I)\sigma^{S'} \\
&+ S(r + \sigma_r^S\xi_r + \sigma_S^S\xi_S + \frac{1-\delta}{\delta}M'Q^{-1}\Sigma^A\sigma^{S'})F_S + \frac{1}{2}S^2\left|\sigma^S\right|^2 F_{SS} + S(\sigma_r^S\sigma^r)F_{rS} \\
&+ (1-\delta)S\frac{F_S}{F}F_z'\Pi(\Sigma^{A\prime}Q^{-1}\Sigma^A - I)\sigma^{S\prime} = 0,
\end{aligned} \tag{6.19}$$

where we have already simplified the common term $\delta\,\frac{J}{F}$.

By the arbitrage free and complete market assumptions, we have that[2]

$$\Sigma^{A\prime}Q^{-1}\Sigma^A = I, \qquad \text{and} \qquad \Sigma^A\xi = M, \tag{6.20}$$

taking into account (5.90). Then we obtain

$$M'Q^{-1}\Sigma^A = \xi', \qquad \text{and} \qquad M'Q^{-1}M = \xi'\xi,$$

so that, taking into account that $\xi'\Pi' = (\Pi\xi)'$ and $\xi'(\sigma^S)' = (\sigma^S\xi)'$, the equation (6.19) becomes

$$\begin{aligned}
&F_t + e^{-\frac{\rho}{\delta}t} z_0^{\frac{1}{\delta}} - z_0\lambda F_{z_0} + (\mu_z' + \frac{1-\delta}{\delta}(\Pi\xi)')F_z + \frac{1-\delta}{\delta}\left(r + \frac{1}{2\delta}\xi'\xi\right)F + \frac{1}{2}tr[\Pi\,\Pi' F_{zz}] \\
&+ S(r + \sigma_r^S\xi_r + \sigma_S^S\xi_S + \frac{1-\delta}{\delta}(\sigma^S\xi)')F_S + \frac{1}{2}S^2\left|\sigma^S\right|^2 F_{SS} + S(\sigma_r^S\sigma^r)F_{rS} = 0.
\end{aligned} \tag{6.21}$$

Then by applying the Feynman-Kač representation to the function $F(t, z_0, z, S)$ satisfying the partial differential equation (6.21), we obtain the announced result.

□

In such a case we are able to compute a solution in a quasi-explicit form only when the financial market is complete. The results are shown in the following proposition. In the sequel for the notational convenience we denote $t_0 = t$.

Proposition 6.2.2. *Under the completeness Assumption 5.2.1, let F be the function given by (6.13). If $F \in C^{1,2}(\mathbb{R}_+, (0,1)\times\mathbb{R}^3)$, then $J = \frac{V^{1-\delta}}{1-\delta}F^\delta$ is the value function given by (6.5), and the corresponding optimal controls are*

$$c_{sup}(t) = \bar{c}_{sup}(t, z_0(t), z(t), S(t); F), \tag{6.22}$$
$$u_{sup}^B(t) = \bar{u}_{sup}^B(t, z_0(t), z(t), S(t); F), \tag{6.23}$$
$$u_{sup}^L(t) = \bar{u}_{sup}^L(t, z_0(t), z(t), S(t); F), \tag{6.24}$$
$$u_{sup}^S(t) = \bar{u}_{sup}^B(t, z_0(t), z(t), S(t); F), \tag{6.25}$$

[2] Recall that (see Theorem 1.4.6) if the market is arbitrage free then the market price of risk $\xi = (\xi_r, \xi_\lambda, \xi_S)'$ must verify $\Sigma^A\xi = \mu^A(t) - r(t)\mathbf{1} - \lambda(t)\mathbf{1}_\lambda = M$. Furthermore

$$Q = \Sigma^A(\Sigma^A)',$$

and the arbitrage free market is complete if and only if the matrix Σ^A is invertible (see Corollary 1.4.5).

where[3]

$$\bar{c}_{sup}(t,z_0,z,S;F)=\frac{z_0^{\frac{1}{\delta}}\,e^{-\frac{\rho}{\delta}t_0}}{F}, \tag{6.26}$$

$$\bar{u}^{B}_{sup}(t,z_0,z,S;F)=\frac{1}{\delta}\frac{\xi_r\sigma^S_S-\sigma^S_r}{\frac{B^T_r}{B^T}\,\sigma^r\,\sigma^S_S}+\frac{\frac{F_r}{F}}{\frac{B^T_r}{B^T}}-\frac{\frac{L^T_r}{L^T}}{\frac{B^T_r}{B^T}}\left(\frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\,\sigma^\lambda}+\frac{\frac{F_\lambda}{F}}{\frac{L^T_\lambda}{L^T}}\right), \tag{6.27}$$

$$\bar{u}^{L}_{sup}(t,z_0,z,S;F)=\frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\,\sigma^\lambda}+\frac{\frac{F_\lambda}{F}}{\frac{L^T_\lambda}{L^T}}, \tag{6.28}$$

$$\bar{u}^{S}_{sup}(t,z_0,z,S;F)=\frac{1}{\delta}\frac{\xi_S}{\sigma^S_S}+S\frac{F_S}{F}, \tag{6.29}$$

and $z_0(t)$, $z(t)$ and $S(t)$ are solutions of (5.50), (5.51) and (5.52) with initial conditions (5.55). Furthermore the function $F(t,z_0,z,S)$ is given by (6.13) with $z_0(s)$, $z(s)$ and $S(s)$ solving (6.14), (6.15) and (6.16), respectively.

Observe that if we consider the arguments (t,z_0,z,S) instead of $(t,z_0(t),z(t),S(t))$ in above expression, then we write $\hat{\xi}^r$, $\hat{\xi}^\lambda$, $\hat{\xi}^S$ instead of ξ^r, ξ^λ, ξ^S respectively, and similarly for the other terms. Observe that the optimal consumption and portfolio $\bar{c}_{sup}$ and $\bar{u}^A_{sup}$ not depend on V, then we can conclude that the optimal consumption and portfolio for the CRRA investor not depend on the value of the portfolio.

Proof. Taking into account (6.3) and (6.12), and substituting the partial derivative (6.7), (6.9), (6.17) and (6.18) into (5.72)-(5.75), by Proposition 5.3.5 we have that

$$\bar{c}_{sup}(t,z_0,z,S;F)=\frac{1}{V}\dot{U}^{-1}\left(\frac{J_V}{e^{-\rho t}z_0}\right)=\frac{1}{V}z_0^{\frac{1}{\delta}}e^{-\frac{\rho}{\delta}t}\left(\frac{J(1-\delta)}{V}\right)^{-\frac{1}{\delta}}=\frac{1}{F}z_0^{\frac{1}{\delta}}e^{-\frac{\rho}{\delta}t}, \tag{6.30}$$

and the compact form of (6.27)-(6.29)

$$\begin{aligned}\left(\bar{u}^A_{sup}\right)'(t,z_0,z,S;F)&=-\frac{J_V}{V\,J_{VV}}Q^{-1}M-\frac{1}{V\,J_{VV}}Q^{-1}\Sigma^A\Pi'\,J_{zV}-\frac{S\,J_{VS}}{V\,J_{VV}}Q^{-1}\Sigma^A\sigma^{S\prime}\\&=\frac{1}{\delta}Q^{-1}M+\frac{1}{F}Q^{-1}\Sigma^A\Pi'F_z+\frac{S}{F}Q^{-1}\Sigma^A\sigma^{S\prime}F_S,\end{aligned} \tag{6.31}$$

where $Q=\Sigma^A\Sigma^{A\prime}$ and $J_{zV}=(J_{rV},J_{\lambda V})'$. Finally, we can write the expression (6.31) in explicit form substituting the partial derivative (6.7), (6.9), (6.17) and (6.18) into and (5.73), (5.74) and (5.75), so that, after some reshuffling, we obtain (6.27)-(6.29). Finally, since $F\in C^{1,2}(\mathbb{R}_+,(0,1)\times\mathbb{R}^3)$, by the Verification Theorem (see Theorem D.1.2) we obtain the announced results.

□

Remark 6.2.1. *Recall that we search the optimal consumption-investment strategy corresponding to the problem with value function $J(t,z,S,V)$ defined in (5.43). The value function $J(t,z,S,V)$ is obtained by computing in $z_0=1$ the value function $\bar{J}(t,z_0,z,S,V)$ defined in (5.49) (see (5.91)). It is interesting to observe that in the CRRA case the optimal consumption $\bar{c}_{sup}$ and optimal portfolio $\bar{u}^A_{sup}$ do not depend on the initial condition z_0, though the value function $\bar{J}$ depends explicitly on z_0: indeed on the one hand, for any fixed initial condition z_0, the representation (6.13) for F becomes*

$$\begin{aligned}F(t_0,z_0,z,S)&=z_0^{\frac{1}{\delta}}\,E^{\bar{Q}}_{t_0,z,S}\left(\int_{t_0}^{\infty}e^{-\frac{\rho}{\delta}s}e^{-\frac{1}{\delta}\int_{t_0}^{s}\lambda(u)du}\,e^{-\frac{\delta-1}{\delta}\int_{t_0}^{s}\left(r(u)+\frac{1}{2\delta}\xi'(u)\xi(u)\right)du}ds\right)\\&=z_0^{\frac{1}{\delta}}\,K(t_0,z,S),\end{aligned} \tag{6.32}$$

by Proposition 6.2.2, we obtain that $\bar{c}_{sup}(t,z_0,z,S;F)=\bar{\bar{c}}_{sup}(t,z,S;K)$ and $\bar{u}^A_{sup}(t,z_0,z,S;F)=\bar{\bar{u}}^A_{sup}(t,z,S;K)$ where K is the defined in (6.32), and

[3]For the notational convenience, the arguments $(t,z_0(t),z(t),S(t))$ "have been suppressed" and we we used the shorthand notation of the form

$$F=F(t,z_0,z,S),$$

and similarly for the partial derivatives.

$$\bar{\bar{c}}_{sup}(t,z,S;K) = \frac{1}{E^{\bar{Q}}_{t_0,z,S}\left(\int_{t_0}^{\infty} e^{-\frac{\rho}{\delta}(s-t_0)} e^{-\frac{1}{\delta}\int_{t_0}^{s}\lambda(u)du} e^{-\frac{\delta-1}{\delta}\int_{t_0}^{s}\left(r(u)+\frac{1}{2\delta}\xi'(u)\xi(u)\right)du} ds\right)}, \tag{6.33}$$

$$\bar{\bar{u}}^{B}_{sup}(t,z,S;K) = \frac{1}{\delta}\frac{\xi_r\sigma^{r,S}_S - \sigma^{r,S}_r}{\frac{B^T_r}{B^T}\sigma^r\sigma^{r,S}_S} + \frac{\frac{K_r}{K}}{\frac{B^T_r}{B^T}} - \frac{\frac{L^T_r}{L^T}}{\frac{B^T_r}{B^T}}\left(\frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\sigma^\lambda} + \frac{\frac{K_\lambda}{K}}{\frac{L^T_\lambda}{L^T}}\right) \tag{6.34}$$

$$\bar{\bar{u}}^{L}_{sup}(t,z,S;K) = \frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\sigma^\lambda} + \frac{\frac{K_\lambda}{K}}{\frac{L^T_\lambda}{L^T}} \tag{6.35}$$

$$\bar{\bar{u}}^{S}_{sup}(t,z,S;K) = \frac{1}{\delta}\frac{\xi_S}{\sigma^{r,S}_S} + S\frac{K_S}{K}\quad . \tag{6.36}$$

On the the other hand, we get that $\bar{J}(t,z_0,z,S,V) = \frac{V^{1-\delta}}{1-\delta}\, z_0\, K^\delta(t_0,z,S)$; and therefore, for $z_0 = 1$, we can rewrite (6.12) as

$$J(t_0,z,V,S) = \frac{V^{1-\delta}}{1-\delta}F^\delta(t_0,z,S) \tag{6.37}$$

where F coincides with K and has the representation

$$F(t_0,z,S) = E^{\bar{Q}}_{t_0,z,S}\left(\int_{t_0}^{\infty} e^{-\frac{\rho}{\delta}s} e^{-\frac{1}{\delta}\int_{t_0}^{s}\lambda(u)du} e^{-\frac{\delta-1}{\delta}\int_{t_0}^{s}\left(r(u)+\frac{1}{2\delta}\xi'(u)\xi(u)\right)du} ds\right). \tag{6.38}$$

6.2.1 The Optimal Consumption and Portfolio for $(\hat{\mathbf{P}})$

In this section we follow the same steps for the problem $(\bar{\mathbf{P}})$ and here we obtain the similar results. Recall that the difference between the problem $(\bar{\mathbf{P}})$ and $(\hat{\mathbf{P}})$ is that the latter concern control processes such that not depend directly on the dynamics of the risk asset, but only through the processes z_0, z and V. Let J be the value function for the control problem $(\bar{\mathbf{P}})$, i.e.,

$$J(t_0,\check{Z}) = \mathbf{1}_{\{\tau>t_0\}} \sup_{\hat{u}^A,\hat{c}\in\hat{\mathcal{U}}_{ad}} E^P\left[\int_{t_0}^{\infty} e^{-\rho t} z_0(t)U\big(V(t)c(t)\big)dt \,\Big|\, \check{Z}(t_0) = \check{Z}\right] \tag{6.39}$$

where $\check{Z}(t) = (z_0(t), r(t), \lambda(t), V(t))'$ with dynamics given by (5.108)-(5.112). According to the results of Section 5.3.2, now we solve the partial differential equation (5.114). Let the CRRA utility function given by (6.2), and taking into account (6.3), we have that the partial differential equation (5.114) becomes

$$\frac{\partial J}{\partial t}(t,\check{Z}) + \frac{\delta}{1-\delta}e^{-\frac{\rho}{\delta}t}z_0^{\frac{1}{\delta}}\left(\frac{\partial J}{\partial V}(t,\check{Z})\right)^{1-\frac{1}{\delta}} + \mathcal{L}^{\hat{u}^A_{sup},\hat{c}_{sup}}J(t,\check{Z}). \tag{6.40}$$

Using the explicit form of $V(t)$ given by (5.29), we have that the value function J is given by

$$J(t_0,z_0,z,V) = \frac{V^{1-\delta}}{1-\delta}F^\delta(t_0,z_0,z,V) \tag{6.41}$$

where

$$\begin{aligned} &F^\delta(t_0,z_0,z,V)\\ &= \sup_{\hat{u}^A,\hat{c}\in\underline{\mathcal{U}}_{ad}} E^P_{z_0,z,V}\Big[\int_{t_0}^{\infty} e^{-\rho t}\, z_0(t)c(t)^{1-\delta} e^{(1-\delta)\int_{t_0}^{t}\left(r(s)+u^A(s)M(s)-c(s)-\frac{1}{2}|u^A(s)\Sigma^A(s)|^2 ds\right)ds}\\ &\qquad\qquad e^{(1-\delta)\int_{t_0}^{t}u^A(s)\Sigma^A(s)dW(s)}\,dt\Big] \end{aligned}$$

with $M = \mu^A - r\mathbf{1} - \lambda\mathbf{1}_\lambda$, and $u^A(t)$, $c(t)$, $\mu^A(t)$ and $\Sigma^A(t)$ given by (5.99) and (5.100), (5.97) and (5.98), respectively.

Exactly as in the previous section, substituting the value function $J(t_0,z_0,z,V)$ and its partial derivatives[4] in Hamilton-Jacobi-Bellman equation (6.40), we obtain that the function $F(t_0,z_0,z,V)$ also solves a partial

[4]In this setting the partial derivatives of $J(t_0,z_0,z,V)$ are given by (6.7), (6.8), (6.9) and (6.10) taking into account that J and F are a functions depend on (t_0,z_0,z,V). In particular (6.9) and (6.10) are valid for $i = 0,1,2,3$.

differential equation, but in general we do not know how to solve it explicitly. Nevertheless we can find a solution independent of V, i.e.,

$$F(t_0, z_0, z, V) = \check{F}(t_0, z_0, z). \tag{6.42}$$

With a little abuse of notation in the sequel we will continue to denote $\check{F}$ as F.

If (6.42) holds, then

$$J(t_0, z_0, z, V) = \frac{V^{1-\delta}}{1-\delta} F^{\delta}(t_0, z_0, z) \tag{6.43}$$

is our candidate value function for the optimal Markov control problem.

Furthermore, since the market is complete, (see Assumption 5.2.1), analogously to the Proposition 6.2.1, the function $F(t_0, z_0, z)$ can be represented through the Feynman-Kač theorem as shown in the following proposition.

Proposition 6.2.3. *If the market is arbitrage free and complete, then the function F in (6.43) has the representation*

$$F(t_0, z_0, z) = E^{\bar{Q}}_{t_0, z_0, z}\left(\int_{t_0}^{\infty} e^{-\frac{\rho}{\delta}s} z_0^{\frac{1}{\delta}}(s) e^{-\frac{\delta-1}{\delta}\int_{t_0}^{s}\left(r(u)+\frac{1}{2\delta}\xi'(u)\xi(u)\right)du} ds \right) \tag{6.44}$$

where the measure $\bar{Q}$ and the subscripts t_0, z_0, and z denote that the expectation are taken using the following dynamics

$$dz_0(s) = -\lambda(s) z_0(s) ds \tag{6.45}$$

$$dz(s) = \left(\mu^z(s) + \frac{1-\delta}{\delta}\Pi(s)\xi(s) \right) ds + \Pi(s) dW^{\bar{Q}}(s) \tag{6.46}$$

$$dS(s) = S(s)\left[\left(r(s) + \sigma^S_r \xi_r(s) + \sigma^S_S \xi_S(s) + \frac{1-\delta}{\delta}\sigma^S_S \xi(s) \right) ds + \sigma^S(s) dW^{\bar{Q}}(s) \right] \tag{6.47}$$

with the initial conditions given by $z_0(t_0) = z_0$, and $z(t_0) = z$.

Proof. The proof follow immediately by Proposition 6.2.1. Indeed, taking into account that in this setting the partial derivative of $J(t_0, z_0, z, V)$ respect to S are null, the equation (6.21) becomes[5]

$$F_t + e^{-\frac{\rho}{\delta}t} z_0^{\frac{1}{\delta}} - z_0 \lambda F_{z_0} + (\mu_z' + \frac{1-\delta}{\delta}\xi'\Pi')F_z + \frac{1-\delta}{\delta}\left(r + \frac{1}{2\delta}\xi'\xi \right) F + \frac{1}{2} tr[\Pi\,\Pi' F_{zz}] \tag{6.48}$$

Then by applying the Feynman-Kač representation to the function $F(t_0, z_0, z)$ such that satisfies the partial differential equation (6.48), we obtain the announced result.

□

So again we are able to compute a solution in a quasi-explicit form only when the financial market is complete. The results are shown in the following proposition analogous to Proposition 6.2.2.

Proposition 6.2.4. *Under the completeness Assumption 5.2.1 if the function F in (6.44) is $C^{1,2}(\mathbb{R}_+, \mathbb{R}^3)$, then, for the Markov control problem ($\hat{\mathbf{P}}$), the optimal consumption and portfolio are given by*

$$\hat{c}_{sup} = \frac{z_0^{\frac{1}{\delta}}\, e^{-\frac{\rho}{\delta}t_0}}{F}, \tag{6.49}$$

$$\hat{u}^B_{sup} = \frac{1}{\delta}\frac{\xi_r \sigma^{r,S}_S - \sigma^{r,S}_r}{\frac{B^T_r}{B^T}\sigma^r \sigma^{r,S}_S} + \frac{\frac{F_r}{F}}{\frac{B^T_r}{B^T}} - \frac{\frac{L^T_r}{L^T}}{\frac{B^T_r}{B^T}}\left(\frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\sigma^\lambda} + \frac{\frac{F_\lambda}{F}}{\frac{L^T_\lambda}{L^T}} \right) \tag{6.50}$$

$$\hat{u}^L_{sup} = \frac{1}{\delta}\frac{\xi_\lambda}{\frac{L^T_\lambda}{L^T}\sigma^\lambda} + \frac{\frac{F_\lambda}{F}}{\frac{L^T_\lambda}{L^T}} \tag{6.51}$$

$$\hat{u}^S_{sup} = \frac{1}{\delta}\frac{\xi_S}{\sigma^{r,S}_S}\ , \tag{6.52}$$

where for the notational convenience, the arguments $(t, z_0(t), z(t))$ "have been suppressed" and we we used the short notations previously introduced. Furthermore the $F(t_0, z_0, z)$ is given by (6.44) with $z_0(s)$, $z(s)$ and $S(s)$ solving (6.45), and (6.46), respectively.

[5] For the notational convenience, in the sequel we denote t_0 by t.

Observe that if we consider the arguments (t_0, z_0, z) instead of $(t, z_0(t), z(t))$ in above expression, then we write $\hat{\xi}^r$,$\hat{\xi}^\lambda$, $\hat{\xi}^s$ instead of ξ^r ,ξ^λ, ξ^s respectively, and similarly for the other terms. Observe that the optimal consumption and portfolio $\hat{c}_{sup}$ and $\hat{u}^A_{sup}$ not depend on V, then we can conclude that the optimal consumption and portfolio for the CRRA investor not depend on the value of the portfolio.

Proof. The proof follow immediately by Proposition 6.2.4 taking into account that in this setting the partial derivative of $F(t_0, z_0, z, V)$ respect to S are null.

□

Observe that in this section we have the same results obtained by Menoncin [18] as a particular case of the general problem $(\bar{\mathbf{P}})$, where in the latter case the risk asset $S(t)$ may not be a Black-Scholes model.

6.3 The Optimal Consumption and Portfolio for $(\bar{\mathbf{P}})$ in UOS market

6.4 A specific market model

The results shown in the previous section are quite general and characterize different models for the financial market and state variables. This section introduces a slight modification of the model for the financial market of Chapter 5, i.e., the classic bond-stock market with a longevity bond (in the sequel shortly denoted as RLS market model).

The key point is that we have chosen to take the model where the interest rate $r(t)$ and the stochastic mortality intensity $\lambda(t)$ are dependent, but with uncorrelated driving noises. In particular we take as assets traded on the RLS market a rolling bond, a T-zero coupon longevity bond and a stock, where T is a suitable deterministic time such that, on the basis of demographic considerations, the investor will be dead at time T with probability 1.

The model can be specified as follows.

1. We take as reference model for the interest rate $r(t)$ the Cox-Ingersoll-Ross (CIR) model, given by

$$\begin{cases} dr(t) = a_r\,(b_r - r(t))\,dt + \bar{\sigma}_r\sqrt{r(t)}dW^r(t), \\ r(t_0) = r \end{cases} \tag{6.53}$$

   where W^r is a 1-dimensional Wiener process, and a_r, b_r, $\bar{\sigma}_r$ and r are strictly positive deterministic constants such that $2\,a_r\,b_r > \bar{\sigma}_r^2$, so that the process $r(t)$ remains strictly positive[6].

2. We set the stochastic mortality intensity $\lambda(t)$ as

$$\lambda(t) = \frac{\lambda^{(c)}(t) + D}{T - t}, \tag{6.54}$$

   where D is a positive constant and the process $\lambda^{(c)}(t)$ is given by (3.51), so that, under P, the process $\lambda(t)$ satisfies

$$\begin{cases} d\lambda(t) = \left(\frac{1}{T-t} - a_\lambda\right)\lambda(t)dt + \frac{a_\lambda\left(b_\lambda + cr(t) + D\right)}{T-t}dt + \frac{1}{\sqrt{T-t}}\bar{\sigma}_\lambda\sqrt{\lambda(t) - \frac{D}{T-t}}\,dW^\lambda(t), \\ \lambda(t_0) = \lambda \end{cases} \tag{6.55}$$

   where W^λ is a 1-dimensional Wiener process independent of W^r, and a_λ, b_λ, $\bar{\sigma}_\lambda$, and c are strictly positive deterministic constants such that $2\,a_\lambda\,b_\lambda > \bar{\sigma}_\lambda^2$. Observe that the stochastic mortality intensity $\lambda(t)$ is not assumed to be independent of r since $\lambda^{(c)}(t)$ depend on $r(t)$ and $\lambda(t)$ is strictly positive. Furthermore, assuming that the death time τ is the first jump time of a doubly stochastic Poisson process with intensity $\lambda(t)$, we get $\tau \leq T$ a.s., as should be in this setting.

3. We fix a time $T_B \leq T - t_0$ (in the applications we take $T_B = 25$ years) and take a rolling bond $R(t, T_B)$, i.e., a self-financing strategy that involves holding at any time one unit of a T_B-sliding bond. We recall (see, e.g., Rutkowski [20]) that the price of a T_B-sliding bond is $B(t, T_B + t)$, the price at time t of a

[6] As shown in Shreve [21].

$T_B + t$-bond. Moreover taking the market price of risk $\hat{\xi}_r(t,r(t)) = \bar{\xi}_r\sqrt{r(t)}$, the value of these bonds evolves according to[7]

$$\begin{aligned}\frac{dR(t,T_B)}{R(t,T_B)} &= \frac{d\hat{R}(t,r(t);T_B)}{\hat{R}(t,r(t);T_B)}\\ &= C_0\Big(f(t,t+T_B) - a_r b_r \psi_{r,\xi_r}(T_B) + r(t)\Big(a_r\psi_{r,\xi_r}(T_B) + \frac{\bar{\sigma}_r^2}{2}\frac{1-e^{\alpha_r T_B}}{\beta_r + \gamma_r e^{\alpha_r T_B}}\Big)\Big)dt\\ &+ \psi_{r,\xi_r}(T_B)\bar{\sigma}_r\sqrt{r(t)}\,dW^r(t) \end{aligned} \tag{6.56}$$

under P, while

$$\begin{aligned}\frac{dR(t,T_B)}{R(t,T_B)} &= \frac{d\hat{R}(t,r(t);T_B)}{\hat{R}(t,r(t);T_B)}\\ &= C_0\Big(f(t,t+T_B) + a_r b_r \psi_{r,\xi_r}(T_B) + r(t)\Big((a_r - \bar{\sigma}_r\bar{\xi}_r)\psi_{r,\xi_r}(T_B) + \frac{\bar{\sigma}_r^2}{2}\psi_{r,\xi_r}(T_B)\Big)\Big)dt\\ &+ \bar{\sigma}_r\sqrt{r(t)}\psi_{r,\xi_r}(T_B)dW^r(t) \end{aligned} \tag{6.57}$$

under Q, where

$$\psi_{r,\xi_r}(T_B) = \frac{1-e^{\alpha_{r,\xi_r} T_B}}{\beta_{r,\xi_r} + \gamma_{r,\xi_r} e^{\alpha_{r,\xi_r} T_B}},$$

$$\alpha_{r,\xi_r} = -\sqrt{(a_r - \bar{\sigma}_r\bar{\xi}_r)^2 + 2\bar{\sigma}_r^2}, \quad \beta_{r,\xi_r} = \frac{\alpha_r - a_r + \bar{\sigma}_r\bar{\xi}_r}{2}, \quad \gamma_{r,\xi_r} = \frac{\alpha_r + a_r - \bar{\sigma}_r\bar{\xi}_r}{2},$$

and as usual, $f(t,T)$ stands for the instantaneous forward rate prevailing at time t for the future infinitesimal time period $[T, T+dT]$,

4. We take a T-longevity bond $L(t,T)$ given by (4.55). Fixed the maturity time T, and the market price of risk $\hat{\xi}_\lambda(t,\lambda(t)) = \bar{\xi}_\lambda\sqrt{\lambda(t) - \frac{D}{T-t}}$, and as above $\hat{\xi}_r(t,r(t)) = \bar{\xi}_r\sqrt{r(t)}$, the value of these bonds evolves according to (see Section 6.5.4)

$$\begin{aligned}\frac{dL(t,T)}{L(t,T)} &= \frac{d\hat{L}(t,z(t);T)}{\hat{L}(t,z(t);T)}\\ &= \Big(r(t) + \lambda(t) - \psi_{r,\xi_r}^{(c)}(t)\bar{\sigma}_r\bar{\xi}_r r(t) - \psi_\lambda^{(D)}(t)\frac{\bar{\xi}_\lambda\bar{\sigma}_\lambda}{\sqrt{T-t}}(\lambda(t) - \frac{D}{T-t})\Big)dt\\ &- \psi_{r,\xi_r}^{(c)}(t)\bar{\sigma}_r\sqrt{r(t)}\,dW^r(t) - \psi_\lambda^{(D)}(t)\frac{\bar{\sigma}_\lambda}{\sqrt{T-t}}\sqrt{\lambda(t) - \frac{D}{T-t}}\,dW^\lambda(t). \end{aligned} \tag{6.58}$$

where the functions $\psi_{r,\xi_r}^{(c)}(t)$ and $\psi_\lambda^{(D)}(t)$ satisfy the following differential equations

$$\dot{\psi}_{r,\xi_r}^{(c)}(t) = -a_r\psi_{r,\xi_r}^{(c)}(t) - \frac{\bar{\sigma}_r^2}{2}\Big(\psi_{r,\xi_r}^{(c)}\Big)^2(t) + \frac{c}{T-t}\psi_\lambda^{(D)}(t) + 1 \tag{6.59}$$

$$\dot{\psi}_\lambda^{(D)}(t) = (\frac{1}{T-t} - a_\lambda)\psi_\lambda^{(D)}(t) - \frac{\bar{\sigma}_\lambda^2}{2}(T-t)^2\Big(\psi_\lambda^{(D)}\Big)^2(t) + 1 \tag{6.60}$$

with the initial conditions $\psi_{r,\xi_r}^{(c)}(0) = 0$ and $\psi_\lambda^{(D)}(0) = 0$.

5. For the risky asset we take the Black and Scholes model, and, taking the market price of risk $\xi_S = \bar{\xi}_S$ the price process $S(t)$ is given by

$$\frac{dS(t)}{S(t)} = \Big(r(t) + \sigma_r^S\bar{\xi}_r\sqrt{r(t)} + \sigma_S^S\bar{\xi}_S\Big)\,dt + \sigma_r^S dW^r(t) + \sigma_S^S dW^S(t), \tag{6.61}$$

where $\bar{\xi}_S$, σ_r^S and σ_S^S are deterministic constants.

[7] See Section 6.5.3).

6.5 Properties of RLS market model

In this section we analyze some properties of the BLS market model briefly describe in the previous section, one of the main properties being that $z(t) = (r(t), \lambda(t))$ is an affine process[8].

6.5.1 CIR interest rate model

We have taken as reference model for the interest rate $r(t)$ the CIR process under the measure P. Moreover taking the market price of risk $\hat{\xi}_r(t, r(t)) = \bar{\xi}_r \sqrt{r(t)}$, the dynamics of $r(t)$ under the measure martingale Q are given by (see (2.42))

$$dr = \big(a_r b_r - (a_r - \bar{\sigma}_r \bar{\xi}_r) r(t)\big) dt + \bar{\sigma}_r \sqrt{r(t)}\, d\bar{W}^r(t) \tag{6.62}$$

i.e., $r(t)$ is still a CIR model under Q.

6.5.2 Mortality intensity model

In this setting, the time T is a suitable deterministic time such that, on the basis of demographic considerations, at time T the investor will be dead. Modelling the death time τ as the first jump time of a doubly stochastic Poisson process with intensity $\lambda(t)$, we have that $\tau \leq T$ a.s., as shown in the following proposition.

Proposition 6.5.1. *Let $\lambda(t)$ be the process with dynamics given by (6.55), with a_λ, b_λ, c and $\bar{\sigma}_\lambda$ strictly positive deterministic constants such that $2\, a_\lambda\, b_\lambda > \bar{\sigma}_\lambda^2$. Then*

$$\lambda(t) > 0 \qquad a.s.$$

Furthermore the first jump time τ of a doubly stochastic Poisson process with intensity $\lambda(t)$ is such that

$$P\big(\tau < T\big) = 1. \tag{6.63}$$

Proof of Proposition 6.5.1. Since $\lambda^{(c)}(t)$ is strictly positive (see Proposition 3.4.2) and $D > 0$, we have

$$\lambda(t) = \frac{\lambda^{(c)}(t) + D}{T - t} \geq \frac{D}{T - t} > 0, \quad \forall t \leq T \tag{6.64}$$

i.e., $\lambda(t)$ is strictly positive. Furthermore we have that the mortality intensity $\lambda(t)$ satisfies the following property

$$P\big(\int_{t_0}^{T} \lambda(u) du = \infty\big) = 1. \tag{6.65}$$

Indeed, by (6.64) we have that

$$\int_{t_0}^{T} \lambda(t) dt \geq \int_{t_0}^{T} \frac{D}{T - t} dt = +\infty.$$

Now we show that τ satisfies (6.63). By the relation (3.38) with $T_1 = \tau$,i.e.,

$$P\big(\tau > T | \mathcal{G} \vee \mathcal{F}_s^N\big) = \mathbf{1}_{\tau > s} e^{-\int_s^T \lambda(u) du}, \tag{6.66}$$

we have that for each $0 \leq t_0 \leq T$

$$P(\tau > T) = E\Big(E\Big(\mathbf{1}_{\tau > T}\Big|\mathcal{G} \vee \mathcal{F}_{t_0}^N\Big)\Big) = \mathbf{1}_{\tau > t_0} E\Big(e^{-\int_{t_0}^T \lambda(u) du}\Big). \tag{6.67}$$

Since (6.65) holds, we obtain that

$$P(\tau < T) = 1 - P(\tau > T) = 1 - \mathbf{1}_{\tau > t_0} E\left(e^{-\int_s^T \lambda(u) du}\right) = 1. \tag{6.68}$$

□

[8] We recall that the convenience of adopting such processes is given by the key property of affine processes, i.e., the property (2.51) of Section 2.3 (see also Section 3.3).

Furthermore given the market price of risk $\hat{\xi}_\lambda(t, \lambda(t)) = (\bar{\xi}_\lambda \sqrt{\lambda(t) - \frac{D}{T-t}})$, under the martingale measure Q the dynamics of $\lambda(t)$ are given by

$$\begin{aligned} d\lambda(t) = &\Big(\frac{1}{T-t} - a_\lambda - \frac{\bar{\xi}_\lambda \bar{\sigma}_\lambda}{\sqrt{T-t}}\Big)\lambda(t)dt + \Big(\frac{\bar{\xi}_\lambda \bar{\sigma}_\lambda D}{\sqrt{(T-t)^3}} + \frac{a_\lambda\Big(b_\lambda + cr(t) + D\Big)}{T-t}\Big)dt \\ &+ \frac{1}{\sqrt{T-t}}\bar{\sigma}_\lambda \sqrt{\lambda(t) - \frac{D}{T-t}}\, d\bar{W}^\lambda(t). \end{aligned} \tag{6.69}$$

6.5.3 Rolling Bonds

In this setting we consider a self-financing strategy that involves holding at any time one unit of a sliding bond. The wealth process of this strategy is referred to as the rolling bond. In contrast to the sliding bond, which does not represent a tradable security in arbitrage-free market, the rolling bond may play the role of a security with infinite lifespan. In particular we take a constant time to maturity $T_B = 25$ years. By (2.52) with $T = T_B + t$, we have that

$$B(t, t + T_B) = \hat{B}(t, r, T_B + t) = e^{\psi^0_{r,\xi_r}(T_B) + \psi_{r,\xi_r}(T_B) r} \tag{6.70}$$

where[9]

$$\begin{aligned} \psi_{r,\xi_r}(T_B) &= \frac{1 - e^{\alpha_{r,\xi_r} T_B}}{\beta_{r,\xi_r} + \gamma_{r,\xi_r} e^{\alpha_{r,\xi_r} T_B}}, \\ \psi^0_{r,\xi_r}(T_B) &= -\frac{2\, a_r\, b_r}{\bar{\sigma}_r^2} ln\left(\frac{\beta_{r,\xi_r} + \gamma_{r,\xi_r} e^{\alpha_{r,\xi_r} z}}{\alpha_{r,\xi_r}}\right) + \frac{a_r\, b_r}{\beta_{r,\xi_r}} T_B \end{aligned}$$

with

$$\alpha_{r,\xi_r} = -\sqrt{(a_r - \bar{\sigma}_r \bar{\xi}_r)^2 + 2\bar{\sigma}_r^2}, \quad \beta_{r,\xi_r} = \frac{\alpha_{r,\xi_r} - a_r + \bar{\sigma}_r \bar{\xi}_r}{2}, \quad \gamma_{r,\xi_r} = \frac{\alpha_{r,\xi_r} + a_r - \bar{\sigma}_r \bar{\xi}_r}{2}. \tag{6.71}$$

Furthermore, taking into account (6.70) the price dynamics of $B(t, T_B + t)$ are given by [10]

$$\begin{aligned} \frac{dB(t, T_B + t)}{B(t, T_B + t)} &= \frac{d\hat{B}(t, r(t), T_B + t)}{\hat{B}(t, r(t), T_B + t)} \\ &= \psi_{r,\xi_r}(T_B) dr(t) + \frac{1}{2}\psi^2_{r,\xi_r}(T_B)\bar{\sigma}_r^2 r(t) dt. \end{aligned} \tag{6.72}$$

Therefore, taking into account (6.53) under P we get

$$\frac{dB(t, T_B + t)}{B(t, T_B + t)} = \psi_{r,\xi_r}(T_B)\Big(a_r\,(b_r - r(t))\,dt + \bar{\sigma}_r \sqrt{r(t)} dW^r(t)\Big) + \frac{1}{2}\psi^2_{r,\xi_r}(T_B)\bar{\sigma}_r^2 r(t) dt, \tag{6.73}$$

while taking into account (6.62) under Q we get

$$\frac{dB(t, T_B + t)}{B(t, T_B + t)} = \psi_{r,\xi_r}(T_B)\Big(\big(a_r b_r - (a_r - \bar{\sigma}_r \bar{\xi}_r) r(t)\big)dt + \bar{\sigma}_r \sqrt{r(t)}\, d\bar{W}^r(t)\Big) + \frac{1}{2}\psi^2_{r,\xi_r}(T_B)\bar{\sigma}_r^2 r(t) dt. \tag{6.74}$$

Moreover by Proposition 3.2 in Rutkowski [20] , we have that

$$R(t, T_B) = C_0 A(t, T_B) B(t, T_B + t), \tag{6.75}$$

where $C_0 = \frac{R(t_0, T_B)}{B(t_0, T_B)}$, $A(t, T_B) = e^{\int_{t_0}^t f(s, T_B + s) ds}$ and so that

$$\begin{aligned} \frac{dR(t, T_B)}{C_0} &= dA(t, T_B) B(t, T_B + t) + A(t, T_B) dB(t, T_B + t) \\ &= R(t, T_B)\Big[\big(f(t, t + T_B) + \frac{dB(t, T_B + t)}{B(t, T_B + t)}\Big]. \end{aligned} \tag{6.76}$$

[9] See Section 2.2.1.

[10] We observe that $\frac{\hat{B}_r}{\hat{B}}(t, r(t), T_B + t) = \psi_{r,\xi_r}(T_B)$.

6.5.4 Longevity Bonds

The aim of this section is the computation of

$$\hat{L}(t,r,\lambda;T) = E^Q_{t,z}\Big(e^{-\int_t^T r(s)ds}e^{-\int_t^T \lambda(s,r)ds}\Big). \tag{6.77}$$

Given the market price of risk $\big(\hat{\xi}_r(t,r(t)),\hat{\xi}_\lambda(t,\lambda(t))\big) = (\bar{\xi}_r\sqrt{r(t)}\,,\bar{\xi}_\lambda\sqrt{\lambda(t)-\frac{D}{T-t}}\,)$, under the martingale measure Q the dynamics of $r(t)$ are given by (6.62), while the dynamics of $\lambda(t)$ are given by (6.69). Similarly to the procedure of Section 4.3.1, we use

$$e^{\psi^0_z+m\psi^{(c)}_{r,\xi_r}(t)r+n\psi^{(D)}_\lambda(t)\lambda}$$

as guess function for

$$E^Q_{t,z}\Big(e^{-m\int_t^T r(s)ds}e^{-n\int_t^T \lambda(s,r)ds}\Big).$$

By Itô's formula and equation (6.69), together with Feynman-Kač representation formula, we obtain that $\psi^{(c)}_{r,\xi_r}(t)$ and $\psi^{(D)}_\lambda(t)$ satisfies (6.59) and (6.60). Thus

$$\hat{L}(t,r,\lambda;T) = e^{\psi^0_z+\psi^{(c)}_{r,\xi_r}(t)r+\psi^{(D)}_\lambda(t)\lambda},$$

and the dynamics follows by Itô's formula.

Chapter 7

The Optimal Portfolio: a numerical simulation of the CRRA utility case

7.1 Numerical Simulation of the interest rate and the stochastic intensity

To compute the price of derivatives with a Monte-Carlo algorithm, we need to simulate paths of $(r(t), \lambda(t))$ and the first difficulty lies in simulating a CIR process. It is well known that a standard Euler scheme can lead to negative values and then to complex values even if $2a_r b_r > \bar{\sigma}_r^2$, which we assume to hold (see Section 6.5.1). Following Brigo and Alfonsi [4] and [5] we present here briefly the implicit positivity-preserving Euler scheme for $r(t)$ and we extend this scheme to the case $z(t) = (r(t), \lambda(t))$. This approximation method ensures that the simulation preserves the positivity property also in this setting.

We start recalling the explicit Euler scheme for a generic autonomous stochastic differential equation given by

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t), \quad X(0) = X_0 \quad 0 \le t \le T \tag{7.1}$$

To apply a numerical method to (7.1) over $[0, T]$, we first discretize the interval with the standard time discretization $t_i = i\frac{T}{n}$, for $i = 0, \dots, n$. For computational purpose it is useful to consider discretized Wiener process, where $W(t)$ is specified at discrete t values. Let W_i denote $W(t_i)$. Then $W_0 = 0$ with probability 1, and

$$W_i = W_{i-1} + \Delta W_i, \quad i = 1, 2, \dots, n$$

where $\{\Delta W_i\}$ is a sequence of independent Gaussian random variables with zero mean and $\sqrt{t_{i+1} - t_i}$ variance. We recall that the (explicit) Eulero-Maruyama method takes the form

$$x^E(t_{i+1}) = \mu(x^E(t_i))\left(t_{i+1} - t_i\right) + \sigma(x^E(t_i))\left(W_{i+1} - W_i\right)$$

and, for $t_i < t < t_{i+1}$

$$\text{either} \qquad x^E(t) = x^E(t_i), \qquad \text{or} \qquad x^E(t) = x^E(t_i) + \frac{t - t_i}{t_{i+1} - t_i}\left(x^E(t_{i+1}) - x^E(t_i)\right)$$

and we known that it has strong order of convergence $\gamma = \frac{1}{2}$.

Brigo and Alfonsi show that the solution of (7.1) is also obtained with an implicit Eulero approximation, $x^E_{imp}(t)$, where as above for $t_i < t < t_{i+1}$,

$$\text{either} \qquad x^E_{imp}(t) = x^E_{imp}(t_i), \qquad \text{or} \qquad x^E_{imp}(t) = x^E_{imp}(t_i) + \frac{t - t_i}{t_{i+1} - t_i}\left(x^E_{imp}(t_{i+1}) - x^E_{imp}(t_i)\right),$$

where in this case $x^E_{imp}(t_{i+1})$ is define as the solution[1] of the following equation

$$\begin{aligned} & x^E_{imp}(t_{i+1}) \\ & = x^E_{imp}(t_i) + \mu(x^E_{imp}(t_{i+1}))\left(t_{i+1} - t_i\right) \\ & - \sigma(x^E_{imp}(t_{i+1}))\sigma'(x^E_{imp}(t_{i+1}))\left(t_{i+1} - t_i\right) + \sigma(x^E_{imp}(t_{i+1}))\left(W_{i+1} - W_i\right), \end{aligned} \tag{7.2}$$

[1]To understand the philosophy of the implicit Euler scheme, observe that if

$$x(t_{i+1}) = x(t_i) + \widetilde{\mu}(x(t_{i+1}))\left(t_{i+1} - t_i\right) + \sigma(x(t_{i+1}))\left(W_{i+1} - W_i\right),$$

i.e.,

$$\frac{\sigma\sigma'(x^E_{imp}(t_{i+1}))}{2}\big(t_{i+1}-t_i\big)+x^E_{imp}(t_{i+1})-\big(\mu(x^E_{imp}(t_{i+1}))\,(t_{i+1}-t_i)\big)$$
$$-\sigma(x^E_{imp}(t_{i+1}))\big(W_{i+1}-W_i\big)-x^E_{imp}(t_i)=0. \tag{7.3}$$

In the CIR model one can find a closed form solution $x^E_{imp}(t_{i+1})$, though generally this is not the case. Indeed, in the CIR model

$$dx(t)=a\big(b-x(t)\big)dt+\sigma\sqrt{x(t)}\,dW(t)$$

the diffusion coefficient $\sigma(x)=\sigma\sqrt{x}$, so that $\sigma'(x)=\frac{1}{2\sqrt{x}}\sigma$, then $\sigma(x)\,\sigma'(x)=\sigma\sqrt{x}\,\frac{1}{2\sqrt{x}}\sigma=\frac{1}{2}\sigma^2$. Thus (7.3) becomes

$$\frac{\sigma^2}{2}\big(t_{i+1}-t_i\big)+x^E_{imp}(t_{i+1})-\big(ab-a\,x^E_{imp}(t_{i+1})\big)\,(t_{i+1}-t_i)-\sigma\sqrt{x^E_{imp}(t_{i+1})}\,\big(W_{i+1}-W_i\big)-x^E_{imp}(t_i)=0.$$

Let

$$y_{i+1}=\sqrt{x^E_{imp}(t_{i+1})}\,,$$

then y_{i+1} is the positive solution of the following equation

$$y^2_{i+1}-\big(ab-a\,y^2_{i+1}\big)\,(t_{i+1}-t_i)-\sigma\,y_{i+1}\big(W_{i+1}-W_i\big)-x^E_{imp}(t_i)+\frac{\sigma^2}{2}\big(t_{i+1}-t_i\big)=0,$$

i.e.,

$$(1+a\,(t_{i+1}-t_i))\;y^2_{i+1}-\sigma\,\big(W_{i+1}-W_i\big)\,y_{i+1}-\left(x^E_{imp}(t_i)+ab\,(t_{i+1}-t_i)-\frac{\sigma^2}{2}\big(t_{i+1}-t_i\big)\right)=0.$$

Then y_{i+1} is given by

$$y_{i+1}=\frac{\sigma\,\big(W_{i+1}-W_i\big)+\sqrt{\sigma^2\,\big(W_{i+1}-W_i\big)^2\;+4\,\left(1+a\,(t_{i+1}-t_i)\right)\,\big(x^E_{imp}(t_i)+(ab-\frac{\sigma^2}{2})\,(t_{i+1}-t_i)\big)}}{2\,(1+a\,(t_{i+1}-t_i))}.$$

---

for some function $\widetilde{\mu}$, then

$$\begin{aligned}x(t_n)-x(0)&=\sum_{i=0}^{n-1}\big(x(t_{i+1})-x(t_i)\big)=\sum_{i=0}^{n-1}\widetilde{\mu}(x(t_{i+1}))\,(t_{i+1}-t_i)+\sum_{i=0}^{n-1}\sigma(x(t_{i+1}))\,(W_{i+1}-W_i)\\&=\sum_{i=0}^{n-1}\widetilde{\mu}(x(t_{i+1}))\,(t_{i+1}-t_i)+\sum_{i=0}^{n-1}\big(\sigma(x(t_{i+1}))-\sigma(x(t_i))\big)\,(W_{i+1}-W_i)+\sum_{i=0}^{n-1}\sigma(x(t_i))\,(W_{i+1}-W_i)\\&\simeq\sum_{i=0}^{n-1}\widetilde{\mu}(x(t_{i+1}))\,(t_{i+1}-t_i)+\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\big(x(t_{i+1})-x(t_i)\big)\,(W_{i+1}-W_i)+\sum_{i=0}^{n-1}\sigma(x(t_i))\,(W_{i+1}-W_i).\end{aligned}$$

As consequence, taking into account that $\big(t_{i+1}-t_i\big)\,\big(W_{i+1}-W_i\big)=o(t_{i+1}-t_i)$, $\big(W_{i+1}-W_i\big)^2\simeq t_{i+1}-t_i$, and therefore

$$\begin{aligned}&\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\big(x(t_{i+1})-x(t_i)\big)\,(W_{i+1}-W_i)\\&=\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\big(\widetilde{\mu}(x(t_{i+1}))\,(t_{i+1}-t_i)+\sigma(x(t_{i+1}))\,(W_{i+1}-W_i)\big)\,(W_{i+1}-W_i)\\&=\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\widetilde{\mu}(x(t_{i+1}))\,\big(t_{i+1}-t_i\big)\,\big(W_{i+1}-W_i\big)\;+\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\sigma(x(t_{i+1}))\,\big(W_{i+1}-W_i\big)^2\\&\simeq\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\sigma(x(t_{i+1}))\,(t_{i+1}-t_i)\end{aligned}$$

we obtain

$$\begin{aligned}x(t_n)-x(0)&=\simeq\sum_{i=0}^{n-1}\widetilde{\mu}(x(t_{i+1}))\,\big(t_{i+1}-t_i\big)+\sum_{i=0}^{n-1}\sigma'(x(t_i))\,\sigma(x(t_{i+1}))\,(t_{i+1}-t_i)+\sum_{i=0}^{n-1}\sigma(x(t_i))\,(W_{i+1}-W_i)\\&\simeq\int_0^{t_n}\big(\widetilde{\mu}(x(s))+\sigma\,\sigma'(x(s))\big)\,ds+\int_0^{t_n}\sigma(x(s))\,dW_s\end{aligned}$$

where for $t_i<t<t_{i+1}$, $x(t)=x(t_i)$.

The latter formula explain why the correct choice for $\widetilde{\mu}$ is

$$\widetilde{\mu}(x)=\mu(x)-\sigma(x)\sigma'(x).$$

Observe that we have assumed $a > 0$, $b > 0$ and $ab > \frac{\sigma^2}{2}$, so that if $x^E_{imp}(t_i) > 0$ then $y_{i+1} > 0$, as should be. Finally we obtain

$$x^E_{imp}(t_{i+1}) = y^2_{i+1}.$$

In order to simulate $r(t)$ and $\lambda(t)$, we apply the previous scheme to $r(t)$ with $a = a_r$, $b = b_r$ and $\sigma = \bar{\sigma}_r$, and we get $r^E_{imp}(t_i)$. Subsequently, following a similar reasoning, we can extend this approach also to the process $\lambda^{(c)}(t)$ given by (3.51), i.e., setting $\gamma(\lambda, r) = a_\lambda\,(b_\lambda - \lambda + c\,r)$ and $\sigma(\lambda) = \bar{\sigma}_\lambda\sqrt{\lambda}$,

$$\begin{aligned} d\lambda^{(c)}(t) &= a_\lambda\left(b_\lambda - \lambda^{(c)}(t) + c\,r(t)\right)dt + \bar{\sigma}_\lambda\sqrt{\lambda^{(c)}(t)}dW^\lambda(t) \\ &= \gamma(\lambda^{(c)}(t), r(t))dt + \sigma(\lambda^{(c)}(t))dW^\lambda(t). \end{aligned}$$

The implicit scheme $\lambda^{c,E}_{imp}(t_{i+1})$ is given by

$$\begin{aligned} \lambda^{c,E}_{imp}(t_{i+1}) &= \lambda^{c,E}_{imp}(t_i) + \left(\gamma(\lambda^{c,E}_{imp}(t_{i+1}), r^E_{imp}(t_i)) - \sigma(\lambda^{c,E}_{imp}(t_{i+1}))\sigma'(\lambda^{c,E}_{imp}(t_{i+1}))\right)(t_{i+1} - t_i) \\ &+ \sigma(\lambda^{c,E}_{imp}(t_{i+1}))\big(W^\lambda_{i+1} - W^\lambda_i\big) \end{aligned} \tag{7.4}$$

which becomes

$$\begin{aligned} &\frac{\sigma\sigma'(\lambda^{c,E}_{imp}(t_{i+1}))}{2}\,(t_{i+1} - t_i) + \lambda^{c,E}_{imp}(t_{i+1}) - \gamma\big(\lambda^{c,E}_{imp}(t_{i+1})r^E_{imp}(t_i)\big)(t_{i+1} - t_i) \\ &- \sigma(\lambda^{c,E}_{imp}(t_{i+1}))\big(W^\lambda_{i+1} - W^\lambda_i\big) - \lambda^{c,E}_{imp}(t_i) = 0. \end{aligned} \tag{7.5}$$

Now proceeding exactly as above for $r(t)$, again we can find a closed form solution $\lambda^{c,E}_{imp}(t_{i+1})$. Let

$$v_{i+1} = \sqrt{\lambda^{c,E}_{imp}(t_{i+1})}\,,$$

then v_{i+1} is the positive solution of the following equation

$$v^2_{i+1} - \big(a_\lambda b_\lambda - a_\lambda\,v^2_{i+1} + cr^E_{imp}(t_i)\big)\,(t_{i+1} - t_i) - \bar{\sigma}_\lambda\,v_{i+1}\left(W^\lambda_{i+1} - W^\lambda_i\right) - \lambda^{c,E}_{imp}(t_i) + \frac{\bar{\sigma}^2_\lambda}{2}\left(t_{i+1} - t_i\right) = 0,$$

so that v_{i+1} is given by

$$\begin{aligned} v_{i+1} = &\frac{\bar{\sigma}_\lambda\left(W^\lambda_{i+1} - W^\lambda_i\right)}{2\,(1 + a_\lambda\,(t_{i+1} - t_i))} \\ &\frac{\sqrt{\bar{\sigma}^2_\lambda\left(W^\lambda_{i+1} - W^\lambda_i\right)^2 + 4\,(1 + a_\lambda\,(t_{i+1} - t_i))\left(\lambda^{c,E}_{imp}(t_i) + (a_\lambda b_\lambda + cr^E_{imp}(t_i) - \frac{\bar{\sigma}^2_\lambda}{2})\,(t_{i+1} - t_i)\right)}}{2\,(1 + a_\lambda\,(t_{i+1} - t_i))}. \end{aligned} \tag{7.6}$$

Observe that the expression inside the square root is strictly positive since $r^E_{imp}(t_i) > 0$, $a_\lambda > 0$ and $a_\lambda b_\lambda > \frac{\bar{\sigma}^2_\lambda}{2}$. Finally we obtain

$$\lambda^{c,E}_{imp}(t_{i+1}) = v^2_{i+1} > 0,$$

so that

$$\lambda(t_{i+1}) = \frac{\lambda^{c,E}_{imp}(t_{i+1}) + D}{T - t_{i+1}}$$

and for $t_i < t < t_{i+1}$

$$\lambda(t) = \frac{\lambda^{c,E}_{imp}(t) + D}{T - t} \tag{7.7}$$

where

$$\text{either}\qquad \lambda^{c,E}_{imp}(t) = \lambda^{c,E}_{imp}(t_i), \qquad \text{or} \qquad \lambda^{c,E}_{imp}(t) = \lambda^{c,E}_{imp}(t_i) + \frac{t - t_i}{t_{i+1} - t_i}\left(\lambda^{c,E}_{imp}(t_{i+1}) - \lambda^{c,E}_{imp}(t_i)\right).$$

To calculate the optimal consumption and portfolio weights we need to simulate paths of $(r(t), \lambda(t))$. In the sequel we apply the above implicit positivity-preserving Euler scheme for the processes $r(t)$ and $\lambda(t)$. The value of the parameters are shown in table 7.1.

Interest Rate ($r(t)$)	Parameters	Mortality Intensity ($\lambda(t)$)	Parameters
a_r	0.20	a_λ	0.05
b_r	0.031	b_λ	0.001
$\bar{\sigma}_r$	0.01	$\bar{\sigma}_\lambda$	0.001
r	0.05	λ	0.037
		D	3.75
		c	0.01

Table 7.1: Value of parameters

Recalling that T is a suitable deterministic time such that, on the basis of demographic considerations, the investor will be dead at time T with probability 1, we take $T = 100$. The results of the simulations[2] carried out on a 90 years period are drawn in Figs. 7.1 and 7.2, where the solid line represents the mean value of $M = 20$ paths and the two dashed lines represents the 95% confidence intervals. Fig. 7.1 shows the simulated path for the interest rate $r(t)$, while Fig. 7.2 shows the simulated path for the stochastic intensity $\lambda(t)$

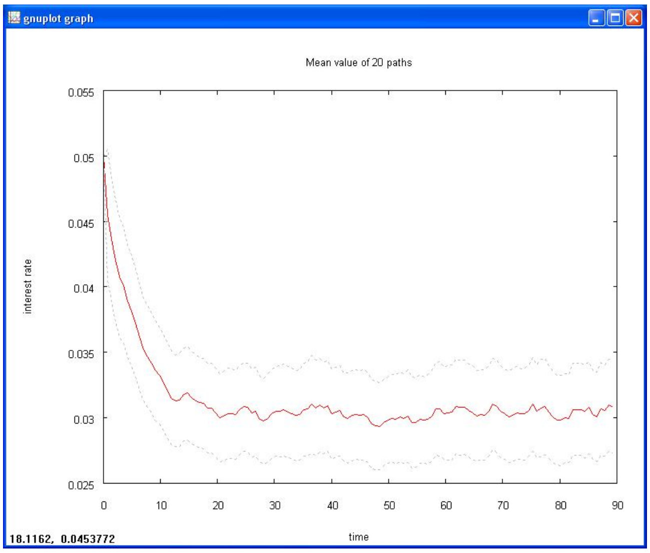

Figure 7.1: Mean value of 20 paths for the interest rate $r(t)$

[2]The C program is left to an Appendix ??

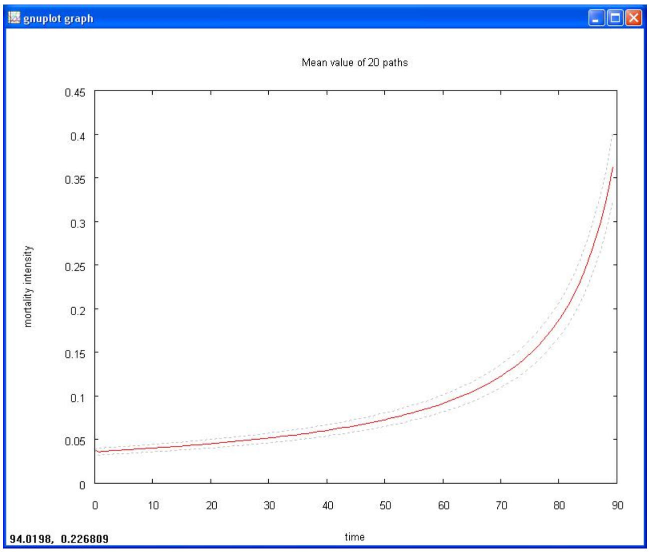

Figure 7.2: Mean value of 20 paths for the stochastic intensity $\lambda(t)$

Appendix A

A.1 Some technical results

Here we recall some basic results which are used in this framework. Let $(\Omega, \mathcal{F}, P)$ be a probability space endowed with a filtration $\mathbb{F}$ satisfying the usual conditions. Then we have the following results.

Lemma A.1.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Assume that $\mathcal{A}$ and $\mathcal{M}$ are σ-algebras contained in $\mathcal{F}$, independent of each other. Let ξ be a random variable taking values in measurable space $(S, \mathcal{S})$, and assume that ξ is measurable with respect to $\mathcal{M}$. Let $\psi : S \times \Omega \to \mathbb{R}$, $(x, \omega) \mapsto \psi(x, \omega)$ be a real valued function, $\mathcal{S} \times \mathcal{A}$ jointly measurable, and such that $\psi(\xi(\omega), \omega)$ is integrable. Then the conditional expectation of $\Psi(\omega) := \psi(\xi(\omega), \omega)$ with respect to $\mathcal{M}$ is given by*

$$E\left[\Psi|\mathcal{M}\right] = E[\psi(x,\omega)]\Big|_{x=\xi(\omega)}. \tag{A.1}$$

Proof. The proof is based on the observation that ***(i)*** relation (A.1) is straightforward when $\psi(x, \omega) = f(x)\, Z(\omega)$, with f a (deterministic) measurable function, and Z is a $\mathcal{A}$-measurable random variable, and therefore for any linear combination of such functions, ***(ii)*** without loss of generality one can assume $\psi(x, \omega)$ non negative, ***(iii)*** the class of non negative functions $\psi(x, \omega)$ such that (A.1) holds is a monotone class.

□

Lemma A.1.2. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space. Let $\alpha = (\alpha(t), t \in [0, T])$ be some stochastic process with $\int_0^T E(|\alpha(t)|dt) < \infty$, and let $\mathcal{S}$ be some sub-σ-algebras of $\mathcal{F}$. Then*

$$E\left(\int_0^t \alpha(s)ds|\mathcal{S}\right) = \int_0^t E\left(\alpha(s)ds|\mathcal{S}\right) ds \quad a.s., \quad 0 \le t \le T. \tag{A.2}$$

Proof. Let μ be a bounded $\mathcal{S}$-measurable random variable. Then using the Fubini theorem, we find that

$$E\left(\mu \int_0^t \alpha(s)ds\right) = \int_0^t E\left(\mu\alpha(s)\right) ds = \int_0^t E\left(\mu E\left(\alpha(s)|\mathcal{S}\right)\right) ds = E\left(\mu \int_0^t E\left(\alpha(s)|\mathcal{S}\right) ds\right). \tag{A.3}$$

On the other hand

$$E\left(\mu \int_0^t \alpha(s)ds\right) = E\left(\mu E\left(\int_0^t \alpha(s)ds|\mathcal{S}\right)\right). \tag{A.4}$$

Hence

$$E\left(\mu \int_0^t \alpha(s)ds\right) = E\left(\mu E\left(\int_0^t \alpha(s)ds|\mathcal{S}\right)\right). \tag{A.5}$$

From this, because of the arbitrariness of μ, we obtain (A.2).

□

Lemma A.1.3. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual conditions and $X(t)$ be a Markov process with respect to $\mathbb{F}$. If $\mathbb{A} = \{\mathcal{A}_t : t \in [0, T]\}$ is a filtration such that*

$$\mathcal{F}_t^X \subseteq \mathcal{A}_t \subseteq \mathcal{F}_t, \ \forall t \tag{A.6}$$

then $X(t)$ is a Markov process with respect to $\mathbb{A}$.

Proof. It is easily seen that $X(t)$ is a $\mathcal{A}_t$-Markov process.

In fact, for all Borel measurable, bounded functions f, we have

$$\begin{aligned} E\left[f\left(X(t+s)\right)|\mathcal{A}_t\right] &= E\left[E\left[f\left(X(t+s)\right)|\mathcal{F}_t\right]|\mathcal{A}_t\right] \\ &= E\left[E\left[f\left(X(t+s)\right)|X(t)\right]|\mathcal{A}_t\right] \\ &= E\left[f\left(X(t+s)\right)|X(t)\right], \end{aligned}$$

where in the last step we have used that $\mathcal{F}_t^X \subseteq \mathcal{A}_t$ so that $E\left[f\left(X(t+s)\right)|X(t)\right]$ is $\mathcal{A}_t$-measurable. □

Appendix B

Stochastic Differential Equations

We give a brief summary of the definitions and results which are the background in this framework. For proofs and more information we refer to Øksendal [19] and Karatzas and Shreve [16].

B.1 Itô Diffusion

Now introduce the concept of stochastic differential equation with respect to Wiener process and its solution in the so called strong sense. We discuss the questions of existence and uniqueness of such solutions, as well as an comparison result in one dimensional case and the connection with partial differential equations.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathbb{F}$ be a filtration satisfying the usual conditions and $W(t)$ be a r-dimensional Wiener process with respect to $\mathbb{F}$. Let

$$\mu(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d,$$
$$\sigma(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \to M(d,r),$$

be a measurable functions, where $M(d,r)$ denoted the class of $d \times r$ matrices. The intent is to assign a meaning to the stochastic differential equation

$$\begin{cases} dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \\ X(u) = x \end{cases} \tag{B.1}$$

where $x \in \mathbb{R}^d$.

Definition B.1.1. *We say that a continuous stochastic process $X(t)$ is a solution of the stochastic differential equation (B.1) if*

1. *$X(t)$ is $\mathbb{F}$-adapted;*
2. *For every $1 \le i \le d$, $1 \le j \le r$ and $u \le t \le \infty$*

$$\int_u^t |\mu_i(s, X(s))| ds + \int_u^t \sigma_{ij}^2(s, X(s)) ds < \infty, \tag{B.2}$$

   holds[1] *a.s.;*

3. *For all $t \ge u$ we have that*

$$X(t) = x + \int_u^t \mu\big(s, X(s)\big)dt + \int_u^t \sigma\big(s, X(s)\big)dW(s). \tag{B.3}$$

Furthermore we recall that the coefficients of this equation, $b(t,x)$ and $\sigma(t,x)$ are called the drift and diffusion term of $X(t)$.

Definition B.1.2 (Strong Existence)**.** *We say that the stochastic differential equation (B.1) admit a strong existence if for each filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, and Wiener process (with respect to $\mathbb{F}$) $W(t)$, there exists a stochastic process $X(t)$ which is solution of (B.1).*

[1] We are assuming the integrability conditions so that the deterministic and stochastic integrals in (B.3) are defined.

Theorem B.1.1 ([19]). *Suppose that there exist a constant K such that the following conditions are satisfied*

$$\|\mu(t,x)-\mu(t,y)\| \le K\|x-y\|, \quad x,y\in\mathbb{R}^N,\ t\in[0,T] \tag{B.4}$$

$$\|\sigma(t,x)-\sigma(t,y)\| \le K\|x-y\|, \quad x,y\in\mathbb{R}^N,\ t\in[0,T] \tag{B.5}$$

$$\|\mu(t,x)\|+\|\sigma(t,x)\| \le K(1+\|x\|), \quad x\in\mathbb{R}^N,\ t\in[0,T] \tag{B.6}$$

where $|\cdot|$ denotes the Euclidean norm and $\|\sigma(t,x)\|^2=\sum_{i=1}^N\sum_{j=1}^M\sigma_{ij}^2(t,x)$. Then, for any $t\ge 0$, the stochastic differential equation (B.1) admits a unique solution.

Now we discuss a notion of solvability for the stochastic differential equation (B.1) which, although weaker that the one introduced previously, is yet extremely usefel and fruitful in both teory and applications. In particular one can prove existence and uniqueness of solutions under assumptions much weaker than those of the previous theorem.

Definition B.1.3 (Weak Existence). *We say that the stochastic differential equation (B.1) admit a weak existence if there exists a filtered probability space $(\Omega,\mathcal{F},\mathbb{F},P)$, a Wiener process (with respect to $\mathbb{F}$) $W(t)$, a stochastic process $X(t)$ which is solution of (B.1).*

Sometimes equation (B.1) may have solutions which are unique in the weaker sense that only their probability laws coincide, but not necessarily their sample paths. We shall say then that we have a unique weak solution.

Definition B.1.4 (Uniqueness in the sense of probability law). *We say that the stochastic differential equation (B.1) admit a unique solution in the sense of probability law if, for any two weak solutions $X(t)$ and $\tilde{X}(t)$ of (B.1), the two processes $X(t)$ and $\tilde{X}(t)$ have the same law.*

Now we recall a comparison result in the following proposition.

Proposition B.1.2 (Proposition 2.18 of [16]). *Suppose that on a certain probability space $(\Omega,\mathcal{F},P)$ equipped with a filtration $\mathbb{F}$ which satisfies the usual conditions, we have a standard, one dimensional Wiener process $W(t)$, and two continuous, adapted processes $X^i(t)$, for $i=1,2$, such that*

$$\begin{cases} dX^i(t)=b^i(t,X^i(t))dt+\sigma(t,X^i(t))dW(t), & 0\le t\le\infty,\\ X^i(0)=X_0^i \end{cases} \tag{B.7}$$

holds a.s. for $i=1,2$. We assume that

1. *the coefficients $\sigma(t,x)$, $b^i(t,x)$ are continuous, real valued function on $[0,\infty)\times\mathbb{R}$;*
2. *$\sigma(t,x)$ is such that*

$$|\sigma(t,x)-\sigma(t,y)|\le h(|x-y|), \tag{B.8}$$

for every $0\le t<\infty$, and $x\in\mathbb{R}$, $y\in\mathbb{R}$, where $h:[0,\infty)\to[0,\infty)$ is a strictly increasing function with $h(0)=0$ and

$$\int_0^\epsilon \frac{1}{h^2(u)}du=\infty, \quad \forall\epsilon>0; \tag{B.9}$$

3. *$X_0^1\le X_0^2$ a.s.;*
4. *$b^1(t,x)\le b^2(t,x)$, $\forall t\in[0,\infty)$, $x\in\mathbb{R}$ and either $b^1(t,x)$ or $b^2(t,x)$ are such that*

$$|b(t,x)-b(t,y)|\le K(|x-y|), \tag{B.10}$$

for every $0\le t<\infty$, and $x\in\mathbb{R}$, $y\in\mathbb{R}$, where K is positive constant.

Then

$$P\left(X^1(t)\le X^2(t)\right)=1 \quad \forall t\in[0,\infty). \tag{B.11}$$

Finally we explore the connection which exist between stochastic differential equation and certain parabolic partial differential equations. For example the so called Cauchy problem. Now, instead of solve the Cauchy problem using purely analytical tools, we will produce a so called stochastic representation formula, which gives the solution to the Cauchy problem in terms of the solution to an stochastic differential equation associated to the Cauchy problem in a natural way.

In the sequel we shall be considering a solution to the stochastic integral equation (B.3) under the following assumptions

1. the coefficients $\mu(t,x)$ and $\sigma(t,x)$ are continuous and satisfy the linear growth condition (B.6);
2. the stochastic integral equation has a weak solution $X(t)$ for every pair (t,x) and this solution $X(t)$ is unique in the sense of probability law.

For every $t \geq 0$, we introduce the partial differential operator

$$\mathcal{A}f(t,x) = \sum_{i=1}^{d} \mu_i(t,x)\frac{\partial f}{\partial x_i}(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x)\frac{\partial^2 f}{\partial x_i \partial x_j}(t,x), \quad f \in C^2, \tag{B.12}$$

where μ_i, x_i represent the ith component of the vector μ, x respectively, and

$$a(t,x) = \sigma(t,x)\sigma'(t,x).$$

With an arbitrary but fixed $T > 0$ and appropriate constants $L > 0$, $c \geq 1$, we consider functions $f(x)$: $\mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, $g(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and $k(t,x) : [0,T] \times \mathbb{R}^d \to [0,\infty)$ which are continuous and satisfy

$$\text{(i) } |f(x)| \leq L(1+|x|^{2c}) \quad \text{or} \quad \text{(ii) } f(x) \geq 0, \ \forall x \in \mathbb{R}^d \tag{B.13}$$

as well as

$$\text{(i) } |g(t,x)| \leq L(1+|x|^{2c}) \quad \text{or} \quad \text{(ii) } g(t,x) \geq 0, \ \forall t \in [0,T], x \in \mathbb{R}^d, \tag{B.14}$$

where $|\cdot|$ denotes the Euclidean norm.

Theorem B.1.3 ([16]). *Under the preceding assumptions, we suppose that*

$$v(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$$

is continuous, $v \in C^{1,2}([0,T),\mathbb{R}^d)$, and satisfies the Cauchy problem

$$-\frac{\partial v}{\partial t}(t,x) + k(t,x)v(t,x) = \mathcal{A}v(t,x) + g(t,x) \quad t \in [0,T], x \in \mathbb{R}^d \tag{B.15}$$

$$v(T,x) = f(x) \quad x \in \mathbb{R}^d \tag{B.16}$$

as well as the polynomial growth condition

$$\max_{0 \leq t \leq T} |v(t,x)| \leq M(1+|x|^{2\nu}), \quad x \in \mathbb{R}^d, \tag{B.17}$$

for some $M > 0$, $\nu \geq 1$. Then $v(t,x)$ admits the stochastic representation

$$v(t,x) = E_{t,x}\Big[f(X(T))e^{-\int_t^T k(u,X(u))du} + \int_t^T g(s,X(s))e^{-\int_t^T k(u,X(u))du}\,ds\Big] \tag{B.18}$$

on $[0,T] \times \mathbb{R}^d$ and where the subscripts t, x denote that the expectation are taken using the following dynamics for $s \in [0,T]$

$$\begin{cases} dX(s) = & \hat{\mu}(s,X(s))ds + \sigma(s,X(s))dW(s) \\ X(t) = & x \end{cases} \tag{B.19}$$

In particular, such a solution is unique.

Remark B.1.1. *A set of conditions sufficient under which the Cauchy problem (B.15) and (B.16) has a solution satisfying the polynomial growth condition (B.17) is*

1. *Uniform ellipticity: there exists a positive constant δ such that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$*

$$\sum_{i,j=1}^{d} a_{ij}(t,x)\xi_i\xi_j \geq \delta \left|\xi\right|^2. \tag{B.20}$$

2. *Boundedness: the functions $a_{i,j}(t,x)$, $b_i(t,x)$ and $k(t,x)$ are bounded in $[0,\infty] \times \mathbb{R}^d$.*

3. *Hölder continuity: the functions $a_{i,j}(t,x)$, $b_i(t,x)$, $k(t,x)$ and $g(t,x)$ are uniformly Hölder continuous in $[0,T] \times \mathbb{R}^d$.*

4. *Polynomial growth: the functions $f(x)$ and $g(t,x)$ satisfy (B.13)(i) and (B.14)(i), respectively.*

B.2 Itô Process

Now we extend the comparison result of Proposition B.1.2 at the case of a generic Itô Process. Before proceeding to give the following definition.

Definition B.2.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathbb{F}$ be a filtration satisfying the usual conditions and $W(t)$ be a 1-dimensional Wiener process with respect to $\mathbb{F}$. A (1-dimensional) Itô process is a stochastic process $X(t)$ on $(\Omega, \mathcal{F}, P)$ of the form*

$$X(t) = X_0 + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) \tag{B.21}$$

where W is a Wiener process, $\mu(t)$ and $\sigma(t)$ are adapted[2].

If $X(t)$ is an Itô process of the form (B.21), the equation (B.21) is sometimes written in the shorter differential form

$$\begin{cases} dX(t) = \mu(t)dt + \sigma(t)dW(t), \\ X(0) = X_0 \end{cases} \tag{B.22}$$

Now we have the following comparison theorem.

Theorem B.2.1. *Suppose that on a certain probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F}$ which satisfies the usual conditions, we have a standard, one dimensional Wiener process $W(t)$, and two continuous, adapted processes $X^i(t)$, for $i = 1, 2$, such that*

$$\begin{cases} dX^1(t) = b^1(t, X^1(t))dt + \sigma(t, X^1(t))dW(t), \quad 0 \leq t \leq \infty, \\ X^1(0) = X_0^1 \end{cases}$$

and

$$\begin{cases} dX^2(t) = \big(b^2(t, X^2(t)) + \alpha(t,\omega)\big)dt + \sigma(t, X^2(t))dW(t), \quad 0 \leq t \leq \infty, \\ X^2(0) = X_0^2 \end{cases}$$

hold a.s. We assume that

1. *the coefficient $\sigma(t,x)$ is continuous, real valued function on $[0,\infty) \times \mathbb{R}$;*

2. *$\sigma(t,x)$ is such that*

$$|\sigma(t,x) - \sigma(t,y)| \leq h(|x-y|), \tag{B.23}$$

   for every $0 \leq t < \infty$, and $x \in \mathbb{R}$, $y \in \mathbb{R}$, where $h : [0,\infty) \to [0,\infty)$ is a strictly increasing function with $h(0) = 0$ and (B.9) holds;

3. *$X_0^1 \leq X_0^2$ a.s.;*

[2]We assume implicitely the integrability conditions that are necessary to define the right hand side of (B.21).

4. $b^1(t,x) \le b^2(t,x)$, $\forall t \in [0,\infty)$, $x \in \mathbb{R}$ *and either* $b^1(t,x)$ *or* $b^2(t,x)$ *are such that*

$$|b(t,x) - b(t,y)| \le K|x-y|, \tag{B.24}$$

for every $0 \le t < \infty$, *and* $x \in \mathbb{R}$, $y \in \mathbb{R}$, *where* K *is positive constant.*

5. $\alpha(t,\omega) \ge 0$ *a.s.*

Then

$$P\left(X^1(t) \le X^2(t)\right) = 1 \quad \forall t \in [0,\infty). \tag{B.25}$$

Appendix C

C.1 A bidimensional CIR: an approximation for the term structure equation

In this section we prove the series expansion announced in Remark 4.3.2 for the functions x_1 and x_2, that give the representation (4.85) for the function $\psi_r^{(c)}$. We recall that the function $\psi_r^{(c)}$ appears in the expression (4.72) for $\hat{L}^T(t,z)$ and is bounded above by 0 and below by $-\frac{1}{|\beta^-|}$ (see (4.77) in Proposition 4.3.7).

We recall that in the proof of Proposition 4.3.7 the functions x_1 and x_2 are defined as $x_1 = y_r$ and $x_2 = \dot{y}_r$, where y_r is given by (4.83). Therefore, using the above bound for $\psi_r^{(c)}$, we observe that

$$0 \le x_1(s) = y_r(s) = e^{-\int_0^s \frac{\bar{\sigma}_r^2}{2}\,\psi_r^{(c)}(u)\,du} \le e^{\frac{1}{|\beta^-|}\,s}$$

and

$$0 \le x_2(s) = \dot{y}_r(s) = -\frac{\bar{\sigma}_r^2}{2}\,\psi_r^{(c)}(s)\,e^{-\int_0^s \frac{\bar{\sigma}_r^2}{2}\,\psi_r^{(c)}(u)\,du} \le \frac{\bar{\sigma}_r^2}{2|\beta^-|}\,e^{\frac{1}{|\beta^-|}\,s}.$$

The functions x_1 and x_2 solve the system $\dot{x}_1(s) = x_2(s)$, $\dot{x}_2(s) = A(s)x_1(s) - Bx_2(s)$, with initial conditions $x_1(0) = 1$, $x_2(0) = 0$, where $A(s) = \frac{\bar{\sigma}_r^2}{2}\big(a_\lambda\, c\, \psi_\lambda(s) + 1\big)$, and $B = a_r$ (see (4.84)). Then, clearly

$$\begin{aligned}
x_1(t) &= 1 + \int_0^t x_2(s)ds \\
x_2(t) &= \int_0^t \big(A(s)x_1(s) - Bx_2(s)\big)\,ds = \int_0^t \Big(A(s)\Big(1 + \int_0^s x_2(u)du\Big) - Bx_2(s)\Big)\,ds \\
&= \int_0^t A(s)\,ds + \int_0^t A(s)\int_0^s x_2(u)du\,ds - \int_0^t Bx_2(s)\,ds \\
&= \int_0^t A(s)\,ds + \int_0^t \Big(\int_u^t A(s)ds - B\Big)x_2(u)\,du
\end{aligned}$$

Let us denote by

$$C(u,t) := \int_u^t A(s)ds - B = \Im_A(t) - \Im_A(u) - B, \qquad 0 \le u \le t,$$

where

$$\Im_A(t) = \frac{\bar{\sigma}_r^2}{2}\Big(a_\lambda\, c \int_0^t \psi_\lambda(s) + t\Big) = \frac{\bar{\sigma}_r^2}{2}\,a_\lambda\, c\, \psi_\lambda^0(t) + \frac{\bar{\sigma}_r^2}{2}\,t$$

Note that

$$\frac{\partial}{\partial t}C(u,t) := A(t), \qquad \text{and} \qquad C(t,t) := -B$$

and that, for c sufficiently small (recall that $B = a_r$)

$$-a_r \le C(u,t) := \int_u^t A(s)ds - B = \Im_A(t) - \Im_A(u) - B \le \frac{\bar{\sigma}_r^2}{2}\,(t-u) - a_r \le \frac{\bar{\sigma}_r^2}{2}\,t - a_r, \qquad 0 \le u \le t,$$

i.e.,

$$|C(u,t)| \le \max\big(a_r, \big|\frac{\bar{\sigma}_r^2}{2}\,t - a_r\big|\big) \le \frac{\bar{\sigma}_r^2}{2}\,t + a_r =: k(t), \qquad 0 \le u \le t.$$

(Observe that, even if c is not small, there exists a linear function $k(t)$ such that $|C(u,t)| \le k(t)$ for $0 \le u \le t$)

With the above notation we get

$$x_2(t) = \int_0^t A(s)\,ds + \int_0^t C(u,t)\,x_2(u)\,du$$

so that

$$\begin{aligned}
&x_2(t)\\
&= \int_0^t A(s_1)\,ds_1 + \int_0^t C(u_1,t)\,x_2(u_1)\,du_1\\
&= \int_0^t A(s_1)\,ds_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s_2)\,ds_2 + \int_0^{u_1} C(u_2,u_1)\,x_2(u_2)\,du_2\Big)\,du_1\\
&= \int_0^t A(s_1)\,ds_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s_2)\,ds_2\Big)\,du_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,x_2(u_2)\,du_2\Big)\,du_1\\
&= \int_0^t A(s_1)\,ds_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s_2)\,ds_2\Big)\,du_1\\
&+ \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(\int_0^{u_2} A(s)\,ds + \int_0^t C(u_3,u_2)\,x_2(u_3)\,du_3\Big)\,du_2\Big)\,du_1\\
&= \int_0^t A(s_1)\,ds_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s_2)\,ds_2\Big)\,du_1 + \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(\int_0^{u_2} A(s)\,ds\Big)\,du_2\Big)\,du_1\\
&+ \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\Big(\int_0^{u_2} C(u_3,u_2)\,x_2(u_3)\,du_3\Big)\,du_2\Big)\,du_1\\
&= \int_0^t A(s)\,ds + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s)\,ds\Big)\,du_1\\
&+ \sum_{k=1}^{m} \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 \qquad \text{(C.1)}\\
&+ \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\,x_2(u_m)\Big)...\Big)\,du_2\Big)\,du_1. \qquad \text{(C.2)}
\end{aligned}$$

Taking into account that

$$|A(s)| \le \overline{A},$$

$$0 \le x_2(u) \le \frac{\bar\sigma_r^2}{2|\beta^-|}\,e^{\frac{1}{|\beta^-|}\,u} \le \frac{\bar\sigma_r^2}{2|\beta^-|}\,e^{\frac{1}{|\beta^-|}\,t}, \qquad 0 \le u \le t,$$

and that, for c sufficiently small,

$$|C(u,v)| \le k(v) \le k(t) = \frac{\bar\sigma_r^2}{2}\,t + a_r, \qquad 0 \le u \le v \le t.$$

we get that the above sum (C.1) converges and that the rest (C.2) goes to zero, for each fixed $t > 0$.

Finally we can show that x_2 coincides with the function

$$\begin{aligned}
x(t) := &\int_0^t A(s)\,ds + \int_0^t C(u_1,t)\Big(\int_0^{u_1} A(s)\,ds\Big)\,du_1\\
&+ \sum_{k=1}^{\infty} \int_0^t C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{k-1}} C(u_k,u_{k-1})\Big(\int_0^{u_k} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1
\end{aligned}$$

which can also be written as

$$\begin{aligned}
:= &\int_0^t A(s)\,ds + \int_0^t A(s)\Big(\int_s^t C(u_1,t)\,du_1\Big)\,ds\\
&+ \sum_{k=1}^{\infty} \int_0^t A(s)\Big(\int_s^t du_m\Big(\int_{u_m}^t du_{m-1}...\int_{u_3}^t du_2\,\int_{u_2}^t du_1 C(u_1,t)\Big)\,C(u_2,u_1)\,...\,C(u_m,u_{m-1})\Big)\Big)\,ds.
\end{aligned}$$

It is not hard to see that x is not only well defined, but also has a derivative, such that

$$\dot x(t) = -B\,x(t) + A(t)\Big[1 + \int_0^t x(u_1)\,du_1\Big] \qquad \text{(C.3)}$$

Indeed

$$
\begin{aligned}
\dot{x}(t) &= A(t) + C(t,t)\Big(\int_0^t A(s)\,ds\Big) + \int_0^t \frac{\partial}{\partial t}C(u_1,t)\Big(\int_0^{u_1} A(s)\,ds\Big)\,du_1 \\
&+ \sum_{k=1}^{\infty} C(t,t)\Big(\int_0^t C(u_2,t)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 + \\
&+ \sum_{k=1}^{\infty} \int_0^t \frac{\partial}{\partial t}C(u_1,t)\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 \\
&= A(t) - B\,\Big(\int_0^t A(s)\,ds\Big) + \int_0^t A(t)\,\Big(\int_0^{u_1} A(s)\,ds\Big)\,du_1 \\
&- \sum_{k=1}^{\infty} B\,\Big(\int_0^t C(u_2,t)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 + \\
&+ \sum_{k=1}^{\infty} \int_0^t A(t)\,\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 \\
&= -B\,\Big(\int_0^t A(s)\,ds\Big) \\
&- \sum_{k=1}^{\infty} B\,\Big(\int_0^t C(u_2,t)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 \\
&+ A(t) + A(t)\,\int_0^t \Big(\int_0^{u_1} A(s)\,ds\Big)\,du_1 \\
&+ A(t)\,\int_0^t \sum_{k=1}^{\infty}\Big(\int_0^{u_1} C(u_2,u_1)\,\Big(...\Big(\int_0^{u_{m-1}} C(u_m,u_{m-1})\Big(\int_0^{u_m} A(s)\,ds\Big)\Big)...\Big)\,du_2\Big)\,du_1 \\
&= -B\,x(t) + A(t)\Big[1 + \int_0^t x(u_1)\,du_1\Big]
\end{aligned}
$$

Then $x(t) = x_2(t)$ since there exists a unique solution of (C.3).

Finally we observe that one could easily find an upper bound for the truncation error of this representation, and use the above expansion to get an approximation for x_1 and x_2 and therefore of $\psi_r^{(c)}$. However in the simulations we prefer to approximate $\psi_r^{(c)}$ by numerical schemes for ordinary differential equations.

Appendix D

D.1 Stochastic Optimal Control: general results

In this section we describe a brief review of the theory of stochastic optimal control problems with the dynamic programming method. In particular, we refer to Fleming and Soner [12].

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space under Assumptions 1.2.1 and 1.2.2, and let $\mu(t,x,u)$ and $\Sigma(t,x,u)$

$$\mu : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n,$$
$$\Sigma : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{n\times d},$$

be given functions. For a given point $x_0 \in \mathbb{R}^n$ we consider the following controlled stochastic differential equation,

$$\begin{cases} dX(t) = \mu(t, X(t), u(t))dt + \Sigma(t, X(t), u(t))dW(t), \\ X(t_0) = x_0, \end{cases} \tag{D.1}$$

where W is a d-dimensional Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. The n-dimensional process X is called the state process (or state variable), the process u is called control process. We can control the state process X by choosing the k-dimensional control process u in a suitable way.

Our first modelling problem concerns the class of the admissible control processes. In general we require that the control process u is $\mathbb{F}$-adapted. In most concrete cases we also have to satisfy some control constraints, and we model this by taking as given a fixed subset U_{ad}, with $U_{ad} \subseteq \mathbb{R}^k$, and requiring that $u(t) \in U_{ad}$ for each t. We can now define the class of admissible control process.

Definition D.1.1 (Admissible Control Process). *A control process u is called admissible if*

- *$u(t)$ is $\mathbb{F}$-adapted;*
- *$u(t) \in U_{ad}$ for all $t \in \mathbb{R}_+$;*
- *For any given initial point (t,x) the stochastic differential equation for $s \in [t, \infty)$*

$$\begin{cases} dX(s) = \mu(s, X(s), u(s))ds + \Sigma(s, X(s), u(s))dW(s), \\ X(t) = x, \end{cases} \tag{D.2}$$

  has a unique solution.

The class of admissible control process is denoted $\mathcal{U}_{ad}^{\mathbb{F}}$.

For a given control process u, the solution process X of (D.2) will of course depend on the initial value, as well as on the chosen control process u. To be precise we denote the process X by $X^{u(\cdot)}$. Fixing $X^{u(\cdot)}(t) = x$, we will denote the unique solution $X^{u(\cdot)}(s)$ of (D.2) also by $X^{u(\cdot)}(s; t, x)$, $s \geq t$.

Now given a function

$$F : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}, \tag{D.3}$$

we have the following definition.

Definition D.1.2 (Optimal Control Problem). *Let $\mathcal{J} : \mathbb{R}_+ \times \mathcal{U}_{ad}^{\mathbb{F}} \to \mathbb{R}$ be defined as*

$$\mathcal{J}\big(t, u(\cdot)\big) = E\left[\int_t^\infty F(s, X^{u(\cdot)}(s), u(\cdot))ds \Big| \mathcal{F}_t\right], \tag{D.4}$$

The optimal control problem is defined as the problem of maximizing $\mathcal{J}\big(t, u(\cdot)\big)$ over $u(\cdot) \in \mathcal{U}_{ad}^{\mathbb{F}}$.

In order to ensure that $\mathcal{J}$ is well defined, we always assume that F is continuous, together with further (integrability) assumptions.

In most concrete cases it is natural to require that the control process u is $\mathbb{F}^X$-adapted. In other words, at time t the value $u(t)$ of the control process is only allowed to depend on past observed values of the state process X. One natural way to obtain an adapted control process is by choosing a deterministic function $\hat{u}(t,x)$

$$\hat{u} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^k,$$

and then defining the control process u by $u(t) = \hat{u}(t, X(t))$. Such a function $\hat{u}$ is called a Markov control policy[1], and in the sequel we will restrict ourselves to consider only Markov control policies.

Suppose now that we have chosen a fixed Markov control policy $\hat{u}(t,x)$. Then we can insert $\hat{u}$ into (D.1) to obtain the standard stochastic differential equation[2]

$$dX(t) = \mu(t, X(t), \hat{u}(t, X(t)))dt + \Sigma(t, X(t), \hat{u}(t, X(t)))dW(t). \tag{D.5}$$

Similarly to Definition D.1.1, we now define the class of admissible Markov control policies. By abuse of notation we use the same notation for the class of admissible control process and the class of admissible Markov control policies.

Definition D.1.3 (Admissible Markov Control policy). *A measurable deterministic function $\hat{u}$ is an admissible Markov control policy if:*

- *the control process u has the form*
$$u(t) = \hat{u}(t, X(t)); \tag{D.6}$$
- $\hat{u}(t,x) \in U_{ad}$ *for all* $t \in \mathbb{R}_+$ *and all* $x \in \mathbb{R}^n$*;*
- *for any given initial point (t,x) the stochastic differential equation for $s \in [t, \infty)$*
$$\begin{cases} dX(s) &= \mu(s, X(s), \hat{u}(s, X(s)))ds + \Sigma(s, X(s), \hat{u}(s, X(s)))dW(s), \\ X(t) &= x, \end{cases} \tag{D.7}$$
has a unique solution.

The class of admissible Markov control policies is denoted $\mathcal{U}_{ad}$.

For a given Markov control policy $\hat{u}$, the solution process X of (D.7) will of course depend on the initial value, as well as on the chosen Markov control policy $\hat{u}$. To be precise we denote the process X by $X^{\hat{u}}$. Fixing $X^{\hat{u}}(t) = x$, we will denote the unique solution $X^{\hat{u}}(s)$ of (D.7) also by $X^{\hat{u}}(s;t,x)$, $s \geq t$. We observe that if (D.6) is assumed, then $X^{\hat{u}}(s;t,x)$ is an Itô diffusion, and for all Borel measurable, bounded functions f, we have

$$E\left[f(X^{\hat{u}}(s'))|\mathcal{F}_s\right] = E\left[f(X^{\hat{u}}(s'))|X^{\hat{u}}(s)\right] = g\left(X^{\hat{u}}(s)\right)$$

for fixed s, s' such that $t \leq s \leq s'$, with[3] $g(y) := E\left[f(X^{\hat{u}}(s;t,y))\right]$, $y \in \mathbb{R}^k$.

Finally, given a Markov control policy $\hat{u}$, with the above notations, we can rewrite the dynamics (D.7) of the process $X^{\hat{u}}$ as

$$dX^{\hat{u}}(s) = \mu\left(s, X^{\hat{u}}(s), \hat{u}\big(s, X^{\hat{u}}(s)\big)\right) ds + \Sigma\left(s, X^{\hat{u}}(s), \hat{u}\big(s, X^{\hat{u}}(s)\big)\right) dW(s). \tag{D.8}$$

We now introduce the objective function of the Markov control problem, and therefore we consider as given a function $F : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ as in (D.3).

[1] In Fleming and Rischel [11], and Björk [2], $\hat{u}$ is called a feedback control law.

[2] When using Markov control policies then as we will see the solution of (D.5) is a Markov process.

[3] In the sequel the equalities analogous to (2.5) and (2.6) hold, substituting r with $X^{\hat{u}}$, i.e.,

$$E\left[f(X^{\hat{u}}(s;t,x))\right] = E_{t,x}\left[f(X^{\hat{u}}(s))\right],$$

and

$$g(y) = E_{t,y}\left[f(X^{\hat{u}}(s))\right].$$

More in general we will use the same kind of notation for functionals of the trajectory $X^{\hat{u}}(s;t,x)$, $s \geq t$.

Definition D.1.4 (Optimal Markov Control Problem)**.** *Let $\mathcal{J} : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{U}_{ad} \to \mathbb{R}$ be defined as*

$$\mathcal{J}(t, x, \hat{u}) = E_{t,x}\left[\int_t^\infty F(s, X^{\hat{u}}(s), \hat{u}(s, X^{\hat{u}}(s)))ds\right], \tag{D.9}$$

where the subscripts t and x denote that the expectation is taken using the dynamics given by (D.8) for $s \in [t, \infty)$ and $X^{\hat{u}}(t) = x$, and where F is called[4] *the (running) cost function, and $\mathcal{J}$ the total expected cost corresponding to $\hat{u}$. The optimal Markov control problem $\mathcal{P}(t, x)$ is defined as the problem of maximizing $\mathcal{J}(t, x, \hat{u})$ over $\hat{u} \in \mathcal{U}_{ad}$.*

As before, in order to ensure that $\mathcal{J}$ is well defined, we always assume that F is continuous, together with further (integrability) assumptions.

Associated to the total expected cost $\mathcal{J}$, we now define the value function (or optimal cost function).

Definition D.1.5 (Value Function and Optimal Markov Control policy)**.** *The value function*

$$J : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R},$$

is defined by

$$J(t, x) = \sup_{\hat{u} \in \mathcal{U}_{ad}} \mathcal{J}(t, x, \hat{u}).$$

Furthermore if there exists[5] *an admissible Markov control policy $\hat{u}_{sup}$ such that*

$$\mathcal{J}(t, x, \hat{u}_{sup}) = \sup_{\hat{u} \in \mathcal{U}_{ad}} \mathcal{J}(t, x, \hat{u}),$$

then we say that $\hat{u}_{sup}$ is an optimal Markov control policy [6] *for the given optimal Markov control problem $\mathcal{P}(t, x)$.*

Given an optimal control problem, there are two main problems: prove the existence of an optimal Markov control policy and determine such a policy. In many case the strategy to solve these problems consists of two steps:

1. find necessary conditions for the optimal policies,
2. if a policy $\hat{u}$ satisfies such conditions, verify that $\hat{u}$ is optimal, so that the problem of existence reduces to a verification.

Dynamics programming is the methodology we will use. The main idea is to embed our problem into a class of control problems, and then to tie all these problems together with a partial differential equation, known as the Hamilton-Jacobi-Bellman equation. The value function satisfies such equation, and the control problem is then shown to be equivalent to the problem of finding a solution to the Hamilton-Jacobi-Bellman equation. A gratifying fact is that the Hamilton-Jacobi-Bellman equation also acts as a sufficient condition for the optimal control problem, thus we obtain the existence of the optimal Markov control policy found. This result is known as the verification theorem for the dynamics programming.

For the notational convenience, we introduce the following notations. For any fixed vector $v \in U_{ad} \subseteq \mathbb{R}^k$ the partial differential operator $\mathcal{A}^v$ is defined by

$$\mathcal{A}^v f(t, x) = \sum_{i=1}^n \mu_i(t, x, v)\frac{\partial f}{\partial x_i}(t, x) + \frac{1}{2}\sum_{i,j=1}^n C_{i,j}(t, x, v)\frac{\partial^2 f}{\partial x_i \partial x_j}(t, x), \tag{D.10}$$

where μ_i, x_i represent the ith component of the vector μ, x respectively, and

$$C(t, x, v) = \Sigma(t, x, v)\Sigma'(t, x, v).$$

We recall that the partial differential operator $\mathcal{A}^v$ is called uniformly elliptic if there exists a constant $K > 0$ such that for all $(t, x, v) \in \mathbb{R}_+ \times \mathbb{R}^n \times U_{ad}$ and $\xi \in \mathbb{R}^n$

$$\sum_{i,j=1}^n C_{ij}(t, x, v)\xi_i\xi_j \geq K\left|\xi\right|^2. \tag{D.11}$$

[4] In Yong and Zhou [22], $\mathcal{J}$ is called the cost functional.

[5] Note that, as for any optimization problem, the optimal Markov control policy may not exist.

[6] In Fleming and Rischel [11], and Björk [2], $\hat{u}$ is called a optimal feedback control law.

For any Markov control policy $\hat{u}$, the partial differential operator $\mathcal{L}^{\hat{u}}$ is defined by

$$\mathcal{L}^{\hat{u}} f(t,x) = \mathcal{A}^{\hat{u}(t,x)} f(t,x), \tag{D.12}$$

and $\mathcal{L}^{\hat{u}}$ is the operator associated to the solution $X^{\hat{u}}$ of the stochastic differential equation (D.8).

We now have the following result, namely the Hamilton-Jacobi-Bellman equation.

Theorem D.1.1 (Hamilton-Jacobi-Bellman equation [11])**.** *Let $\mathcal{P}(t,x)$ be the control problem defined in Definition D.1.4. If the value function J is regular, i.e., $J \in C^{1,2}(\mathbb{R}_+, \mathbb{R}^n)$, then J satisfies the Hamilton-Jacobi-Bellman equation*

$$\frac{\partial J}{\partial t}(t,x) + \sup_{v \in U_{ad} \subseteq \mathbb{R}^k} \{F(t,x,v) + \mathcal{A}^v J(t,x)\} = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}^n, \tag{D.13}$$

with boundary condition

$$\lim_{t \to \infty} J(t,x) = 0. \tag{D.14}$$

Furthermore, if $\hat{u}_{sup}$ is optimal Markov control policy, then for each $(t,x) \in (0,\infty) \times \mathbb{R}^n$,

$$\hat{u}_{sup}(t,x) \in \arg\max_{v \in U_{ad}} \{F(t,x,v) + \mathcal{A}^v J(t,x)\}.$$

It is important to note that when (D.11) holds, results from the theory of second order nonlinear partial differential equations of parabolic type imply existence and uniqueness[7] of a solution to the problem (D.13)-(D.14), and the Hamilton-Jacobi-Bellman equation (D.13) is called uniformly elliptic. When the uniform elliptic condition (D.11) does not hold, the Hamilton-Jacobi-Bellman equation (D.13) is called degenerate parabolic type. In this case a smooth solution $J(t,x)$ cannot be expected. Instead, the value function will be interpreted as a solution in some broader sense, for instance as a generalized solution. Another convenient interpretation of the value function is as viscosity solution to the Hamilton-Jacobi-Bellman equation. Furthermore Theorem D.1.1 has the form of a necessary condition. It says that if J is the value function, and if $\hat{u}_{sup}$ is the optimal control, then J satisfies Hamilton-Jacobi-Bellman equation (D.13) and (D.14), and $\hat{u}_{sup}$ realizes the supremum in the equation. The following result belongs to a class of theorems known as verification theorems for the dynamics programming, and shows that Hamilton-Jacobi-Bellman equation (D.13) and (D.14) are also a sufficient condition for the optimal control problem.

Theorem D.1.2 (Verification Theorem [11])**.** *Suppose that we have two functions H and $\hat{u}^*$ such that*

- *$H \in C^{1,2}$, is integrable[8] and solves the Hamilton-Jacobi-Bellman equation*

  $$\frac{\partial H}{\partial t}(t,x) + \sup_{v \in U_{ad} \subseteq \mathbb{R}^k} \{F(t,x,v) + \mathcal{A}^v H(t,x)\} = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}^n, \tag{D.15}$$

  with the boundary condition

  $$\lim_{t \to \infty} H(t,x) = 0. \tag{D.16}$$

- *The function $\hat{u}^*$ is an admissible Markov control policy.*
- *For each fixed (t,x), $\hat{u}^*(t,x) \in \arg\max_{v \in U_{ad}} \{F(t,x,v) + \mathcal{A}^v H(t,x)\}$, i.e.,*

  $$F(t,x,\hat{u}^*(t,x)) + \mathcal{A}^{\hat{u}^*(t,x)} H(t,x) = \sup_{v \in U_{ad}} \{F(t,x,v) + \mathcal{A}^v H(t,x)\}.$$

  Then the following results hold.

  1. *The value function J to the control problem $\mathcal{P}(t,x)$ coincides with the function H, i.e.,*

     $$J(t,x) = H(t,x).$$

  2. *There exists an optimal Markov control policy $\hat{u}_{sup}(t,x)$, and coincides with $\hat{u}^*$, i.e.,*

     $$\hat{u}_{sup}(t,x) = \hat{u}^*(t,x).$$

[7] A sufficient condition for the uniqueness is that (D.14) holds.

[8] The assumption that H is integrable is made in order to guarantee that the Dynkin formula (used in the proof) holds. This will be the case if, for example, H satisfies the condition $H_x(s, X^{\hat{u}}(s))\Sigma(t, X^{\hat{u}}(s), \hat{u}(s, X^{\hat{u}}(s))) \in \mathcal{L}^2(0,\infty;\mathbb{F})$, for all admissible Markov control policy.

Thanks to the previous results, the optimal Markov control problem $\mathcal{P}(t,x)$ may be solved by using the Hamilton-Jacobi-Bellman equation with the corresponding boundary condition, as in the following scheme.

1. Fixed an arbitrary point $(t,x) \in (0,\infty) \times \mathbb{R}^n$ and any function $H(t,x)$ sufficiently smooth, we can solve the optimization problem

$$\max_{v \in U_{ad}} \{F(t,x,v) + \mathcal{A}^v H(t,x)\}, \tag{D.17}$$

   where the partial differential operator $\mathcal{A}^v$ is given by (D.12). Note that in (D.17) v is the only variable, whereas t and x are considered to be fixed parameters. The functions F, μ, Σ are given.

2. If the maximum in (D.17) is attained in a unique point v, the optimal choice of v will of course depend on our choice of t and x, but it will also depend on the function H, and its partial derivatives, which appear in $\mathcal{A}^v H$, and the argument of the maximum in (D.17) is denoted by $\hat{u}^*(t,x;H) = \hat{u}^*_H(t,x)$ (more in general $\hat{u}^*_H(t,x) \in \arg\max_{v \in U_{ad}} \{F(t,x,v) + \mathcal{A}^v H(t,x)\}$).

3. The function $\hat{u}^*_H(t,x)$ is our candidate for the optimal Markov control policy. If we knew the value function J, then our candidate would be $\hat{u}^*_J(t,x)$, but we do not know J and therefore we substitute the expression for $\hat{u}^*_H(t,x)$ of the previous point 2 into (D.15), giving us the partial differential equation

$$\frac{\partial H}{\partial t}(t,x) + F(t,x,\hat{u}^*_H(t,x)) + \mathcal{L}^{\hat{u}^*_H} H(t,x) = 0. \tag{D.18}$$

4. Now we solve the partial differential equation (D.18) under condition (D.16), and we assume that H^* is a classical solution. Then we can use the Verification Theorem D.1.2 with $H = H^*$, $\hat{u}^* = \hat{u}^*_{H^*}(t,x)$, and conclude that $J = H^*$, $\hat{u}_{sup} = \hat{u}^*_{H^*}$.

The hard work of dynamic programming consists in solving the non linear partial differential equation (D.18). There are no general analytic methods available for this, so the number of known optimal control problems with an analytic solution is very small indeed.

Observe that if the hypotheses of the Verification Theorem D.1.2 do not hold, then we cannot follow the above scheme. When ellipticity condition (D.11) holds, then we have existence and uniqueness for the problem (D.13)-(D.14) and in this case we have a smooth solution $J(t,x)$.

Bibliography

[1] Enrico Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance Math. Econom.*, 37(3):443–468, 2005.

[2] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, first edition, 1998.

[3] D. Blake and W. Burrows. Survivor bonds: Helping to hedge mortality risk. *Journal of Risk and Insurance*, 68(3..):339–348, February.. 2001.

[4] D. Brigo and A. Alfonsi. Credit default swaps calibration and option pricing with the ssrd stochastic intensity and interest-rate model. *Working paper at www.damianobrigo.it and reduced version in Proceedings of the* 6*-th Columbia-JAFEE Conference, Tokyo*, March 2003.

[5] D. Brigo and A. Alfonsi. Credit default swaps calibration and derivatives pricing with the ssrd stochastic intensity model. *Finance and Stochastics*, 9(9):29–42, March 2005.

[6] P. Brémaud. *Point processes and queues*. Springer-Verlag, first edition, 1981.

[7] G. Deelstra and F. Delbaen. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Date Analysis*, 14(1):77–84, March 2008.

[8] D. Duffie. *Dynamics Asset Pricing Theory*. Princeton University Press, third edition, 2001.

[9] D. Duffie and K.J. Singleton. *Credit risk*. Princeton University Press, first edition, 2003.

[10] First International Conference on Longevity Risk and Capital Markets Solutions. *The longevity bond*, February 2005.

[11] W. H. Fleming and R. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, first edition, 1975.

[12] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, second edition, 2005.

[13] M. Hardy. A matter of life and death. *Financial Engineering News*, 46(3..):17–20, February.. 2005.

[14] J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, June 1979.

[15] J. M. Harrison and S. R. Pliska. A stochastic calculus model of continuous trading: complete market. *Stochastic Processes and their Applications*, 15(3):313–316, August 1983.

[16] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, second edition, 1991.

[17] Elisa Luciano, Jaap Spreeuw, and Elena Vigna. Modelling stochastic mortality for dependent lives. *Insurance Math. Econom.*, 43(2):234–244, 2008.

[18] F. Menoncin. The role of longevity bonds in optimal portfolios. *Insurance: Mathematics and Economics*, 42(3):343–358, February 2008.

[19] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, sixth edition, 2003.

[20] M. Rutkowski. Self-financing trading strategies for sliding, rolling-horizon, and consol bonds. *Insurance: Mathematical Finance*, 9(4):361–385, October 1999.

[21] Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer-Verlag, first edition, 2004.

[22] J. Yong and X.J. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer-Verlag, first edition, 1999.