Risk-neutral valuation of GLWB riders in variable annuities

Anna Rita Bacinello [a], Rosario Maggistro [a,*], Ivan Zoccolan [b]

[a] University of Trieste, Department of Economics, Business, Mathematics and Statistics "Bruno de Finetti", Trieste, Italy
[b] Generali Italia S.p.a., Milano, Italy

A B S T R A C T

In this paper we propose a model for pricing GLWB variable annuities under a stochastic mortality framework. Our set-up is very general and only requires the Markovian property for the mortality intensity and the asset price processes. The contract value is defined through an optimization problem which is solved by using dynamic programming. We prove, by backward induction, the validity of the bang-bang condition for the set of discrete withdrawal strategies of the model. This result is particularly remarkable as in the insurance literature either the existence of optimal bang-bang controls is assumed or it requires suitable conditions. We assume constant interest rates, although our results still hold in the case of a Markovian interest rate process. We present extensive numerical examples, modelling the mortality intensity as a non mean reverting square root process and the asset price as an exponential Lévy process, and compare the results obtained for different parameters and policyholder behaviours.

1. Introduction

Variable annuities (VAs) are very flexible life insurance investment products that package living and death benefits endowed with a number of possible guarantees in respect of financial or biometric risks. Usually, a lump-sum premium is paid when the product is bought, and is invested in well-diversified mutual funds chosen by the policyholder among a range of alternative opportunities. This initial investment establishes a reference portfolio (*policy account*), and all guarantees are financed through periodical deductions from the policy account value.

The market for variable annuities has shown phenomenal growth since the 1990s: according to Chopra et al. (2009), the annual rate of growth in the U.S. VA market during that decade was 21%, and the level of total assets reached almost USD 1 trillion by 2001. A good overview of VA products and the development of their market can be found, e.g., in Bauer et al. (2008) and Ledlie et al. (2008). By the end of 2019, their net assets level was around USD 2 trillion and constituted the largest class of liabilities for U.S. life insurers (see Moenig (2021)). In 2021, total VA sales amounted to USD 125.6 billion, 27% higher than prior year and 49.29% of the total U.S. annuity sales (that includes fixed annuities).[1]

A possible rider that can be included in a VA contract in order to provide a post-retirement income is the Guaranteed Lifelong Withdrawal Benefit (GLWB), that offers a lifelong withdrawal guarantee. Therefore, there is no explicit limit on the total amount that can be withdrawn over the term of the policy because, if the policy account is depleted while the policyholder is still alive, she continues to receive the guaranteed amount until death. Any remaining account value at the time of death is paid to the beneficiary as a death benefit. The GLWB take-up rate was 72% in 2015.[2]

The guaranteed withdrawal is computed by applying a fixed percentage (*withdrawal rate*) to the so called *benefit base* (also referred to as *base amount*, or *withdrawal base*), that is initially set equal to the value of the investment portfolio and then evolves according to the contractually stated rules. A plain GLWB can be enriched by adding elements such as a *ratchet* (or *step-up*) feature and a *roll-up* (or *bonus*) feature. Under the ratchet provision the benefit base is increased to the policy account value on predetermined (ratchet) dates if the latter exceeds the previous benefit base recorded on the last withdrawal date. Under the roll-up provision, the benefit base may also be increased by a proportional amount if the policyholder chooses not to withdraw on a withdrawal

* Corresponding author.
E-mail addresses: bacinel@units.it (A.R. Bacinello), rosario.maggistro@deams.units.it (R. Maggistro), ivanzoccolan@gmail.com (I. Zoccolan).

[1] Source: LIMRA, *U.S. Individual Annuities survey*, https://www.limra.com/en/newsroom/news-releases/2022/secure-retirement-institute-total-annuity-sales-jump-16-in-2021--marking-highest-sales-since-2008/.

[2] Source: Society of Actuaries and LIMRA, *Variable Annuity Guaranteed Living Benefits Utilization* (page 35, last paragraph), https://www.soa.org/493476/globalassets/assets/files/resources/research-report/2018/variable-annuity-guaranteed-utilization.pdf.

date. Therefore, roll-ups are commonly used as a disincentive to withdraw during the first years of contract. Moreover, complete surrender refers to the withdrawal of the whole policy account, if its value exceeds the guaranteed withdrawal amount. The actual withdrawal can be lower than the guaranteed amount, or even exceed it, but in this second case the net amount received by the policyholder is subject to a penalty, that can be accompanied by further disincentives. There are also deferred versions of the contract, as opposed to immediate withdrawals, where the policyholder can only withdraw money after the deferred period. Details on additional features embedded in GLWB can be found, e.g., in Piscopo and Haberman (2011).

There has been a number of papers dealing with pricing of the VA products. Most of them are focused on pricing VA guarantees under the *static* policyholder behaviour in withdrawal and surrender (see e.g., Milevsky and Salisbury, 2006), meaning that the policyholder always withdraws exactly the guaranteed amount, and never surrenders the contract. Some studies include pricing under the *dynamic* approach, when the policyholder optimally decides the amount to withdraw at each withdrawal date depending on the information available at that date (see, e.g., Steinorth and Mitchell, 2015). According to whether withdrawals are assumed to occur continuously or discretely, the dynamic approach is usually tackled employing, respectively, stochastic control (e.g. Chen and Forsyth, 2008 and Dai et al., 2008) and dynamic programming (Steinorth and Mitchell, 2015, Huang et al., 2017, Fontana and Rotondi, 2023, Bacinello et al., 2016 and Alonso-García et al., 2018). Recently, dynamic programming was applied to the valuation of a large VA portfolio with different guarantee types and policyholder characteristics (see Moenig, 2021). A range of numerical methods are also used, including standard and regression-based Monte Carlo methods (Bacinello et al., 2011 and Huang and Kwok, 2016), Partial Differential Equation (PDE) and direct integration methods (Chen and Forsyth, 2008, Dai et al., 2008, Luo and Shevchenko, 2015a, Luo and Shevchenko, 2015b, Forsyth and Vetzal, 2014, Shevchenko and Luo, 2017). A comprehensive overview of numerical methods for the pricing of VA guarantees is provided in Shevchenko and Luo (2016).

The main contribution of our paper is the definition of a pricing model for GLWB variable annuities within a stochastic mortality framework that allows to capture systematic longevity improvements as well as mortality shocks due, e.g., to pandemics. A very general setting which only requires the Markovian property for the mortality intensity and the asset price processes is considered. Following a risk-neutral perspective in an incomplete market, we define the initial contract value, under the dynamic approach, through an optimization problem that is solved by exploiting dynamic programming. We adopt the same approach proposed by Bacinello et al. (2016), that construct a dynamic programming algorithm for the valuation of a Guaranteed Minimum Withdrawal Benefit (GMWB)[3] in a VA contract, using exponential Lévy processes to model the evolution of the reference asset price and incorporating alternative policyholder behaviours as special cases in the dynamic programming algorithm. In particular, in the numerical implementation of our model we assume a non mean reverting square root process to describe the stochastic evolution of the mortality intensity (see e.g., Fung et al., 2014 and Dacorogna and Apicella, 2016), and, as in Bacinello et al. (2016), an exponential Lévy process for the asset price. In this way we get a flexible stochastic model for pricing and risk management of the GLWB.

The valuation of GLWB VAs with dynamic withdrawals and through dynamic programming was also studied in Steinorth and Mitchell (2015) and Huang et al. (2017). In Steinorth and Mitchell (2015), a lifecycle model is set up, where risk-averse individuals maximize the expected utility of the future cash-flows generated by their VA contract, so as to balance longevity protection with financial returns. In particular, the authors use a Geometric Brownian Motion (GBM) to model the investment returns, and find that an investment in a GLWB VA dominates that in a mutual fund, in terms of Money's Worth Ratio, but is dominated by a fixed annuity. Unlike almost all the papers dealing with the GLWB rider, in which withdrawals (possibly $= 0$) take place immediately, in Huang et al. (2017) the withdrawal phase is preceded by an accumulation phase, during which the policyholder can make additional investments in the VA, up to a fixed limit. Then these investments, whose returns are modelled with the Heston stochastic volatility process (Heston, 1993), can be seen as negative withdrawals. The end of the accumulation phase and the beginning of the withdrawal one is chosen by the policyholder, and surrender is always admitted.

Other papers to mention in the context of GLWB are Holz et al. (2012), Kling et al. (2014) and Bacinello et al. (2012), although their contract valuation is not made under a (fully) dynamic approach. In particular, Holz et al. (2012) analyse different product designs, with roll-up or ratchet provisions, under the GBM dynamics of the underlying fund value process, and compare numerically the contract values, obtained by Monte Carlo simulation, corresponding to fixed withdrawal behaviours. These behaviours include either *deterministic* strategies, or a mixture of them (that the authors call *probabilistic*), or *stochastic* strategies based on the value of state variables (e.g., on the moneyness of the guarantees), i.e., in the language of stochastic processes, *adapted* strategies. As opposed to our dynamic approach, their contract value is not the result of an optimization process. The authors also formalize the optimization problem under the dynamic approach, but do not solve it stating that the corresponding optimal strategy is not obvious and cannot be easily determined. For this reason they limit themselves to describe how Monte Carlo simulation can be used to approximate this strategy. After, they argue that the optimal strategy for a GLWB contract without ratchet and roll-up features consists only of either a withdrawal of the guaranteed amount or complete surrender, that is what in Bacinello et al. (2011) is called the *mixed* approach. Taking this into account, they compute the contract value using a recursive algorithm designed to determine the optimal surrender boundary. In Kling et al. (2014) the previous analysis is extended and includes also hedging, not only pricing, of the guarantees. Concerning, in particular, the pricing, the contract value under the mixed approach is computed by using the Least Squares Monte Carlo (LSMC) method (see Longstaff and Schwartz, 2001), and the Heston volatility model for the fund value is employed as well. Also Bacinello et al. (2012) value the GLWB under the mixed approach, by employing the LSMC technique implemented with a very sophisticated model framework.

Another important contribution of our paper is the verification, through backward induction, of the bang-bang condition for the set of discrete withdrawal strategies of the GLWB model. This means that the set of the optimal withdrawals consists of three choices only: zero withdrawal, withdrawal at the contractual amount, complete surrender. This result, proven in our discrete time framework for the evolution of the policy account value, is particularly remarkable as in the insurance literature either the existence of optimal bang-bang controls is assumed or it requires suitable conditions. For example, in Azimzadeh and Forsyth (2015) a GBM for the reference asset price, as well as the convexity and monotonicity of the contract value function for all times, are assumed to ensure the existence of optimal bang-bang strategies. A regression-based Monte Carlo method for pricing GLWB under the bang-bang strategy in the case of stochastic volatility is also developed in Huang and Kwok (2016), while in the more general model by Huang et al. (2017) again the convexity preservation of the price of any convex terminal payoff function is required, as well as a suitable scaling property.

The bang-bang condition, beyond drastically reducing the computational time needed to search the optimal withdrawal in the backward recursive step of our dynamic algorithm, allows to clearly separate the various contract components: the first one is the *basic GLWB contract* value, obtained under the static approach, the second is the value of

[3] It differs from the GLWB in that withdrawals take place for a fixed duration, regardless of the policyholder's survival.

the *surrender option*, and the last component is the value of the *roll-up option*.

The paper is structured as follows. In Section 2, we describe the structure of the variable annuity contract. In Section 3, we introduce our valuation framework and define the optimal withdrawal problem. In Section 4, we first define the dynamic programming equations which allow to solve the problem, then we verify the validity of the bang-bang condition for the strategy space of discrete withdrawal policies of the GLWB model, after we show how the previous results can be adapted in order to deal with the case of stochastic interest rates, then we outline our valuation algorithm and finally present the contract decomposition. In Section 5, we perform a sensitivity analysis comparing the numerical results obtained for different parameters and policyholder behaviours. Section 6 concludes the paper.

2. The contract structure

In this section we describe the GLWB rider in our variable annuity contract. At time 0 (contract inception), the policyholder, aged x, pays a single premium P which is entirely invested in a well-diversified and non-dividend paying mutual fund of her own choice. We denote by S_t the market price at time t of each unit of this fund, that drives the return on the investment portfolio built up with the policyholder's payment. The value at time t of such portfolio, that is called *personal account* (or *policy account*) is denoted by W_t.

The GLWB rider gives the policyholder the right to make periodical withdrawals from her personal account at some specified dates for the whole life, even if the account value is reduced to zero. The cost of the guarantee is financed by periodical proportional deductions from the policy account value (*insurance fees*). The guaranteed withdrawal amount is calculated as a fixed proportion g (*withdrawal rate*) of the *benefit base* (or *withdrawal base*, or *base amount*), denoted by A_t. The benefit base is initially set equal to the single premium and is adjusted upward via the *roll-up* (or *bonus*) feature, that applies when no withdrawal is made on a specified withdrawal date.

Complete surrender of the policy refers to the withdrawal of the whole personal account, if it exceeds the guaranteed amount. As a result, the variable annuity contract terminates. Another event that causes the closure of the contract is the policyholder's death. The value that remains in the policy account when the policyholder dies is paid to the beneficiary as a death benefit.

In particular, from now on we assume that: (i) withdrawals are allowed on a predetermined set of equidistant dates and, without loss of generality, we take the distance between two consecutive withdrawal dates as unit of measurement of time; (ii) the death benefit is paid to the beneficiary on the next upcoming withdrawal date.

We now formalize what just described. Let τ denote the time of death of the policyholder, so that withdrawals are allowed only at times $i = 1, 2, \dots$, provided that $\tau > i$. The guaranteed amount that can be withdrawn at time i is equal to gA_i, $i = 1, 2, \dots$, and the return on the reference fund over the interval $[i-1, i]$ is

$$R_i = \frac{S_i}{S_{i-1}} - 1, \quad i = 1, 2, \dots . \tag{1}$$

As already mentioned in Section 1, the policyholder is not obliged to withdraw the guaranteed amount, but she can decide to withdraw less than this amount, or even more if not exceeding the policy account value. We denote by y_i the actual withdrawal made by the policyholder at time i. Hence, under our dynamic approach, we assume that the set of possible withdrawals at this time is given by the interval $[0, \max\{gA_i, W_i\}]$.[4] If the policyholder does not withdraw anything at time i, the benefit base is proportionally increased according to the roll-up rate, that we denote by b_i (with $0 < b_i < 1$), while, if the withdrawal exceeds gA_i, it is proportionally reduced according to the so called 'pro-rata' adjustment rule. Then the benefit base evolves as follows:

$$A_{i+1} = f_{i+1}^{A}(W_i, A_i, y_i) = \begin{cases} A_i(1+b_i) & \text{if} \quad y_i = 0, \\ A_i & \text{if} \quad 0 < y_i \le gA_i, \\ A_i \dfrac{W_i - y_i}{W_i - gA_i} & \text{if} \quad gA_i < y_i \le W_i \end{cases}, \quad i = 1, 2, \dots, \tag{2}$$

with $A_1 = A_0 = P$.[5]

Moreover, in case of withdrawals exceeding the guaranteed amount, there is also a proportional penalization on the surplus according to a penalty rate, that we denote by k_i (such that $0 \le k_i < 1$). Therefore, the net amount (cash-flow) received by the policyholder at time i is given by

$$\begin{aligned} B_i^{(s)} = f_i^{(s)}(y_i, A_i) &= \begin{cases} y_i & \text{if} \quad 0 \le y_i \le gA_i \\ gA_i + (1-k_i)(y_i - gA_i) & \text{if} \quad y_i > gA_i \end{cases} \\ &= y_i - k_i \max\{y_i - gA_i, 0\}, \quad i = 1, 2, \dots . \end{aligned} \tag{3}$$

The policy account value evolves according to the following equation:

$$W_{i+1} = f_{i+1}^{W}(W_i, R_{i+1}, y_i) = \max\{W_i - y_i, 0\}(1 + R_{i+1})(1 - \varphi), \quad i = 0, 1, \dots, \tag{4}$$

where φ (such that $0 < \varphi < 1$) is the insurance fee rate, $W_0 = P$ and $y_0 = 0$. Note that 0 is an absorbent barrier for W because, once it becomes null, it remains so for ever. The contract, however, continues while $A_t > 0$ (and the insured is still alive).

Finally, in case of death in the time interval $(i-1, i]$, the death benefit, paid at time i, is

$$B_i^{(d)} = W_i, \quad i - 1 < \tau \le i, \quad i = 1, 2, \dots . \tag{5}$$

We notice that, in case of surrender at time i, i.e., when $y_i = W_i > gA_i$, the contract is automatically closed because Equations (2) and (4) imply $A_t = W_t = 0$ for all $t > i$, hence no further withdrawals are admitted, nor a death benefit will be paid.

3. The valuation framework

We assume to act in a frictionless market where all agents are price-takers, the assets are perfectly divisible and there are no short-sale constraints, no taxes or transaction costs. Moreover, we assume that markets are arbitrage-free, so that equivalent martingale measures (or risk-neutral, or risk-adjusted, measures) do exist. However, insurance markets are typically incomplete, especially when one acts at single contract level rather than at portfolio one, since it is impossible to replicate cash-flows depending on the time of death of a specific insured. Sometimes market completeness is assumed at portfolio level, when the insurance portfolio is well diversified away because the insurer has been able to sell a sufficiently high number of contracts, but this cannot be the case when instead sources of systematic risk, e.g. systematic longevity risk, are involved. Then the risk-neutral measures are infinitely many, and we assume that the insurer selects one among them, according to his preferences, for pricing purposes. In particular, the absence of transaction costs implies that there are no management fees or, more generally, no expense loadings of the premium. The cost

[4] Under the static approach this set is the singleton $\{gA_i\}$ while, under the mixed approach, it is again the singleton $\{gA_i\}$ if $gA_i \ge W_i$, and the two-points set $\{gA_i, W_i\}$ otherwise; see also Bacinello et al. (2016).

[5] Note that A_0 is only a fictitious value because we assume that no withdrawals can be made at contract inception and that the roll-up feature does not apply between times 0 and 1.

of all guarantees is recovered through the application of the insurance fee and, partly, also through the penalties on withdrawals exceeding the guaranteed amount or surrender. Of course this does not mean that there are no safety loadings (or risk-premiums) embedded in the contract, but this implicitly derives from the fact that we put ourselves in a risk-neutral framework, where the insurer computes expectations not under the physical measure but under a suitable risk-adjusted one.

Given the above, we consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ supporting all sources of financial and biometric uncertainty, where all random variables and processes are defined. The filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, that represents the flow of information available to the insurer and the policyholder over time, satisfies the usual conditions of right continuity and completeness, and is such that $\mathcal{F}_0$ is Q-trivial. Q is the risk-neutral probability measure selected by the insurer for pricing the contract.

In this setting, the residual lifetime of the policyholder τ is a stochastic $\mathbb{F}$-stopping time. Let $\mu_t := \mu_{x+t}(t)$ be the mortality intensity which determines the probability of death at time t conditional on survival for the policyholder aged x at time 0. Letting moreover

$$\mathcal{G}_t^S = \sigma(S_u, u \leq t), \quad \mathcal{G}_t^\tau = \sigma(\mathbf{1}_{\{\tau \leq u\}}, u \leq t), \quad \mathcal{G}_t^\mu = \sigma(\mu_u, u \leq t),$$

the natural filtrations associated to the fund value process S_t, the death indicator process $\mathbf{1}_{\{\tau\leq t\}}$ and the mortality process μ_t, respectively, we take the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ as

$$\mathcal{F}_t = \mathcal{G}_t^S \vee \mathcal{G}_t^\tau \vee \mathcal{G}_t^\mu,$$

with $\mathcal{G}_t^S$ independent of $\mathcal{G}_t^\tau \vee \mathcal{G}_t^\mu$. In other words, there is independence between financial- and biometric-related variables. Therefore, we can define

$$Q\left(\tau > s | \tau > t, \mathcal{G}_t^\mu\right) = \mathbb{E}^Q\left[\mathrm{e}^{-\int_t^s \mu_u \,\mathrm{d}u} | \mathcal{G}_t^\mu\right], \quad 0 \leq t < s \tag{6}$$

as the probability of surviving $s-t$ periods for the policyholder aged $x+t$ at time t, considering all the information about the mortality until this time. With the usual actuarial notation this probability should be indicated with ${}_{s-t}p_{x+t}(t)$. However, with specific reference to the one-period death and survival probabilities, we use slightly different symbols. In particular, assuming for μ a (strictly positive) Markovian process, we denote by

$$p_i(\mu_i) = Q\left(\tau > i+1 \,|\, \tau > i, \mu_i\right) = \mathbb{E}^Q\left[\mathrm{e}^{-\int_i^{i+1} \mu_u \,\mathrm{d}u} | \mu_i\right], \quad i = 0, 1, \ldots, \tag{7}$$

the probability of survival up to $i+1$ for the policyholder aged $x+i$ given the mortality intensity's values up to i. Consequently, $q_i(\mu_i) = 1 - p_i(\mu_i)$ is the probability of death before $i+1$ conditional on survival at time i.

Concerning the financial uncertainty, in order to keep the curse of dimensionality of our valuation algorithm manageable, we assume the instantaneous interest rate to be deterministic and constant, and denote it by r. Nonetheless, all the results reported in the next section, in particular the dynamic valuation procedure and the bang-bang analysis, can be straightforwardly extended to the case of a Markovian interest rate stochastic process $(r_t)_{t\geq 0}$ (see Section 4.2).

Also the reference price S, in principle, could be any Markovian process whose discounted value is a martingale under Q. However, just to fix the ideas, we restrict ourselves to the flexible and wide class of exponential Lévy processes, i.e., we assume that

$$S_t = S_0 \mathrm{e}^{(r+d)t + X_t}, \tag{8}$$

where $(X_t)_{t\geq 0}$, with $X_0 = 0$, is a Lévy process,[6] and d represents an adjustment so that $S_t \mathrm{e}^{-rt}$ is a martingale under Q. Hence X_t is right-continuous with left limits paths, $X_s - X_t$ is independent of $(X_u)_{0\leq u\leq t}$ and is distributed as X_{s-t}, for $0 \leq t < s$. Moreover, X_t can be obtained as a combination of a linear drift, a Brownian motion, and a jump component, and is completely determined by its characteristic function

$$\Phi_t(u) := \mathbb{E}^Q\left[\mathrm{e}^{iuX_t}\right] = [\Phi_1(u)]^t, \tag{9}$$

where i is the unit imaginary number. Then the drift adjustment d is given by

$$d = -\frac{1}{t} \ln \Phi_t(-i) = -\ln \Phi_1(-i). \tag{10}$$

All moments of X_t can be numerically recovered from the knowledge of Φ_t (when not available in closed form), and if Φ_t is integrable, then X_t has density given by:

$$f_t(\phi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-iz\phi} \Phi_t(z) \mathrm{d}z. \tag{11}$$

Now we want to define the initial value of the GLWB variable annuity. As already mentioned at the beginning of this section, markets are incomplete and arbitrage-free, so that this value is not uniquely determined through the (common) value of some replicating strategy. Moreover, in addition to the uncertainty driven by biometric and financial risk factors, the cash-flows generated by the contract, and hence their value, crucially depend on the policyholder behaviour, that the insurer would like to anticipate correctly. However, this is a very hard task. Even under the assumption of a fully dynamic policyholder who makes her withdrawal choices by pursuing some optimization criterion, her optimization problem can be formulated, e.g., as an expected utility maximization, under the physical measure (Steinorth and Mitchell, 2015), or as the maximization of the expected present value of her future cash-flows under a "subjective" risk-neutral measure, not necessarily coinciding with Q. The uncertainty about the preferences of the policyholder (i.e., the specific utility function, or the subjective risk-neutral measure) is further amplified in the presence of information asymmetry concerning, in particular, her health status. To overcome this trouble we assume that the insurance company fixes the measure Q that, in some way, translates its preferences, and computes the initial contract value by maximizing the expected present value, under Q, of all future cash-flows produced by the variable annuity. Hence we follow a risk-neutral approach in an incomplete market.[7] We remark that the withdrawal strategy solving this maximization problem does not necessarily lead to the optimization of the policyholder's preferences, but leads to the worst-case scenario from the insurance company's point of view.

To formally translate our problem we introduce a withdrawal strategy $y = (y_i)_{i\in\mathbb{N}^+}$, where y_i denotes the actual withdrawal made at time i (in case of survival). This is a stochastic process, adapted to the filtration $\mathbb{F}$, because at each withdrawal date the policyholder takes her withdrawal decision once she knows the values of all state variables. This strategy is *admissible* if it belongs to the set of admissible withdrawal strategies $Y = (Y_i)_{i\in\mathbb{N}^+}$ that, as already mentioned, is the sequence of intervals $Y_i = \left[0, \max\{W_i, gA_i\}\right]$. Then we define the initial value of the GLWB variable annuity as the solution of the following optimization problem:

$$V_0 = \sup_{y\in Y} \mathbb{E}^Q\left[\sum_{i=1}^{\infty} \mathrm{e}^{-ri}\left(\mathbf{1}_{\{\tau>i\}} f_i^{(s)}(y_i, A_i) + \mathbf{1}_{\{i-1<\tau\leq i\}} W_i\right)\right], \tag{12}$$

where the account value and the benefit base satisfy Equations (4) and (2) respectively.

Note that the policyholder pays the lump-sum premium P at the contract inception. The contract value V_0 depends on P through the initial level of the benefit base and the policy account, but does not

[6] For a comprehensive description of Lévy processes, their properties and applications we refer to Tankov and Cont (2003) and Schoutens (2003).

[7] For an in-depth and comprehensive analysis of policyholders' behaviour and categorization of the valuation approaches proposed in the literature see Bauer et al. (2017).

necessarily coincide with it. If this happens, we say that the contract is *fairly priced* and, recalling the above remark concerning the worst-case scenario, in this case the insurance company puts itself on the safety side. To produce a contract value exactly equal to P a suitable choice of the set of contract parameters should be made. In particular, the key-parameter usually chosen in such a way that the contract is fairly priced, once all the other parameters are fixed, is the insurance fee rate φ. For this reason the particular value of φ which makes $V_0 = P$ from now on will be referred to as *fair fee rate*.

4. Dynamic programming

In order to solve (12), in this section we implement a dynamic programming algorithm for discrete stochastic control problems (see, e.g., Bertsekas, 2005 and Seierstad, 2009). As we act in a Markovian framework, for each $i = 0, 1, \dots$ and each value of policy account W_i, benefit base A_i and mortality intensity μ_i, we denote by $V_i(W_i, A_i, \mu_i)$ the contract value at time i (before the periodic withdrawal) and by $v_i(W_i, A_i, \mu_i)$ the contract value at the same time when, moreover, the policyholder is then alive. Clearly $V_i(W_i, A_i, \mu_i) = \mathbf{1}_{\{\tau>i\}} v_i(W_i, A_i, \mu_i)$ and $V_0 = V_0(P, P, \mu_0) = v_0(P, P, \mu_0)$.

Since the algorithm proceeds backward, we need a starting point. To this end, as is common in life insurance mathematics, we assume that there is an ultimate age for the policyholder beyond which her survival probability is null. We denote by ω this age, that typically is in the range 110-120 years, and let $n = \lfloor \omega - x \rfloor$, hence $n \le \omega - x < n + 1$.[8] Then Problem (12) can be rewritten as

$$V_0 = \sup_{y\in Y} \mathbb{E}^Q \left[\sum_{i=1}^{n} \mathrm{e}^{-ri} \left(\mathbf{1}_{\{\tau>i\}} f_i^{(s)}(y_i, A_i) + \mathbf{1}_{\{i-1<\tau\le i\}} W_i \right) + \mathrm{e}^{-r(n+1)} \mathbf{1}_{\{\tau>n\}} W_{n+1} \right]. \tag{13}$$

We take $n+1$ as starting point of our backward dynamic algorithm, and define the following terminal condition:

$$v_{n+1}(W_{n+1}, A_{n+1}, \mu_{n+1}) \equiv 0. \tag{14}$$

Note that the particular value of v_{n+1} is quite irrelevant because it will be multiplied for the survival indicator $\mathbf{1}_{\{\tau>n+1\}}$, almost surely $= 0$, in order to get the contract value V_{n+1}.

Then, to proceed backward, we define the Bellman recursive equation of the problem as follows:

$$\begin{aligned} &v_i(W_i, A_i, \mu_i) \\ &= \sup_{y_i\in Y_i} \Bigg(f_i^{(s)}(y_i, A_i) + \mathbb{E}^Q\left[\mathbf{1}_{\{\tau\le i+1\}} f_{i+1}^W(W_i, R_{i+1}, y_i) \mathrm{e}^{-r} | W_i, A_i, \mu_i, \tau > i \right] \\ &\quad + \mathbb{E}^Q\Big[V_{i+1}\big(f_{i+1}^W(W_i, R_{i+1}, y_i), \\ &\qquad f_{i+1}^A(W_i, A_i, y_i), \mu_{i+1}\big) \mathrm{e}^{-r} | W_i, A_i, \mu_i, \tau > i \Big] \Bigg), \\ &\qquad i = n, n-1, \dots, 1. \end{aligned} \tag{15}$$

In the first expectation of (15) we can exploit the stochastic independence between financial and demographic factors, the definition of the policy account value in (4), and the martingale property according to which $\mathbb{E}^Q\left[(1+R_{i+1})\mathrm{e}^{-r}\right] = 1$, to get

$$\begin{aligned} &\mathbb{E}^Q\left[\mathbf{1}_{\{\tau\le i+1\}} f_{i+1}^W(W_i, R_{i+1}, y_i) \mathrm{e}^{-r} | W_i, A_i, \mu_i, \tau > i \right] \\ &= q_i(\mu_i) \max\{W_i - y_i, 0\}(1-\varphi). \end{aligned} \tag{16}$$

In the second expectation we first replace $V_{i+1}(\dots)$ with $\mathbf{1}_{\{\tau>i+1\}} v_{i+1}(\dots)$, then we further condition on μ_u, $i < u \le i+1$, so that we can factorize the conditional Q-expectation inside, and after apply the tower property, obtaining

$$\begin{aligned} &\mathbb{E}^Q\left[V_{i+1}\left(f_{i+1}^W(W_i, R_{i+1}, y_i), f_{i+1}^A(W_i, A_i, y_i), \mu_{i+1}\right) \mathrm{e}^{-r} | W_i, A_i, \mu_i, \tau > i \right] \\ &= \mathbb{E}^Q\Big[\mathrm{e}^{-\int_i^{i+1} \mu_u \,\mathrm{d}u} v_{i+1}\big(f_{i+1}^W(W_i, R_{i+1}, y_i), \\ &\qquad f_{i+1}^A(W_i, A_i, y_i), \mu_{i+1}\big) \mathrm{e}^{-r} | W_i, A_i, \mu_i \Big]. \end{aligned} \tag{17}$$

Finally, the initial contract value is given by

$$\begin{aligned} v_0(P, P, \mu_0) &= q_0(\mu_0) \mathbb{E}^Q\left[f_1^W(P, R_1, 0) \mathrm{e}^{-r} \right] \\ &\quad + \mathbb{E}^Q\left[\mathrm{e}^{-\int_0^1 \mu_u \,\mathrm{d}u} v_1(f_1^W(P, R_1, 0), P, \mu_1) e^{-r} \right] \\ &= q_0(\mu_0) P(1-\varphi) \\ &\quad + \mathbb{E}^Q\left[\mathrm{e}^{-\int_0^1 \mu_u \,\mathrm{d}u} v_1(P(1+R_1)(1-\varphi), P, \mu_1) e^{-r} \right]. \end{aligned} \tag{18}$$

4.1. Bang-bang analysis

The Bellman recursive equation (15) requires, at each time step $i = n, n-1, \dots 1$, to solve a real-valued optimization problem where the domain of the single variable y_i is the interval $Y_i = [0, \max\{W_i, gA_i\}]$. Even if withdrawals can take place only once a year, for, e.g., a 65-years old policyholder this would imply around 50 time steps. Moreover, at each time step the problem must be solved for every possible triplet of state variables (W_i, A_i, μ_i). Then the computational effort could be substantial. A property that would drastically reduce this effort is the *bang-bang* condition, which states that the set of the optimal withdrawals consists of three choices only: zero withdrawal, withdrawal at the contractual amount, complete surrender, i.e., the research of the maximum can be restricted to the subset $\{0, gA_i, W_i\}$ of Y_i.

The intuition behind this condition is clear: due to the roll-up feature it may be convenient to withdraw nothing ($y_i = 0$) in order to increase the benefit base for future withdrawals, especially if the expected residual lifetime of the policyholder is sufficiently long, but it is never convenient to withdraw something less than the guaranteed amount, i.e., $0 < y_i < gA_i$. In fact, if $y_i > 0$, the policyholder loses the roll-up incentive without taking full advantage of the guarantee ($y_i < gA_i$): if she does not need the entire amount gA_i and prefers to benefit from the investment of what exceeds her needs, it would be better to invest this money in a risk-free asset because the policy account, on average, yields less than the risk-free rate r due to the charge of the insurance fees. On the other hand, if $W_i > gA_i$, it is possible to withdraw an amount $y_i > gA_i$, but it is subject to the withdrawal penalty $k_i(y_i - gA_i)$. If, moreover, $gA_i < y_i < W_i$, there is a second penalization due to the reduction of the benefit base according to the pro-rata rule: this suggests that, rather than suffering the second penalization as well, it would be more convenient to withdraw the entire account value W_i, i.e., to surrender the contract. The following result states that this intuition is right.

Proposition 1. *The optimal solution of Problem* (15) *is* $y_i = 0$*, or* $y_i = gA_i$*, or* $y_i = W_i$*.*

Proof. We prove this result by backward induction, starting from the case $i = n$.

Initial step. Taking into account that $\mathbf{1}_{\{\tau\le n+1\}} = 1$ almost surely, and replacing Equations (3) and (4) in Problem (15), we get

$$v_n(W_n, A_n, \mu_n) = \sup_{0\le y_n\le \max\{gA_n, W_n\}} F_n(y_n; W_n, A_n, \mu_n),$$

[8] Recall that our unit of measurement of time is the common distance between two consecutive withdrawal dates, so that also x and ω are expressed according to this measure.

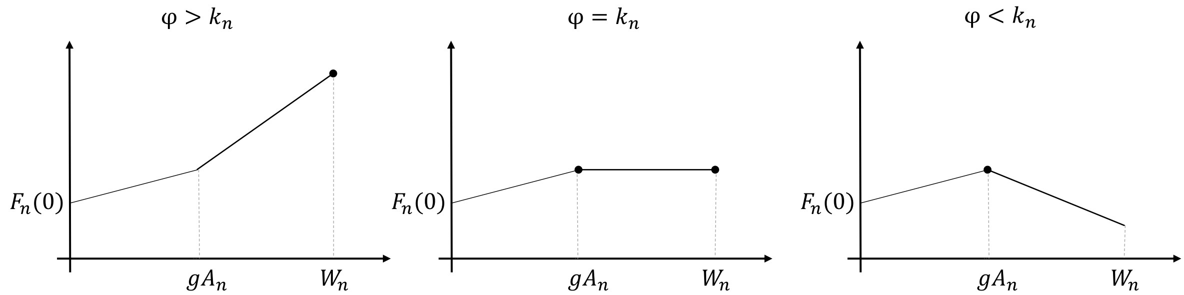

Fig. 1. The maximizer $y_n^* \in \{gA_n, W_n\}$.

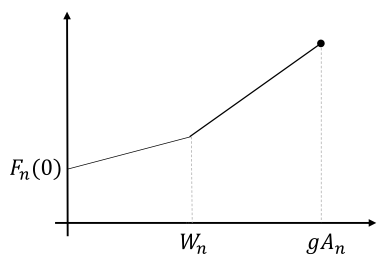

Fig. 2. The maximizer $y_n^* = gA_n$.

where

$$F_n(y_n; W_n, A_n, \mu_n) = y_n - k_n \max\{y_n - gA_n, 0\} + \max\{W_n - y_n, 0\}(1-\varphi).$$

- Then, if $gA_n < W_n$,

$$F_n(y_n; W_n, A_n, \mu_n) = \begin{cases} \varphi y_n + (1-\varphi)W_n & \text{if } 0 \le y_n \le gA_n \\ (\varphi - k_n)y_n + k_n gA_n + (1-\varphi)W_n & \text{if } gA_n < y_n \le W_n \end{cases},$$

so that the maximizer of F_n is (see Fig. 1)

1. $y_n^* = W_n$ if $\varphi > k_n$,
2. any element of the interval $[gA_n, W_n]$ (hence both $y_n^* = W_n$ and $y_n^* = gA_n$ are OK) if $\varphi = k_n$,
3. $y_n^* = gA_n$ if $\varphi < k_n$.

- If $gA_n \ge W_n$, then

$$F_n(y_n; W_n, A_n, \mu_n) = \begin{cases} \varphi y_n + (1-\varphi)W_n & \text{if } 0 \le y_n \le W_n \\ y_n & \text{if } W_n < y_n \le gA_n \end{cases},$$

so that the maximizer of F_n is $y_n^* = gA_n$ (see Fig. 2).

Summing up, replacing the maximizers in F_n, we have the following three alternatives:

$$v_n(W_n, A_n, \mu_n) = \begin{cases} (1-k_n)W_n + k_n gA_n & \text{if } gA_n < W_n \text{ and } \varphi > k_n \\ (1-\varphi)W_n + \varphi gA_n & \text{if } gA_n < W_n \text{ and } \varphi \le k_n \\ gA_n & \text{if } gA_n \ge W_n \end{cases}.$$

Then v_n is linear in both W_n and gA_n, and independent of μ_n. More generally, we can express it as

$$v_n(W_n, A_n, \mu_n) = C_n(\mu_n)W_n + D_n(\mu_n)gA_n,$$

where the two (constant) functions C_n and D_n are such that $0 \le C_n(\mu_n) \le 1$ and $D_n(\mu_n) \ge 0$ since the penalty and fee rates k_n and φ are between 0 and 1.

Iterative step. Assume now that $v_{i+1}(W_{i+1}, A_{i+1}, \mu_{i+1}) = C_{i+1}(\mu_{i+1})W_{i+1} + D_{i+1}(\mu_{i+1})gA_{i+1}$, where the two functions C_{i+1} and D_{i+1} are such that $0 \le C_{i+1}(\mu_{i+1}) \le 1$ and $D_{i+1}(\mu_{i+1}) \ge 0$ (almost surely), $i = n-1, n-2, \ldots, 1$. Then we have

$$v_i(W_i, A_i, \mu_i) = \sup_{0 \le y_i \le \max\{gA_i, W_i\}} F_i(y_i; W_i, A_i, \mu_i),$$

where

$$\begin{aligned} &F_i(y_i; W_i, A_i, \mu_i) \\ &= y_i - k_i \max\{y_i - gA_i, 0\} + q_i(\mu_i)\max\{W_i - y_i, 0\}(1-\varphi) \\ &+ \mathbb{E}^Q\left[e^{-\int_i^{i+1}\mu_u\,du} C_{i+1}(\mu_{i+1})e^{-r}\max\{W_i - y_i, 0\}(1+R_{i+1})(1-\varphi)|W_i, A_i, \mu_i\right] \\ &+ \mathbb{E}^Q\left[e^{-\int_i^{i+1}\mu_u\,du} D_{i+1}(\mu_{i+1})e^{-r} g f_{i+1}^A(W_i, A_i, y_i)|W_i, A_i, \mu_i\right]. \end{aligned} \tag{19}$$

Exploiting independence between financial and demographic factors, as well as the martingale property, we get

$$\begin{aligned} F_i(y_i; W_i, A_i, \mu_i) = y_i - k_i \max\{y_i - gA_i, 0\} + q_i(\mu_i)\max\{W_i - y_i, 0\}(1-\varphi) \\ + G_i(\mu_i)\max\{W_i - y_i, 0\}(1-\varphi) \\ + H_i(\mu_i)e^{-r} g f_{i+1}^A(W_i, A_i, y_i), \end{aligned}$$

where

$$G_i(\mu_i) = \mathbb{E}^Q\left[e^{-\int_i^{i+1}\mu_u\,du} C_{i+1}(\mu_{i+1})|\mu_i\right],$$
$$H_i(\mu_i) = \mathbb{E}^Q\left[e^{-\int_i^{i+1}\mu_u\,du} D_{i+1}(\mu_{i+1})|\mu_i\right].$$

Note that $0 \le C_{i+1}(\mu_{i+1}) \le 1$ and $D_{i+1}(\mu_{i+1}) \ge 0$ imply $0 \le G_i(\mu_i) \le p_i(\mu_i)$ and $H_i(\mu_i) \ge 0$.

- Then, if $gA_i < W_i$,

$$\begin{aligned} &F_i(y_i; W_i, A_i, \mu_i) \\ &= \begin{cases} \left[q_i(\mu_i) + G_i(\mu_i)\right] W_i(1-\varphi) + H_i(\mu_i)e^{-r} gA_i(1+b_i) \\ \qquad\qquad \text{if } y_i = 0 \\ \left\{1 - \left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi)\right\} y_i + \left[q_i(\mu_i) + G_i(\mu_i)\right] W_i(1-\varphi) \\ + H_i(\mu_i)e^{-r} gA_i \qquad\qquad \text{if } 0 < y_i \le gA_i \\ \{1 - k_i - \left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi) - H_i(\mu_i)e^{-r}\frac{1}{W_i - gA_i} gA_i\} y_i \\ + k_i gA_i + \left[q_i(\mu_i) + G_i(\mu_i)\right] W_i(1-\varphi) + H_i(\mu_i)e^{-r}\frac{W_i}{W_i - gA_i} gA_i \\ \qquad\qquad \text{if } gA_i < y_i \le W_i \end{cases}. \end{aligned} \tag{20}$$

Note that, being $G_i(\mu_i) \le p_i(\mu_i)$, in the second line of (20) the slope belongs to $(0,1)$ (hence F_i is increasing with y_i) while the intercept is lower than $F_i(0; W_i, A_i, \mu_i)$ as $b_i > 0$. It is not clear, instead, which is the sign of the slope in the third line. Hence, we can conclude that the maximizers of (20) belong to $\{0, gA_i, W_i\}$ (see Fig. 3).

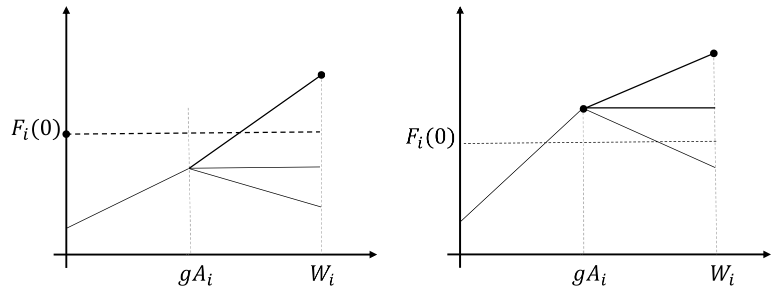

Fig. 3. The maximizer $y_i^* \in \{0, gA_i, W_i\}$.

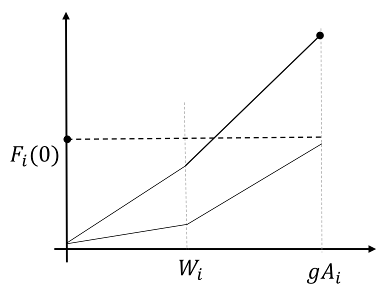

Fig. 4. The maximizer $y_i^* \in \{0, gA_i\}$.

- If instead $gA_i \geq W_i$, then

$$
\begin{aligned}
&F_i(y_i; W_i, A_i, \mu_i) \\
&= \begin{cases}
\left[q_i(\mu_i) + G_i(\mu_i)\right] W_i(1-\varphi) + H_i(\mu_i)\mathrm{e}^{-r} gA_i(1+b_i) & \text{if } y_i = 0 \\
\left\{1 - \left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi)\right\} y_i + \left[q_i(\mu_i) + G_i(\mu_i)\right] W_i(1-\varphi) & \\
+H_i(\mu_i)\mathrm{e}^{-r} gA_i & \text{if } 0 < y_i \leq W_i \\
y_i + H_i(\mu_i)\mathrm{e}^{-r} gA_i & \text{if } W_i < y_i \leq gA_i
\end{cases}
\end{aligned}
\tag{21}
$$

The first two lines in (21) are the same as in (20), while the slope in the third line ($=1$) is greater than in the second one. Hence now we can conclude that the maximizers of (21) belong to $\{0, gA_i\}$ (see Fig. 4).

Summing up, replacing the maximizers in F_i, we have the following four alternatives:

$$
\begin{aligned}
&v_i(W_i, A_i, \mu_i) \\
&= \begin{cases}
\left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi)W_i + H_i(\mu_i)\mathrm{e}^{-r}(1+b_i)gA_i & \text{if } y_i^* = 0 \\
\left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi)W_i & \\
+\left\{1 - \left[q_i(\mu_i) + G_i(\mu_i)\right](1-\varphi) + H_i(\mu_i)\mathrm{e}^{-r}\right\} gA_i & \text{if } y_i^* = gA_i < W_i \\
(1-k_i)W_i + k_i gA_i & \text{if } y_i^* = W_i > gA_i \\
\left[1 + H_i(\mu_i)\mathrm{e}^{-r}\right] gA_i & \text{if } y_i^* = gA_i \geq W_i
\end{cases}
\end{aligned}
$$

Therefore $v_i(W_i, A_i, \mu_i) = C_i(\mu_i)W_i + D_i(\mu_i)gA_i$, with $0 \leq C_i(\mu_i) \leq 1$ and $D_i(\mu_i) \geq 0$ (almost surely), and this concludes the proof. □

4.2. Stochastic interest rates

As already mentioned in Section 3, we have decided to focus on the case of a deterministic interest rate, allowing instead for stochasticity in the mortality intensity so as to capture the longevity risk. Nevertheless, our theoretical results would be preserved also under the assumption of a Markovian interest rate process $(r_t)_{t\geq 0}$ (still independent of mortality factors). In this section we briefly describe the changes needed in order to deal with the extension of these results to the case of stochastic interest rates.

First of all, in Equations (12) and (13), e^{-rt} must be replaced by $\mathrm{e}^{-\int_0^t r_u \mathrm{d}u}$, for $t = i, n+1$. Secondly, the contract value V_i and the value function v_i must depend also on the current interest rate level $r_i, i = 0, \dots, n+1$, since the expectations in (15)-(17) are conditioned on r_i as well. Moreover, in these expectations e^{-r} must be replaced by $\mathrm{e}^{-\int_i^{i+1} r_u \mathrm{d}u}$, and by $\mathrm{e}^{-\int_0^1 r_u \mathrm{d}u}$ in (18).

Concerning, in particular, the proof of Proposition 1, of course also the function $F_i, i = n, n-1, \dots, 1$, must depend on r_i. Then, in the initial step n, the value function v_n keeps its linearity with respect to W_n and gA_n and independence of μ_n and r_n. However, in the iterative step, we need to assume, more generally, that $v_{i+1}(W_{i+1}, A_{i+1}, \mu_{i+1}, r_{i+1}) = C_{i+1}(\mu_{i+1})W_{i+1} + D_{i+1}(\mu_{i+1}, r_{i+1})gA_{i+1}$, with $0 \leq C_{i+1}(\mu_{i+1}) \leq 1$ and $D_{i+1}(\mu_{i+1}, r_{i+1}) \geq 0$ (almost surely) for $i = n-1, n-2, \dots, 1$. Note that only D_{i+1} depends on r_{i+1}, while C_{i+1} does not. Afterwards the proof proceeds with $G_i(\mu_i)$ defined exactly as before, while instead $H_i(\mu_i)\mathrm{e}^{-r}$ must be replaced everywhere by $H_i(\mu_i, r_i) = \mathbb{E}^Q\left[\mathrm{e}^{-\int_i^{i+1}(\mu_u + r_u)\mathrm{d}u} D_{i+1}(\mu_{i+1}, r_{i+1}) | \mu_i, r_i\right]$.

4.3. Algorithm

Now we outline the algorithm employed to value the GLWB variable annuity, coming back to the case of constant interest rates. The execution of the algorithm requires a discretization over the state variables W, A, μ, and interpolation of the value function v over the resulting grid in order to compute the expectation in (17) and (18).

Step 0 For each $i = 1, 2, \dots, n+1$ discretize the state space $[0, \infty)$ for W_i, $(0, \infty)$ for μ_i, and $(0, A^{\max}]$ for A_i, where $A^{\max} = P\prod_{j=1}^{n-1}(1+b_j)$[9]:

$$
\begin{aligned}
\mathbb{W} &= \{w_1, \dots, w_L\}, \quad 0 = w_1 < w_2 < \cdots < w_L, \\
\mathbb{A} &= \{a_1, \dots, a_H\}, \quad 0 < a_1 < a_2 < \cdots < a_H = A^{\max}, \\
\mathbb{M} &= \{m_1, \dots, m_K\}, \quad 0 < m_1 < m_2 < \cdots < m_K.
\end{aligned}
$$

Step 1. Start at $n+1$ by setting $v_{n+1}(w_\ell, a_h, m_k) = 0$ for each $(w_\ell, a_h, m_k) \in \mathbb{W} \times \mathbb{A} \times \mathbb{M}$.

Step 2. Proceed backwards: for $i = n, n-1, \dots, 1$:

(I) interpolate the $L \cdot H \cdot K$ quadruples $\left(w_\ell, a_h, m_k, v_{i+1}(w_\ell, a_h, m_k)\right)$, $\ell = 1, 2, \dots, L$, $h = 1, 2, \dots, H$ and $k = 1, 2, \dots, K$, to construct the function $\tilde{v}_{i+1}(w, a, m)$ for $w \geq 0$, $0 < a \leq A^{\max}$, $m > 0$;

[9] It would be sufficient do discretize A_i until $P\prod_{j=1}^{i-1}(1+b_j)$, but in this case the grid would depend on i.

(II) for each $w_\ell, a_h, m_k \in \mathbb{W}\times\mathbb{A}\times\mathbb{M}$ compute

$$
\begin{aligned}
&v_i(w_\ell, a_h, m_k)\\
&= \sup_{y\in\{0, g a_h, w_\ell\}} \Bigg\{ y - k_i \max\{y - g a_h, 0\} + q_i(m_k)\max\{w_\ell - y, 0\}(1-\varphi)\\
&\qquad + e^{-r} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \mathrm{e}^{-\xi}\, \tilde{v}_{i+1}(\tilde{w}\mathrm{e}^{\delta}, \tilde{a}, \eta)\, f_1(\delta)\, g_1^{(m_k)}(\xi,\eta)\,\mathrm{d}\delta\,\mathrm{d}\xi\,\mathrm{d}\eta \Bigg\},
\end{aligned}
$$

where $\tilde{w} = \mathrm{e}^{r+d}\max\{w_\ell - y, 0\}(1-\varphi)$ with d the drift adjustment (10), $\tilde{a} = f_{i+1}^A(w_\ell, a_h, y)$, $g_1^{(m_k)}$ is the joint density of $\int_i^{i+1}\mu_u \mathrm{d}u$ and μ_{i+1} conditional on $\mu_i = m_k$,[10] and f_1 is the density of the Lévy process given by Equation (11).[11]

Step 3. The value of the contract at inception is

$$
\begin{aligned}
&v_0(P, P, \mu_0)\\
&= q_0(\mu_0)P(1-\varphi)\\
&+ e^{-r} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \mathrm{e}^{-\xi}\, \tilde{v}_1\left(P\mathrm{e}^{r+d+\delta}(1-\varphi), P, \eta\right)\, f_1(\delta)\, g_1^{(\mu_0)}(\xi,\eta)\,\mathrm{d}\delta\,\mathrm{d}\xi\,\mathrm{d}\eta.
\end{aligned}
$$

The density f_1 can be calculated through inversion of the characteristic function (9) (see Bailey and Swarztrauber, 1994) and the constant d via numerical integration. Similarly, numerical integration can be used to compute the integrals in Steps 2(II) and 3 of the algorithm.

4.4. Contract decomposition

It is clear that the valuation algorithm just described, aimed at producing the contract value under the dynamic approach, can be used to obtain, as simplified cases, also the contract values under alternative policyholder behaviours, namely under the static and the mixed approaches. To obtain the value under the static approach, in fact, it is sufficient to fix $y = g a_h$ in Step 2(II) without searching the maximum,[12] while to obtain the value under the mixed approach the search of the maximum must be restricted to the subset $\{g a_h, w_\ell\}$. To distinguish between these three different values we denote them, respectively, by $V_0^{dynamic}$, V_0^{static} and V_0^{mixed}.

Then we can see the dynamic contract as the combination of three components: the *basic GLWB contract*, i.e., the static one, the *surrender option* (with value given by $V_0^{surrender} := V_0^{mixed} - V_0^{static}$), and the *roll-up option* (whose value is $V_0^{rollup} := V_0^{dynamic} - V_0^{mixed}$):

$$
V_0^{dynamic} = V_0^{static} + V_0^{surrender} + V_0^{rollup}.
$$

5. Numerical examples

In this section we perform a sensitivity analysis comparing the numerical results obtained for different parameters and policyholder behaviours.

We focus on a contract with a single premium $P = 100$ and annual withdrawals, so that we use the standard unit of measurement of time, i.e., the year. We consider a policyholder aged $x = 65$ years at inception, which is in line with the retirement age in many countries.

[10] Assuming that the mortality process admits such a density.

[11] One could alternatively skip Step 2(I) and instead apply a numerical interpolation to $\tilde{v}_{i+1}(\cdot)$ over the time-$(i+1)$ grid, for each projected set of values (δ, ξ, η) in the numerical integration.

[12] This value could also be obtained directly, without dynamic programming, as the value of a portfolio composed by a standard whole life annuity with constant instalment equal to gP, and a whole life insurance with death benefit given by the residual policy account value.

5.1. Mortality process

We model the mortality intensity as an affine diffusion, following the non mean-reverting square root process

$$
\mathrm{d}\mu_t = (\alpha + \theta\mu_t)\mathrm{d}t + \sigma_\mu\sqrt{\mu_t}\,\mathrm{d}B_t, \tag{22}
$$

where $\alpha > 0$, $\theta > 0$, B is a Brownian motion and $\mu_0 > 0$ is given. Although this model is very simple (it collapses to the well-known deterministic Gompertz force of mortality when $\alpha = \sigma_\mu = 0$), it has the desirable property of producing strictly positive paths with probability 1, provided that $2\alpha > \sigma_\mu^2$. Moreover, it allows to get a closed-form formula for the survival probability given in (7) (see, e.g., Fung et al., 2014 and Dacorogna and Apicella, 2016). Specifically, we have[13]

$$
p_i(\mu_i) = c_1(1)\mathrm{e}^{-c_2(1)\mu_i}, \quad i = 0, 1, \ldots, n-1, \tag{23}
$$

where, for every $\zeta > 0$,

$$
c_1(\zeta) = \left(\frac{2\gamma \mathrm{e}^{1/2(\gamma-\theta)\zeta}}{(\gamma-\theta)(\mathrm{e}^{\gamma\zeta}-1)+2\gamma}\right)^{2\alpha/\sigma_\mu^2}, \quad c_2(\zeta) = \frac{2(\mathrm{e}^{\gamma\zeta}-1)}{(\gamma-\theta)(\mathrm{e}^{\gamma\zeta}-1)+2\gamma},
$$
$$
\gamma = \sqrt{\theta^2 + 2\sigma_\mu^2}.
$$

The conditional law of μ_s given μ_t (with $t < s$) features a noncentral chi-squared distribution. In particular, letting $Q(\mu_s \le z|\mu_t)$ the conditional cumulative distribution function of μ_s, the corresponding density function is

$$
f_{\mu_s|\mu_t}(z) = K f_{\chi^2_{\nu,\lambda}}(Kz), \tag{24}
$$

where

$$
K = \frac{4\theta}{\sigma_\mu^2(\mathrm{e}^{\theta(s-t)}-1)}, \quad \nu = \frac{4\alpha}{\sigma_\mu^2}, \quad \lambda = K\mu_t\,\mathrm{e}^{\theta(s-t)},
$$

and $f_{\chi^2_{\nu,\lambda}}$ denotes the probability density function of a noncentral chi-squared distribution with ν degrees of freedom and non-centrality parameter λ. More in detail:

$$
f_{\chi^2_{\nu,\lambda}}(h) = \sum_{j=0}^\infty \frac{\mathrm{e}^{-\frac{\lambda+h}{2}}\lambda^j h^{j+\frac{\nu}{2}-1}}{2^{2j+\frac{\nu}{2}}\Gamma(j+\frac{\nu}{2})},
$$

where $\Gamma(z) = \int_0^\infty y^{z-1}\mathrm{e}^{-y}\mathrm{d}y$ is the gamma function (see also Brigo and Mercurio, 2006 for further details).

As we have seen in Section 4.3, in order to get the contract value we need to compute the joint density function of μ_s and $\int_t^s \mu_u \mathrm{d}u$, given μ_t (with $t < s$).[14] Such a density can be characterized by its Laplace transform (see Lamberton and Lapeyre, 1995):

$$
\mathcal{L}_{s-t}^{(\mu_t)}(\beta_1,\beta_2) := \mathbb{E}^Q\left[\mathrm{e}^{-\beta_1\mu_s}\,\mathrm{e}^{-\beta_2\int_t^s\mu_u\mathrm{d}u}|\mu_t\right] = \mathrm{e}^{-\alpha\Psi_1}\,\mathrm{e}^{-\mu_t\Psi_2}, \tag{25}
$$

for any $\beta_1, \beta_2 \ge 0$, where

$$
\begin{aligned}
\Psi_1 &= -\frac{2}{\sigma_\mu^2}\ln\left(\frac{2\tilde{\gamma}\,\mathrm{e}^{(s-t)(\tilde{\gamma}-\theta)/2}}{\sigma_\mu^2\beta_1(\mathrm{e}^{\tilde{\gamma}(s-t)}-1)+\tilde{\gamma}+\theta+\mathrm{e}^{\tilde{\gamma}(s-t)}(\tilde{\gamma}-\theta)}\right),\\
\Psi_2 &= \frac{\beta_1(\tilde{\gamma}-\theta+\mathrm{e}^{\tilde{\gamma}(s-t)}(\tilde{\gamma}+\theta))+2\beta_2(\mathrm{e}^{\tilde{\gamma}(s-t)}-1)}{\sigma_\mu^2\beta_1(\mathrm{e}^{\tilde{\gamma}(s-t)}-1)+\tilde{\gamma}+\theta+\mathrm{e}^{\tilde{\gamma}(s-t)}(\tilde{\gamma}-\theta)},\\
\tilde{\gamma} &= \sqrt{\theta^2+2\sigma_\mu^2\beta_2}.
\end{aligned}
$$

Therefore we can obtain the conditional joint density of μ_s and $\int_t^s \mu_u \mathrm{d}u$, denoted by $g_{s-t}^{(\mu_t)}$, through (numerical) inversion of the Laplace transform. Indeed, if $g_{s-t}^{(\mu_t)}$ possesses first order partial derivatives $\partial g_{s-t}^{(\mu_t)}/\partial x_1$

[13] For $i \ge n$ we set $p_i(\mu_i) = 0$, see Section 4.

[14] For the (marginal) density function of $\int_t^s \mu_u \mathrm{d}u$ we refer the reader to Dufresne (2001).

and $\partial g_{s-t}^{(\mu_t)}/\partial x_2$ and second order derivative $\partial^2 g_{s-t}^{(\mu_t)}/\partial x_1 \partial x_2$, and there exist positive constants M, γ_1, γ_2 such that, for all strictly positive numbers x_1, x_2, it is

$$|g_{s-t}^{(\mu_t)}(x_1, x_2)| < M e^{\gamma_1 x_1 + \gamma_2 x_2}, \quad \left| \frac{\partial^2 g_{s-t}^{(\mu_t)}}{\partial x_1 \partial x_2} \right| < M e^{\gamma_1 x_1 + \gamma_2 x_2},$$

then

$$g_{s-t}^{(\mu_t)}(x_1, x_2) = \frac{1}{(2\pi i)^2} \int_{d_1 - i\infty}^{d_1 + i\infty} \int_{d_2 - i\infty}^{d_2 + i\infty} e^{\beta_1 x_1 + \beta_2 x_2} \mathcal{L}_{s-t}^{(\mu_t)}(\beta_1, \beta_2) d\beta_1 \, d\beta_2, \tag{26}$$

where $d_1 > \gamma_1$, $d_2 > \gamma_2$ and i is the unit imaginary number (see Cohen, 2007).

The parameters of the mortality intensity process used in our numerical experiments are the following: $\mu_0 = 0.00995483$, $\alpha = 0.0001$, $\theta = 0.1006875$, $\sigma_\mu = 0.01$. We have obtained them by fitting the survival probabilities for a 65 years old male implied by the projected life table IPS55. This table is commonly used in the Italian market and is focused on an individual born in 1955, hence with an age close to 65 at present (2022). The expected residual lifetime produced by the calibrated process is equal to 20.04, and we take $\omega = 118$, that is the extreme age in this life table, so that $n = 53$.

5.2. Financial model

As already mentioned in Section 3, for the reference price we assume an exponential Lévy process. In particular, we choose the Carr, Geman, Madan, Yor (CGMY) model whose characteristic function is

$$\Phi_t(u) = \exp\left(C\,\Gamma(-\mathcal{Y}) t \left[(M - iu)^{\mathcal{Y}} - M^{\mathcal{Y}} + (G + iu)^{\mathcal{Y}} - G^{\mathcal{Y}} \right] \right), \tag{27}$$

with G, $M \geq 0$, $C > 0$, $\mathcal{Y} < 2$ and Γ is the gamma function.

Such a model, introduced in Carr et al. (2002), allows for both a diffusion and a jump component. Moreover, it can be parametrized in order to capture finite or infinite variation. The parameters of the process are the same used in Kirkby and Nguyen (2021), i.e., $C = 0.02$, $G = 5$, $M = 15$, and $\mathcal{Y} = 1.2$.

For the interest rate r, unless otherwise mentioned, we fix a base level of 3%, which is in line with the EUR EIOPA term structure of risk-free interest rates for very long maturities.[15]

5.3. Contract parameters

The contract parameters in the basic case are fixed as follows: penalty rate $k_i \equiv k = 2\%$, withdrawal rate $g = 5\%$, roll-up rate $b_i \equiv b = 6\%$, fee rate $\varphi = 0.5\%$.

The choice of the withdrawal rate, that is consistent with that fixed in many other papers (e.g. Huang et al., 2014, Fung et al., 2014 and Huang and Kwok, 2016), is motivated by the fact that, with a (constant) interest rate of 3% and the survival probabilities implied by our model, the (actuarial) value of a whole-life annuity is 14.80, so that its reciprocal, that would be the conversion rate in a standard lifelong annuity, is 6.76%. However, if one buys such annuity, while being protected against the longevity risk (as in the case of the GLWB rider), completely looses the property of the investment fund. Then, if she dies during the first years of contract, her heirs do not receive anything while, in our case, the whole account value remains to them. In addition, within the annuity surrender is never admitted in order to avoid adverse selection. That is the reason why the withdrawal rate must be lower than the conversion rate of the annuity.

As far as the roll-up rate is concerned, we have first verified, by performing some preliminary numerical calculations of the initial contract value, that lower roll-up levels drive to negligible values of the roll-up option, i.e., to values of the contract under the mixed and the dynamic approaches very close to each other. Then, to disincentive withdrawals, the roll-up rate should be significantly high. On the other hand, this is quite expected. In fact, even neglecting interest rates, with e.g. a 6% roll-up rate it would take 16-17 years to recover a 0 withdrawal through an increase in the subsequent ones produced by the roll-up feature, and moreover this recovery can take place only if the policyholder lives long enough. The same level of roll-up has been considered, e.g., in Huang and Kwok (2016). Anyway, if the reader is interested to comparative static results corresponding to lower levels, she can simply refer to those reported for the mixed approach.

Concerning the level of the fee, our calculations show that, with the basic parameters here considered, the fair fee rate should be equal to 0.41% under the static approach, 0.49% under the mixed approach and 0.58% under the dynamic approach. We have chosen a basic level for φ somehow between these values because several empirical studies show that policyholders are not so active in the optimal exercise of their options.[16] Hence, recalling that the contract values that we have computed under the dynamic and the mixed approaches cover the worst-case scenario for the insurer, a more realistic prediction of the policyholder behaviour would lead to a fee rate not too far from that required under the static approach. In particular, we observe that this fee level is exactly the same applied on the account value to recover the cost of the GLWB offered within the MetLife Series VA,[17] and is also comparable with the fair fee rate obtained by Fung et al. (2014) under the static approach.

5.4. Results

In order to isolate the effect on the initial contract value of each (relevant) parameter of our model, in what follows we make a comparative static analysis by moving one parameter at a time and keeping all the others fixed at the level described in Sections 5.1-5.3. In particular, also the fee rate φ is fixed at the level of 0.5%. Then the contract is not (necessarily) fairly priced: more in detail, it is overpriced when its initial value is lower than the single premium $P = 100$, while it is underpriced when this value is $> P$. In the first case, in fact, the fair fee rate, say φ^*, is lower than the (fixed) rate $\varphi = 0.5\%$, while in the second case it is higher. As a consequence, the decomposition of the dynamic contract into the basic GLWB contract, the surrender option and the roll-up option may be driven not only by the parameter under consideration but also by the mis-pricing of the GLWB guarantee. In particular, surrender incentives are highly affected by this possible mis-pricing. This fact must never be forgotten, even if when we comment the numerical results we make explicit reference only to the parameter that changes.

5.4.1. Sensitivity analysis with respect to the contractual parameters

Coming now to the results of the numerical experiments, in Table 1 we study the initial contract value under the static (V_0^{static}), mixed (V_0^{mixed}) and dynamic ($V_0^{dynamic}$) approaches for different levels of the roll-up rate b. We also report the value in percentage of the static contract ($V_0^{static}(\%)$), the surrender option ($V_0^{surrender}(\%)$) and the roll-up option ($V_0^{rollup}(\%)$) to highlight the three components of the GLWB dynamic contract. Obviously, the contract value under static and mixed scenarios is independent of the roll-up rate, while it increases under the dynamic approach. Moreover, as already anticipated, for levels of the roll-up rate not sufficiently high ($b \leq 5\%$), $V_0^{dynamic} = V_0^{mixed}$. In particular, the roll-up option component $V_0^{rollup}(\%)$ is increasing from 0 to almost 6% to the detriment of the main component, i.e., the static one, while the (decreasing) surrender option remains negligible and rather stable around 0.7%.

[15] See http://www.eiopa.europa.eu/tools-and-data/risk-free-interest-rate-term-structures_en.

[16] See, e.g., Bauer et al. (2017) and the references therein.

[17] See Koijen and Yogo (2022).

Table 1
Contract values and contract components (in %) for different roll-up rates b.

$b(\%)$	V_0^{static}	V_0^{mixed}	$V_0^{dynamic}$	$V_0^{static}(\%)$	$V_0^{surrender}(\%)$	$V_0^{rollup}(\%)$
5	99.25	99.98	99.98	99.27	0.73	0.00
6	99.25	99.98	100.41	98.84	0.73	0.43
7	99.25	99.98	101.49	97.79	0.72	1.49
8	99.25	99.98	102.84	96.51	0.71	2.78
9	99.25	99.98	104.41	95.05	0.70	4.25
10	99.25	99.98	106.23	93.42	0.69	5.88

Table 2
Contract values and contract components (in %) for different penalty rates k.

$k(\%)$	V_0^{static}	V_0^{mixed}	$V_0^{dynamic}$	$V_0^{static}(\%)$	$V_0^{surrender}(\%)$	$V_0^{rollup}(\%)$
0	99.25	100.92	101.37	97.91	1.65	0.44
1	99.25	100.41	100.85	98.41	1.15	0.44
2	99.25	99.98	100.41	98.84	0.73	0.43
3	99.25	99.64	100.07	99.18	0.39	0.43
4	99.25	99.40	99.84	99.41	0.15	0.44
5	99.25	99.28	99.73	99.52	0.03	0.45
6	99.25	99.25	99.71	99.54	0.00	0.46

Table 3
Contract values and contract components (in %) for different withdrawal rates g.

$g(\%)$	V_0^{static}	V_0^{mixed}	$V_0^{dynamic}$	$V_0^{static}(\%)$	$V_0^{surrender}(\%)$	$V_0^{rollup}(\%)$
3	93.26	97.59	97.59	95.56	4.44	0.00
4	95.34	97.78	97.83	97.46	2.49	0.05
5	99.25	99.98	100.41	98.84	0.73	0.43
6	105.54	105.71	106.29	99.29	0.16	0.54
7	114.01	114.05	114.46	99.61	0.03	0.36

In Table 2 the initial contract value in all scenarios is displayed for different levels of the penalty rate k. Clearly, V_0^{static} remains constant not depending on k, while V_0^{mixed} and $V_0^{dynamic}$ obviously decrease. In particular, the weight of the surrender option component decreases from 1.65%, when no surrender penalties are applied, to 0, when $k \geq 6\%$ and the difference between static and mixed approaches completely vanishes. The decrease of this component is (almost) completely absorbed by an increase in the static contract value, while the roll-up option component does not have a monotonic trend, although it remains very stable and almost negligible (around 0.45% of the dynamic contract value).

In Table 3 we move the withdrawal rate g up to 7% even if, being the conversion rate associated to a standard whole life annuity equal to 6.76%, g-levels higher than 6% would be completely impractical as they produce an extremely high contract value (even under the static approach). As expected, the contract value in all scenarios strongly increases with g. In particular, when g is very low (3%), the surrender option component is substantial (4.44%), while very high levels of g (e.g. 7%) make surrender (almost) completely unattractive. In contrast, the roll-up option is valueless when $g = 3\%$ because, after a first year of no withdrawal, the increase in the withdrawal base makes the guaranteed withdrawal equal to 3.18% of the single premium P (4.24% if $g = 4\%$), still too low in order to have a competitive contract where surrender is discouraged. Moreover, as in the case of the penalty rate, the roll-up component does not seem to have a monotonic trend and anyway remains always very low.

To highlight the influence of the fee rate on the initial contract value under the different policyholder behaviours, we plot, in Fig. 5, V_0^{static}, V_0^{mixed} and $V_0^{dynamic}$ against φ maintaining fixed all the remaining basic parameters. From this figure one can capture at first glance the decrease in V_0 with respect to φ as well as the differences among the contract values under the various approaches. In particular, when φ is very low

Table 4
Contract values and contract components (in %) for different interest rates r.

$r(\%)$	V_0^{static}	V_0^{mixed}	$V_0^{dynamic}$	$V_0^{static}(\%)$	$V_0^{surrender}(\%)$	$V_0^{rollup}(\%)$
1	109.47	109.54	112.96	96.91	0.06	3.03
2	103.39	103.65	105.24	98.24	0.25	1.51
3	99.25	99.98	100.41	98.84	0.73	0.43
4	96.65	98.24	98.25	98.37	1.62	0.01
5	95.13	97.75	97.75	97.32	2.68	0.00

(until 0.2%), the surrender option component vanishes since it is not convenient to surrender a contract offering a so cheap guarantee; hence $V_0^{static} = V_0^{mixed}$. On the other hand, when φ approaches 1%, the roll-up component tends to disappear, i.e., V_0^{mixed} and $V_0^{dynamic}$ are very close to each other; moreover, they are far enough from V_0^{static}, thus highlighting a very strong surrender incentive. Finally, note that the fair fee rate φ^* is given by the abscissa of the intersection between the contract graph and the horizontal line at level $P = 100$: $\varphi^* = 0.41\%$ for the static approach, $\varphi^* = 0.49\%$ for the mixed approach and $\varphi^* = 0.58\%$ for the dynamic approach.

5.4.2. *Sensitivity analysis with respect to interest rates and mortality*

A very important issue related to the long term nature of GLWB is implied by possible misspecified parameters. If this misspecification takes place at the contract inception, it can have a significant impact on pricing and profit/loss for the guarantee provider. For this reason, in what follows we conduct an investigation by changing some assumptions that we have so far kept fixed.

We start this analysis by changing the interest rate r, that is a key-parameter which has a very strong effect on the contract value. We show the corresponding results in Table 4. As expected, we observe a general decreasing of the contract value in all scenarios, as well as of the roll-up option weight. In particular, for $r \geq 5\%$, $V_0^{mixed} = V_0^{dynamic}$: this means that when r is sufficiently high the exercise of the roll-up option does not return any benefit. Recall, in fact, that with $r = 0\%$ and $b = 6\%$ it takes almost 17 years to recover one year of non-withdrawal through the increase of the guaranteed amount; when r increases, more time is needed because the future withdrawals are subject to (financial) discount,[18] and hence count less. The weight of the static component does not change so much, and does not display a monotonic trend, while the weight of the surrender option, almost negligible for low levels of the interest rate, increases with r reaching the level of 2.68% when $r = 5\%$.

Focusing on the dynamic approach, we move now the mortality assumptions, in particular the volatility σ_μ of the intensity process, while keeping the other parameters fixed. The results obtained are reported in Table 5, where, along with the initial contract value $V_0^{dynamic}$, we show the expected residual lifetime, denoted by e_{65}, implied by the new volatility level. As we can see, both the expected residual lifetime e_{65} and the contract value $V_0^{dynamic}$ increase with the volatility of the mortality intensity, but the effect of changes in σ_μ is not very strong. We did not consider higher levels of σ_μ because they do not satisfy the Feller condition $2\alpha > \sigma_\mu^2$.

We also investigate the effects of a misspecification of the life table on which the mortality intensity process has been calibrated. To this end we consider, on one hand, the most recent projected life table used in the Italian market, named A62, focused on an individual born in 1962, and fit the survival probabilities for a 65 years old male implied by this table. The expected residual lifetime resulting from the calibration is 21.93, i.e., almost 2 years more than that implied by our base table. On the other hand, we consider the life table HMD IT 1×10 (2000-2009), again with reference to an Italian male aged 65 years. In this case the

[18] In addition to the "demographic discount".

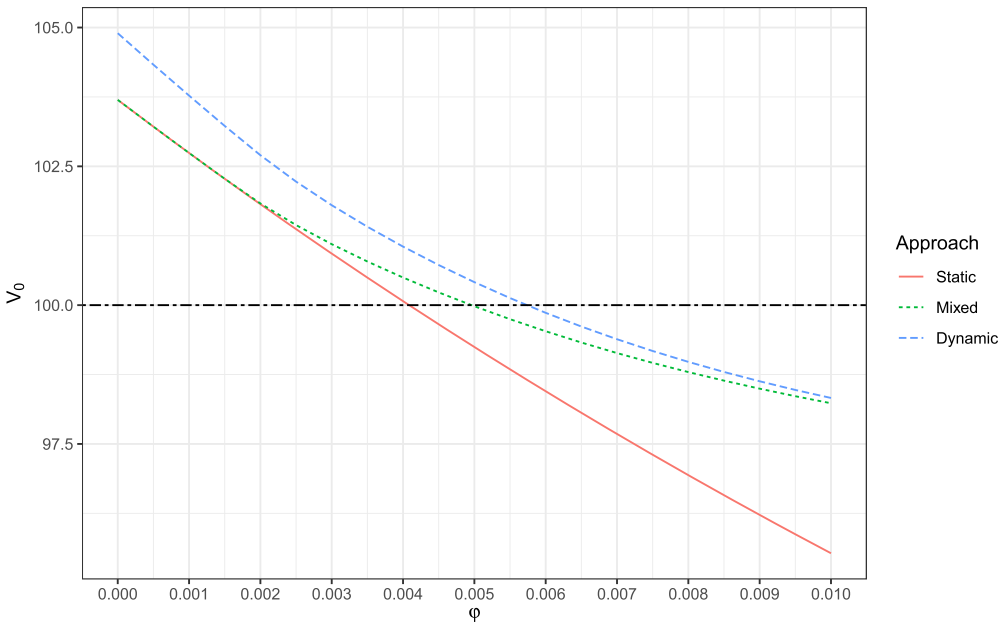

Fig. 5. Contract value against the fee rate φ under the different policyholder behaviours.

Table 5
$V_0^{dynamic}$ and expected residual lifetime for different levels of the volatility σ_μ.

$\sigma_\mu(\%)$	e_{65}	$V_0^{dynamic}$
0.4	19.78	99.99
0.6	19.84	100.09
0.8	19.92	100.23
1.0	20.04	100.41
1.2	20.18	100.68
1.4	20.37	101.05

Table 6
$V_0^{dynamic}$ and calibrated mortality parameters corresponding to different life tables.

table	μ_0	α	θ	σ_μ	e_{65}	$V_0^{dynamic}$
HMD	0.01212986	0.0001	0.1105087	0.01	17.57	98.73
IPS55	0.00995483	0.0001	0.1006875	0.01	20.04	100.41
A62	0.00618973	0.0001	0.1142278	0.01	21.93	101.68

expected residual lifetime $e_{65} = 17.57$, about 2 and a half years less than that of our base case. The calibrated parameters, as well as the results obtained for the initial contract value under the dynamic approach, are reported in Table 6 where, for comparison, we show also the results obtained in our base case. As we can see, with the calibration on A62 the contract value increases from 100.41 to 101.68 (about 1.26%), while with HMD it decreases from 100.41 to 98.73 (about 1.67%).

Summing up, while the interest rate r heavily impacts the value of the contract, the effect of mortality seems to be important but not so dramatic, at least with the non mean reverting square root process here adopted in order to illustrate some numerical examples.

5.4.3. Optimal withdrawal strategies for different contractual parameters

As already remarked in Section 3, a withdrawal strategy solving the optimization problem (12), i.e., an *argsup* of the problem, does not necessarily lead to the optimization of the policyholder's preferences, but leads to the worst-case scenario from the insurance company's point of view. Hence it can be seen as an intermediate result to achieve the final goal, that is pricing. Nonetheless, it is very important to understand what is the implied optimal withdrawal behaviour of the policyholder, in order to provide a useful guidance to practitioners and other academics. That is why we show this behaviour at different points in time and for different underlying contractual parameter specification.

More in detail, in the following tables we highlight the optimal withdrawal at time i, i.e. y_i, as a function of the account value W_i. To this end we have fixed the other two parameters on which y_i depends, namely the benefit base A_i and the mortality intensity μ_i, at the levels of $P = 100$ and $\mathbb{E}^Q[\mu_i] = e^{\theta i}\mu_0 - \frac{\alpha}{\theta}(1 - e^{\theta i})$, respectively. In all the tables we have separated the *surrender zone* ($y_i = W_i > gA_i$) from the *continuation zone* and, within the latter, the *no-withdrawal zone* ($y_i = 0$) from the *guaranteed-withdrawal zone* ($y_i = gA_i$), and specified the values of W_i identifying these zones.

The results concerning the basic set of contractual parameters are reported in Table 7, from which we notice that for high values of the personal account it is convenient to surrender the contract. This is quite expected: in fact, in these cases, the GLWB guarantee is deeply out of the money and it is very unlikely that this rider will be activated before death. Then, it is better to exit the contract immediately rather than continue to pay a fee for nothing. As time goes on, the surrender zone expands to the detriment of the continuation zone; e.g., for $i = 1$ it is given by the set of account values not less than 125, while for $i = 20$, corresponding approximately to the average remaining lifetime of a 65-year-old male, this set includes also the account values between 80 and 125. Of course, this is in line with the decrease in the value of the remaining guaranteed payments, as will become explicit immediately below.

As far as the continuation zone is concerned, we observe that, for relatively low levels of the account value it is always convenient to withdraw the guaranteed amount gA_i while, for moderate levels, it is convenient not to withdraw anything so taking profit of the roll-up feature, at least in the first years of contract. Moreover, for $i = 5$ and account values sufficiently high (between 100 and 115), there is a reversion to the guaranteed amount.

To explain these facts, we start by fixing the ideas on $i = 1$. The actuarial value of all future guaranteed payments in case of immediate withdrawal of the guaranteed amount is equal to 66.6, while it increases by 6% (roll-up rate) and becomes 70.6 in case of no withdrawal. Then, if the contract remains in force until death, the total value of the guaranteed payments in case of withdrawal also includes the current payment of 5 currency units, becoming 71.6. Conversely, in case of no withdrawal, it lies between 70.6 and 75.6 net of fees applied until death on

Table 7
Withdrawal zones in the basic case.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	$65 \le W_i < 125$	$W_i < 65$	$W_i \ge 125$
5	$80 \le W_i \le 100$	$W_i < 80$ or $100 < W_i < 115$	$W_i \ge 115$
10	–	$W_i < 100$	$W_i \ge 100$
20	–	$W_i < 80$	$W_i \ge 80$

Table 8
Withdrawal zones when the roll-up rate $b = 4\%$.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	–	$W_i < 120$	$W_i \ge 120$
5	–	$W_i < 115$	$W_i \ge 115$
10	–	$W_i < 100$	$W_i \ge 100$
20	–	$W_i < 80$	$W_i \ge 80$

Table 9
Withdrawal zones when the roll-up rate $b = 8\%$.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	$W_i < 135$	–	$W_i \ge 135$
5	$20 \le W_i < 120$	$W_i < 20$	$W_i \ge 120$
10	$45 \le W_i < 105$	$W_i < 45$	$W_i \ge 105$
20	–	$W_i < 80$	$W_i \ge 80$

Table 10
Withdrawal zones when the penalty rate $k = 0$.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	$65 \le W_i < 115$	$W_i < 65$	$W_i \ge 115$
5	$75 \le W_i < 105$	$W_i < 75$	$W_i \ge 105$
10	–	$W_i < 90$	$W_i \ge 90$
20	–	$W_i < 65$	$W_i \ge 65$

Table 11
Withdrawal zones when the penalty rate $k = 4\%$.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	$65 \le W_i \le 120$	$W_i < 65$ or $120 < W_i < 135$	$W_i \ge 135$
5	$80 \le W_i \le 105$	$W_i < 80$ or $105 < W_i < 125$	$W_i \ge 125$
10	–	$W_i < 120$	$W_i \ge 120$
20	–	for all W_i	–

Table 12
Withdrawal zones when the withdrawal rate $g = 4\%$.

i	Continuation zone $y_i = 0$	Continuation zone $y_i = gA_i$	Surrender zone $y_i = W_i$
1	$55 \le W_i < 100$	$W_i < 55$	$W_i \ge 100$
5	$65 \le W_i \le 80$	$W_i < 65$ or $80 < W_i < 90$	$W_i \ge 90$
10	–	$W_i < 80$	$W_i \ge 80$
20	–	$W_i < 65$	$W_i \ge 65$

the amount of 5 currency units at the annual rate of 0.5%. More in detail, if the GLWB guarantee will be triggered before death, then nothing will remain in the personal account to be paid as a death benefit, and the "value" of a null withdrawal is simply equal to 70.6, hence less than that of the guaranteed withdrawal (71.6). If instead the GLWB guarantee will never be activated, there will still be some money in the account at death, and the "value" of a null withdrawal may dominate that of gA_i. The fact that the probability of activation of the GLWB before death decreases with the current account value motivates the convenience to withdraw gA_i for low levels of W_i (actually at $i = 1$ the guarantee is already in the money for $W_i < 65$) and 0 for moderate levels. Of course the surrender threshold, 125, is much greater than 65 because even for account values close to 125 there is still the possibility that the GLWB guarantee becomes in the money and will be activated in the future; recall, in fact, that there is a positive probability that the policyholder remains alive up to 118 years!

Similar considerations can be made in the other cases. E.g., for $i = 5$ the total "value" of an immediate withdrawal is $57.7 + 5$ while that of a null withdrawal is between 61.1 and $61.1 + 5 - \text{fees}$. However, in this case, if the account value is sufficiently high (between 100 and 115), there is a concrete possibility that the guarantee will become deeply out of the money in the future and the contract will be surrendered. Then it results more convenient to withdraw gA_i immediately and not let it in the account subject to reductions due to the application of the fees and possibly also of the surrender penalty. In case of early termination, indeed, the surcharge value due to the roll-up provision will not be fully exploited and hence is less than $61.1 - 57.7$, so the decision to withdraw the guaranteed amount can overcome that of a null withdrawal.

To complete the analysis of the basic case, we notice that, as expected, the no-withdrawal zone shrinks until disappearing completely for $i = 10$ and $i = 20$. In these cases the surcharge due to the roll-up feature, if fully exploited until death, boils down to 2.8 and 1.6, respectively.

Tables 8 and 9 allow to analyse the effect on the withdrawal behaviour of the roll-up rate b. We observe that, when $b = 4\%$, the no-withdrawal zone is always empty, which confirms the fact that the contract value under the dynamic approach is the same as under the mixed approach, while the surrender zone remains (almost) the same as in the basic case. When instead $b = 8\%$, the surrender zone shrinks a bit from the basic case, while, as expected, the no-withdrawal zone expands, absorbing the guaranteed-withdrawal zone for $i = 1$ and remaining non-empty even for $i = 10$.

Similar results, but referred now to the penalty rate k, are reported in Tables 10 and 11. When $k = 0$ the surrender zone obviously widens. This widening goes all to the detriment of the no-withdrawal zone for $i = 1$ (compared to the basic case), while for $i = 5$ the no-withdrawal zone loses ground with respect to the guaranteed-withdrawal one. When instead $k = 4\%$ the surrender zone shrinks and disappears completely for $i = 20$: in this case, indeed, the optimal strategy is to withdraw the guaranteed amount, independently of the account value.

In Tables 12 and 13 we report the results concerning the withdrawal rate g. A lower g, with respect to the basic case, expands the surrender zone, while symmetrically a higher g shrinks it. E.g., for $i = 1$ the surrender zone, given by the interval $[125, +\infty)$ in the basic case, moves to $[100, +\infty)$ when $g = 4\%$ and to $[150, +\infty)$ when $g = 6\%$; for $i = 20$ these movements go from $[80, +\infty)$ to $[65, +\infty)$ and $[95, +\infty)$ respectively. Recall, however, that this effect on surrender is not only driven by the change in the contractual parameter but also by the fact that for low g's the contract is overpriced since we are maintaining fixed the fee rate at the level of 0.5%, while for high g's it is underpriced The increase/decrease of the surrender zone all goes to the disadvantage/advantage of the guaranteed-withdrawal zone when $i = 10$ and $i = 20$, and pours more or less uniformly on the two continuation zones when $i = 1$. When instead $i = 5$ the prevailing effect of the increase/reduction of surrenders is on the guaranteed-withdrawal zone if the contract is overpriced (for $g = 4\%$ this zone is reduced by 20 against a reduction of only 5 for the no-withdrawal zone compared to the basic case) and on the no-withdrawal zone if the contract is underpriced ($g = 6\%$).

To conclude this analysis we move now the fee rate φ. Recalling that in the basic case the contract is (slightly) underpriced since the fair fee rate would be equal to 0.58% against 0.5%, we do not consider

Table 13
Withdrawal zones when the withdrawal rate $g = 6\%$.

	Continuation zone		Surrender zone
i	$y_i = 0$	$y_i = gA_i$	$y_i = W_i$
1	$75 \le W_i < 150$	$W_i < 75$	$W_i \ge 150$
5	$90 \le W_i \le 125$	$W_i < 90$ or $125 < W_i < 135$	$W_i \ge 135$
10	–	$W_i < 120$	$W_i \ge 120$
20	–	$W_i < 95$	$W_i \ge 95$

Table 14
Withdrawal zones when the fee rate $\varphi = 0.7\%$.

	Continuation zone		Surrender zone
i	$y_i = 0$	$y_i = gA_i$	$y_i = W_i$
1	$70 \le W_i < 115$	$W_i < 70$	$W_i \ge 115$
5	$85 \le W_i \le 90$	$W_i < 85$ or $90 < W_i < 105$	$W_i \ge 105$
10	–	$W_i < 95$	$W_i \ge 95$
20	–	$W_i < 70$	$W_i \ge 70$

values for φ lower than the current one, but show only an example of overpricing, with $\varphi = 0.7\%$. The effect on the withdrawal zones is shown in Table 14, and it is quite similar to that highlighted in Table 12: we have an expansion, with respect to the basic case, of the surrender zone, which all goes to the detriment of the guaranteed-withdrawal zone when $i = 10$ and $i = 20$, while for low i's this zone slightly expands, to the detriment of the other two (hence not only of surrenders but also of no-withdrawals).

6. Conclusions

In this paper we have proposed a discrete time model, based on dynamic programming, to price GLWB variable annuities under the dynamic approach within a stochastic mortality framework. We have verified, by backward induction, the bang-bang condition for the set of discrete withdrawal strategies of the model, and offered an interesting contract decomposition. We have considered a quite general set-up, only requiring the Markovian property for the mortality intensity and the asset price processes. To maintain the curse of dimensionality of our valuation algorithm manageable, we have assumed constant interest rates, although our results still hold in the case of a Markovian interest rate stochastic process. Then, we have numerically implemented the model by focusing on a non mean-reverting square root process for the mortality intensity and an exponential Lévy process for the asset price. In particular, we have considered the Carr, Geman, Madan, Yor (CGMY) process and computed the initial value of the contract for different contractual parameters and policyholder behaviours. In such analysis we have also highlighted the components into which the contract can be split. Moreover, in order to verify the impact on the contract value of possible misspecified parameters, we have considered alternative calibrations for the mortality intensity, as well as different interest rates. Finally, we have shown how the optimal policyholder behaviour depends on the current account value at different points in time and for different contractual parameters.

Our numerical results, that highlight which are the main drivers of the contract value, can be particularly useful to GLWB providers when designing the product and fixing the contractual parameters. These results, completely in line with the existing literature, have proven to be very sensitive to the financial market situation. In the paper we have fixed the interest rate in accordance with the current situation, but only few months ago this choice would have been different. The asset price model, instead, has been chosen as an example in order to illustrate the effect of the various parameters on the contract value, and the consequent results must be interpreted in relative, rather than absolute, terms. The providers, indeed, should be particularly careful in the model choice, trying to capture as much as possible the specific features of the investment portfolio backing the VA, that is usually selected by the policyholder among a set of alternative opportunities, so as to match her preferences in terms of risk/return profile. Also the mortality intensity process has been chosen as an example, but in this case the results corresponding to different calibrations, implying modifications in the expected residual lifetime even up to 10 - 12%, do not change dramatically.

Summing up, we strongly believe that our approach can be profitably applied in practice.

Declaration of competing interest

There is no conflict of interest.

Data availability

No data was used for the research described in the article.

Acknowledgements

We are very grateful to two anonymous referees for their valuable comments and suggestions that have allowed us to improve the paper. All remaining errors and omissions are our own sole responsibility.

This work was partially supported by the GNAMPA-INdAM Group.

References

Alonso-García, J., Wood, O., Ziveyi, J., 2018. Pricing and hedging guaranteed minimum withdrawal benefits under a general Lévy framework using the COS method. Quantitative Finance 18 (6), 1049–1075.

Azimzadeh, P., Forsyth, P., 2015. The existence of optimal bang-bang controls for GMxB contracts. SIAM Journal on Financial Mathematics 6 (1), 117–139.

Bacinello, A.R., Millossovich, P., Montealegre, A., 2016. The valuation of GMWB variable annuities under alternative fund distributions and policyholder behaviours. Scandinavian Actuarial Journal 2016 (5), 446–465.

Bacinello, A.R., Millossovich, P., Olivieri, A., Pitacco, E., 2011. Variable annuities: a unifying valuation approach. Insurance: Mathematics and Economics 49 (3), 285–297.

Bacinello, A.R., Millossovich, P., Olivieri, A., Pitacco, E., 2012. Variable annuities as life insurance packages: a unifying approach to the valuation of guarantees. In: Vanmaele, M., Deelstra, G., De Schepper, A., Dhaene, J., Schoutens, W., Vanduffel, S., Vyncke, D. (Eds.), Actuarial and Financial Mathematics Conference. Interplay Between Finance and Insurance. Koninklijke Vlaamse Academie van België, pp. 3–15.

Bailey, D.H., Swarztrauber, P.N., 1994. A fast method for the numerical evaluation of continuous Fourier and Laplace transforms. SIAM Journal on Scientific Computing 15 (5), 1105–1110.

Bauer, D., Gao, J., Moenig, T., Ulm, E., Zhu, N., 2017. Policyholder exercise behavior in life insurance: the state of affairs. North American Actuarial Journal 21 (4), 485–501.

Bauer, D., Kling, A., Russ, J., 2008. A universal pricing framework for guaranteed minimum benefits in variable annuities. ASTIN Bulletin 38 (2), 621–651.

Bertsekas, D., 2005. Dynamic Programming and Optimal Control. Athena Scientific.

Brigo, D., Mercurio, F., 2006. Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit. Springer Finance.

Carr, P., Geman, H., Madan, D.B., Yor, M., 2002. The fine structure of asset returns: an empirical investigation. Journal of Business 75 (2), 305–332.

Chen, Z., Forsyth, P., 2008. A numerical scheme for the impulse control formulation for pricing variable annuities with a guaranteed minimum withdrawal benefit (GMWB). Numerische Mathematik 109 (4), 535–569.

Chopra, D., Erzan, O., de Gantès, G., Grepin, L., Slawner, C., 2009. Responding to the variable annuity crisis. McKinsey Working Papers on Risk 10.

Cohen, A.M., 2007. Numerical Methods for Laplace Transform Inversion. Springer.

Dacorogna, N., Apicella, G., 2016. A general framework for modeling mortality to better estimate its relationship to interest rate risks. SCOR Papers 39.

Dai, M., Kwok, K.Y., Zong, J., 2008. Guaranteed minimum withdrawal benefit in variable annuities. Mathematical Finance 18 (4), 595–611.

Dufresne, D., 2001. The integrate square-root process. Working Paper. University of Montreal.

Fontana, C., Rotondi, F., 2023. Valuation of general GMWB annuities in a low interest rate environment. Insurance: Mathematics and Economics 112, 142–167.

Forsyth, P., Vetzal, K., 2014. An optimal stochastic control framework for determining the cost of hedging of variable annuities. Journal of Economic Dynamics and Control 44, 29–53.

Fung, M., Ignatieva, K., Sherris, M., 2014. Systematic mortality risk: an analysis of guaranteed lifetime withdrawal benefits in variable annuities. Insurance: Mathematics and Economics 58, 103–115.

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

Holz, D., Kling, A., Russ, J., 2012. GMWB for life: an analysis of lifelong withdrawal guarantees. Zeitschrift für die Gesamte Versicherungswissenschaft 101 (3), 305–325.

Huang, Y.T., Kwok, Y.K., 2016. Regression-based Monte Carlo methods for stochastic control models: variable annuities with lifelong guarantees. Quantitative Finance 16 (6), 905–928.

Huang, H., Milevsky, M.A., Salisbury, T.S., 2014. Optimal initiation of a GLWB in a variable annuity: no Arbitrage approach. Insurance: Mathematics and Economics 56, 102–111.

Huang, Y.T., Zeng, P., Kwok, Y.K., 2017. Optimal initiation of guaranteed lifelong withdrawal benefit with dynamic withdrawals. SIAM Journal on Financial Mathematics 8 (1), 804–840.

Kirkby, J., Nguyen, D., 2021. Equity-linked guaranteed minimum death benefits with dollar cost averaging. Insurance: Mathematics and Economics 100, 408–428.

Kling, A., Ruez, F., Russ, J., 2014. The impact of policyholder behavior on pricing, hedging, and hedge efficiency of withdrawal benefit guarantees in variable annuities. European Actuarial Journal 4 (2), 281–314.

Koijen, R., Yogo, M., 2022. The fragility of market risk insurance. The Journal of Finance 77 (2), 815–862.

Lamberton, D., Lapeyre, B., 1995. An Introduction to Stochastic Calculus Applied to Finance. Chapman and Hall.

Ledlie, M., Corry, D., Finkelstein, G., Ritchie, A., Su, K., Wilson, D., 2008. Variable annuities. British Actuarial Journal 14 (2), 327–389.

Longstaff, F.A., Schwartz, E.S., 2001. Valuing American options by simulation: a simple least-squares approach. The Review of Financial Studies 14 (1), 113–147.

Luo, X., Shevchenko, P.V., 2015a. Fast numerical method for pricing of variable annuities with guaranteed minimum withdrawal benefit under optimal withdrawal strategy. International Journal of Financial Engineering 2 (3).

Luo, X., Shevchenko, P.V., 2015b. Valuation of variable annuities with guaranteed minimum withdrawal and death benefits via stochastic control optimization. Insurance: Mathematics and Economics 62, 5–15.

Milevsky, M.A., Salisbury, T.S., 2006. Financial valuation of guaranteed minimum withdrawal benefits. Insurance: Mathematics and Economics 38 (1), 21–38.

Moenig, T., 2021. Efficient valuation of variable annuity portfolios with dynamic programming. The Journal of Risk and Insurance 88 (4), 1023–1055.

Piscopo, G., Haberman, S., 2011. The valuation of guaranteed lifelong withdrawal benefit options in variable annuity contracts and the impact of mortality risk. North American Actuarial Journal 15 (1), 59–76.

Schoutens, W., 2003. Lévy Processes in Finance: Pricing Financial Derivatives. Wiley.

Seierstad, A., 2009. Stochastic Control in Discrete and Continuous Time. Springer.

Shevchenko, P.V., Luo, X., 2016. A unified pricing of variable annuity guarantees under the optimal stochastic control framework. Risks 4 (3), 22.

Shevchenko, P.V., Luo, X., 2017. Valuation of variable annuities with guaranteed minimum withdrawal benefit under stochastic interest rate. Insurance: Mathematics and Economics 76, 104–117.

Steinorth, P., Mitchell, O.S., 2015. Valuing variable annuities with guaranteed minimum lifetime withdrawal benefits. Insurance: Mathematics and Economics 64, 246–258.

Tankov, P., Cont, R., 2003. Financial Modelling with Jump Processes. Chapman and Hall/CRC.