Decisions in Economics and Finance (2019) 42:21–49
https://doi.org/10.1007/s10203-019-00255-w

S.I.: MATHEMATICAL & STATISTICAL METHODS FOR ACTUARIAL SCIENCES & FINANCE

Variable annuities with a threshold fee: valuation, numerical implementation and comparative static analysis

Anna Rita Bacinello[1] · Ivan Zoccolan[2]

Abstract
In this paper we deal with a variable annuity which provides guarantees at death and maturity financed through the application of a state-dependent fee structure of the threshold type. Our first aim is to test the use of least squares Monte Carlo methods (LSMC) for the numerical implementation of the valuation model. In fact, special care is needed when applying LSMC, due to the shape of the surrender region. To this end we introduce a quite general framework, under which we derive a theoretical result that allows us to stem the numerical errors arising in the regression step of the valuation algorithm. The second aim of the paper is to analyse numerically the interaction between the various contract components, in particular fee rates/thresholds and surrender penalties, under alternative policyholder behaviours. This analysis turns out to be very useful, in particular when addressing the problem of the contract design.

Keywords Variable annuities · State-dependent fees · Surrender option · LSMC

JEL Classification C61 · C63 · G17 · G22

1 Introduction

Variable annuities are very flexible life insurance investment products that can package living and death benefits, essentially with the aim of constructing a post-retirement income endowed with a number of possible guarantees in respect of financial or biometric risks. Typically, 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

✉ Anna Rita Bacinello
bacinel@units.it

1 Department of Economics, Business, Mathematics and Statistics 'B. de Finetti', University of Trieste, Via Università 1, 34123 Trieste, Italy

2 Oracle Italia S.r.l., Cinisello Balsamo, Italy

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.

Guarantees are commonly referred to as GMxBs (guaranteed minimum benefit of type 'x'), where 'x' stands for accumulation (A), death (D), income (I), surrender (S) or withdrawal (W). In particular, GMAB and GMDB provide guarantees in the accumulation phase, prior to retirement, although sometimes the GMDB is offered also after retirement. The GMIB consists of a deferred life annuity, with guarantees either on the annuitized amount or on the annuitization rate, while the GMWB is similar to an income drawdown, entitling the policyholder to make periodical withdrawals from her account, even when there are no more available funds. Finally, the GMSB provides guarantees in case of surrender.

Guarantees are often set in such a way that at least the lump-sum premium is totally recouped. To fix the ideas, consider the case of a variable annuity with both a GMAB and a GMDB maturing at the same date, in which the guarantee is given by the single premium. Even if no GMSB is present, the policyholder is generally allowed to surrender the contract at any time before maturity by receiving a cash amount equal to the account value net of some possible surrender penalty. Then, when the account value is very high, i.e. the guarantee ('Titanic' put option, see Milevsky and Posner 2001) is out of the money, there is a great incentive for the policyholder to surrender the contract, stopping to pay the high fees (proportional to the account value) for an out-of-the-money guarantee, and to buy a new contract, identical to the old one but with an updated, higher guarantee, equal to the surrender benefit. Conversely, when the account value is low, the policyholder pays a low fee for an in-the-money guarantee. Summing up, not only there is an unfair misalignment between costs incurred by the insurer and premiums (fees) to cover them, but also a huge incentive, for policyholders, to abandon their contracts when they become uneconomical, as defaulting in a swap, with a loss for the insurer that does not recover the total costs for the guarantee. In particular, this fact is highlighted in Milevsky and Salisbury (2001), where the surrender penalties are identified not only as a way to force policyholders to keep their contracts alive or, at least, to allow insurers to recoup some of their costs in case of surrender, but also as a way to complete the market enabling the variable annuity to be hedged.

To eliminate the misalignment between costs and fees and to reduce the surrender incentive insurers can adopt the so-called *threshold expense structure*, a special case of *state-dependent fees*, according to which the fees, still proportional to the account value, are, however, paid only if this value is below a fixed threshold, typically equal to the minimum amount guaranteed, i.e. only when the guarantee is in-the-money. This structure has actually been introduced in the market for optional GMDB's by Prudential UK (see Prudential (UK) 2012) and has been first employed by Bae and Ko (2013) in the framework of refracted Brownian motions to price maturity guarantees.

An extensive analysis of this particular kind of state-dependent fees is carried out by Bernard et al. (2014) to price GMAB and GMDB within the framework of geometric Brownian motions and regime-switching lognormal processes for the assets value, as well as deterministic mortality intensity. In their paper sufficient conditions on the fees in order to eliminate the surrender incentive are provided and explored with the aid of some numerical examples. A similar and wider analysis is conducted in MacKay et al. (2017) within the framework of geometric Brownian motion and deterministic mortality intensity. More in detail, the analysis conducted by MacKay et al. (2017) aims at capturing the interaction between fee rates and surrender penalties on the optimal surrender region, in order to design a marketable insurance product for which surrender is never optimal, allowing then to completely ignore the presence of the surrender option in pricing and hedging such product. In Zhou and Wu (2015) probabilistic properties of the total time of deducting fees are derived within a jump-diffusion processes framework. Moreover, it is worth mentioning the paper by Bae and Ko (2010) where, instead, fees are applied when the account value exceeds a given threshold, i.e. when the guarantee is (close to be) out-of-the-money, and the assets price follows a geometric Brownian motion. Finally, a very rich model is offered in the paper by Delong (2014), where pricing and hedging results for variable annuities with GMAB and quite general state-dependent fees (hence, not only based on the threshold expense structure) are derived within the framework of incomplete financial markets and bidimensional Lévy processes.

One of the main conclusions in Bernard et al. (2014) is that the surrender region when fees are state dependent has a different form than when fees are constant,[1] since the optimal surrender strategy is no longer based on a simple threshold but on a corridor, that can be very strict. Although the authors do not include a full analysis of optimal surrenders in the complex case of state-dependent fees, they claim that the particular shape of the surrender region makes least squares Monte Carlo techniques unsuitable to tackle the optimal surrender problem, because the numerical errors would be too significant.

Driven by this argument but believing the intrinsic flexibility of a Monte Carlo based approach should be preserved, we tested, first of all, the application of these techniques to the valuation problem. In doing so, we indeed verified that a straightforward application of them can actually imply non-negligible numerical errors, particularly for low levels of the fee, where the contract value in the presence of the surrender option often turned out to be lower than the one without the option.[2] This fact motivated us to refine the LSMC algorithm in order to reduce, and possibly eliminate, the regression error. Beyond optimizing number and type of basis functions, one of the arrangements that have allowed us to improve the numerical approximation is based on a theoretical result presented in MacKay et al. (2017), where it is proved that it is never optimal to surrender a contract with both a GMAB and a GMDB, and state-dependent fees, when the surrender penalties are decreasing, strictly positive, and the account value is not below the threshold of application of the fee. Although the underlying assumptions in MacKay et al. (2017) are geometric Brownian motion for the assets price and deterministic mortality intensity, we are able to generalize their result; we just require that, under the pricing measure, the discounted assets price is a martingale and, more-

[1] Unless otherwise stated, we use the (general) term 'state-dependent fees' to refer to the (special) case of a threshold expense structure, while the term 'constant fees' refers to the case in which fees are always applied, independently of the policy account value.

[2] In these cases the surrender incentive had been completely eliminated and hence the real surrender option value was 0.

over, the mortality intensity and the death time of the insured are independent of any financial-related variable.

Secondly, we applied the (adjusted) LSMC algorithm in order to analyse numerically the interaction between the various contract components, in particular fee rates/thresholds and surrender penalties, under alternative policyholder behaviours. This analysis, conducted by assuming a quite complex valuation model, turns out to be very useful, in particular when addressing the problem of the contract design.

This paper is structured as follows. In Sect. 2 we describe the structure of the contract. In Sect. 3 we present our valuation framework. Section 4 is devoted to the numerical analysis that is conducted assuming different policyholder behaviours. In particular, in Sect. 4.3 we describe the problems encountered when applying the LSMC and the arrangements adopted to overcome them. Section 5 concludes the paper, and a technical proof is reported in 'Appendix'.

2 The structure of the contract

Consider a single premium variable annuity contract which provides guarantees at death and maturity. We denote by P the single premium, 0 the time of issuance and T the contract maturity, and we assume that the death benefit is paid upon death within the contract maturity. The single premium is invested in a well-diversified mutual fund, and the (net) value of the accumulated investments in this fund is referred to as the *policy account value*. We denote by A_t this value at time t. The cost of the guarantees is recouped through a periodical deduction, called *fee*, from this account. Assuming that the policyholder is not allowed to make partial withdrawals from her account and, for simplicity, that deductions take place continuously over time, we can describe the instantaneous evolution of the account value while the contract is still in force as follows:

$$\frac{\mathrm{d}A_t}{A_t} = \frac{\mathrm{d}S_t}{S_t} - \psi_t \mathrm{d}t, \quad \text{with } A_0 = P. \tag{1}$$

Here S_t denotes the unit value at time t of the reference fund backing the variable annuity and ψ_t denotes the (instantaneous, state-dependent) fee rate. Then, the net return on the account value is obtained by subtracting the (instantaneous) fees from the return of the reference fund.

Both death and maturity benefits contain a minimum guarantee. Hence, the death benefit is given by

$$b_\tau^{\mathrm{D}} = \max\left\{A_\tau, G_\tau^{\mathrm{D}}\right\}, \quad \tau \leq T, \tag{2}$$

while the survival benefit is

$$b_T^{\mathrm{A}} = \max\left\{A_T, G_T^{\mathrm{A}}\right\}, \quad \tau > T. \tag{3}$$

In (2) and (3) we denote by τ the residual lifetime of the policyholder, assumed to be aged x years at inception, and by G_τ^{D} and G_T^{A} the minimum amount guaranteed at death or maturity, respectively.

We assume that the contract can be surrendered at any time before maturity, if the insured is still alive, and that, in case of surrender at time t, the policyholder receives a cash amount, called surrender value. Since the contract does not contain a GMSB, the surrender value is typically given by

$$b_t^{\mathrm{S}} = A_t(1 - \pi_t), \quad t < T \wedge \tau, \tag{4}$$

where π_t is a (state-dependent) penalty rate such that $0 \leq \pi_t \leq 1$ (a.s.) for any $t < T$.

Note that the instantaneous fee rate ψ_t and the penalty rate π_t are stochastic processes, adapted to the relevant filtration that we will introduce formally in the next section. Hence, in principle, they may depend on any observable variable (interest rate, account value, guaranteed amount, reference fund or other traded asset, volatility index, etc.). However, the term 'state-dependent' fees (or penalties) is usually referred to the case in which they depend only on the account value A_t and possibly on the current date t, hence $\psi_t \doteq \psi(A_t, t)$ and $\pi_t \doteq \pi(A_t, t)$. By suitably choosing the functions ψ and π the insurance company can address the policyholder behaviour. For instance, in order to discourage surrender, which is more likely to take place when the account value is high, it can choose a fee rate $\psi(A_t, t)$ decreasing with respect to the account value A_t. A similar result is pursuable through the surrender penalty rate $\pi(A_t, t)$ that may increase with A_t. If instead the insurance company aims at keeping contracts alive as long as possible, it may decide to fix a penalty rate decreasing with respect to t. A similar, although not completely obvious, effect may be the consequence of a time-decreasing fee rate.

Also the guaranteed benefits G^{D} and G^{A} are in general stochastic because they can depend, for example, on past account values, as happens in the case of *ratchet* and *reset* guarantees (see Bacinello et al. 2011).

However, in what follows we restrict ourselves to a narrow, yet relevant, subset of contract space due to its linkages to the extant literature (Bernard et al. 2014; MacKay et al. 2017) and empirical evidence on contracts available in the market (Prudential (UK) 2012). In particular, we assume

$$\psi_t = \varphi 1_{\{A_t < \beta\}}, \tag{5}$$

where φ and β are real numbers and 1_C denotes the indicator of the event C. Then the deductions for fees are made only when the account value is below a given threshold β, i.e. we adopt a *threshold expense* structure. In case of deduction, fees are proportional to the account value according to a fixed rate φ (between 0 and 1). If β were equal, for example, to the minimum amount guaranteed, then the fees would be deducted only when the guarantee is *in-the-money*. In our case the minimum amount guaranteed can change over time, but for simplicity we have taken a barrier β (deterministic and) constant for all the contract duration. Of course, in the degenerate case of $\beta = \infty$ (no barrier) we recover a *constant fee* structure.

Moreover, we assume that death and maturity guarantees are of the *roll-up* type and with the same, constant, roll-up rate δ. Hence,

$$G_\tau^{\mathrm{D}} = P\mathrm{e}^{\delta\tau} \tag{6}$$

and

$$G_T^{\mathrm{A}} = P\mathrm{e}^{\delta T}. \tag{7}$$

Here δ is assumed to be ≥ 0. In particular, when $\delta = 0$ we recover the so-called *return-of-premium* guarantee.

Finally, we assume

$$\pi_t = p(t), \tag{8}$$

where p is a deterministic penalty rate such that $0 \leq p(t) \leq 1$ for any $t < T$.

Therefore, equations (1), (2), (3) and (4) for account value, death, maturity and surrender benefits, respectively, can be rewritten as follows:

$$\frac{\mathrm{d}A_t}{A_t} = \frac{\mathrm{d}S_t}{S_t} - \varphi 1_{\{A_t < \beta\}}\mathrm{d}t, \quad \text{with } A_0 = P, \tag{9}$$

$$b_\tau^{\mathrm{D}} = \max\left\{A_\tau, P\mathrm{e}^{\delta\tau}\right\}, \quad \tau \leq T, \tag{10}$$

$$b_T^{\mathrm{A}} = \max\left\{A_T, P\mathrm{e}^{\delta T}\right\}, \quad \tau > T, \tag{11}$$

$$b_t^{\mathrm{S}} = A_t\left[1 - p(t)\right], \quad t < T \wedge \tau. \tag{12}$$

To conclude this section we observe that all our contract terms can be summarized through a vector containing the single premium P, the maturity T, the specification of the reference fund with (unit) price process S, the roll-up rate δ, the pair (φ, β) characterizing the fee structure, and the penalty function p. In particular, the surrender penalties have a role not at all ancillary since they contribute, together with the fees, both to addressing the policyholder behaviour and to recovering all costs arising from the variable annuity contract.

3 Valuation framework

3.1 Assumptions

We fix a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, Q)$ supporting all sources of financial and biometric uncertainty, where all random variables and processes are defined. The filtration $\mathbb{F} \doteq (\mathscr{F}_t)_{t\geq 0}$ (satisfying the usual conditions of right continuity and completeness and such that $\mathscr{F}_0$ is Q-trivial) represents the flow of information available to the insurer and the policyholder over time. The probability Q is a risk-neutral probability measure selected by the insurer, for pricing purposes, among the infinitely many equivalent martingale measures existing in incomplete arbitrage-free markets. Then the fair value of any security is given by the (conditional) expectation, under Q, of its expected discounted cash-flows, where discounting is performed at the risk-free rate (see, for example, Duffie 2001). In particular, we denote by r_t the instantaneous risk-free rate at time t.

It is natural to assume that the policyholder's residual lifetime τ is an $\mathbb{F}$-stopping time, meaning that at any time t the information carried by $\mathscr{F}_t$ allows us to tell whether death has occurred or not by t. We denote by $\mathbb{H}$ the filtration generated by the death indicator process $\left(1_{\{\tau\leq t\}}\right)_{t\geq 0}$, which equals 0 as long as the individual is alive and jumps to 1 at death, and assume that $\mathbb{F} \doteq \mathbb{G} \vee \mathbb{H}$ for some filtration $\mathbb{G}$ not including $\mathbb{H}$, with $\mathscr{G}_0$ trivial. The intuition is that $\mathbb{G}$ carries all relevant information about biometric and financial risk factors (in particular, security prices and likelihood of death), but does not yield knowledge of τ. More specifically, we take $\mathbb{G} = \mathbb{G}^{\mathrm{F}} \vee \mathbb{G}^{\mathrm{B}}$, where the filtrations $\mathbb{G}^{\mathrm{F}}$ and $\mathbb{G}^{\mathrm{B}}$ pertain to financial and biometric factors, respectively. In particular, we assume that both processes r and S are adapted to $\mathbb{G}^{\mathrm{F}}$. It is also natural to require independence between $\mathbb{G}^{\mathrm{F}}$ and $\mathbb{G}^{\mathrm{B}} \vee \mathbb{H}$. In other words, there is independence between financial- and biometric-related variables.

We define the residual lifetime by setting

$$\tau \doteq \inf\left\{t : \int_0^t \mu_s \mathrm{d}s > \xi\right\}, \tag{13}$$

with μ a $\mathbb{G}^{\mathrm{B}}$-predictable non-negative process and ξ a unit exponential random variable independent of $\mathscr{G}_\infty$. The force of mortality μ_t drives the instantaneous probability of death at time t conditional on survival for an individual aged x at time 0. The probability of survival at time $s > t$, conditional on survival at $t \geq 0$ and on $\mathscr{G}_t$, is given by

$$Q(\tau > s|\tau > t, \mathscr{G}_t) = E\left[\mathrm{e}^{-\int_t^s \mu_v \mathrm{d}v}\middle|\mathscr{G}_t\right] = E\left[\mathrm{e}^{-\int_t^s \mu_v \mathrm{d}v}\middle|\mathscr{G}_t^{\mathrm{B}}\right], \tag{14}$$

while the (conditional) death probability can also be expressed as

$$Q(\tau \leq s|\tau > t, \mathscr{G}_t) = E\left[\int_t^s \mathrm{e}^{-\int_t^y \mu_v \mathrm{d}v}\mu_y \mathrm{d}y\middle|\mathscr{G}_t\right] = E\left[\int_t^s \mathrm{e}^{-\int_t^y \mu_v \mathrm{d}v}\mu_y \mathrm{d}y\middle|\mathscr{G}_t^{\mathrm{B}}\right]. \tag{15}$$

This construction is equivalent to the so-called conditionally Poisson set-up, which means that τ, conditionally on $\mathscr{G}_\infty$ and under the measure Q, is the first jump time of a Poisson inhomogeneous process with intensity $(\mu_t)_{t\geq 0}$. This set-up ensures that any $\mathbb{G}$-martingale is an $\mathbb{F}$-martingale, a property that yields considerable simplifications in pricing formulae (see, in particular, Biffis 2005).

A key element in the valuation of the contract from the insurer's point of view is constituted by the behavioural risk. The policyholder, in fact, can choose among a set of possible actions such as partial or total withdrawal (i.e. surrender), selection of new guarantees, switch between different reference funds, and so on. In particular, in Bacinello et al. (2011) the possible policyholder behaviours are classified, with respect to the only aspect concerning partial or total withdrawals, into three categories, characterized by an increasing level of rationality: *static*, *mixed* and *dynamic*. The variable annuity contract dealt with in Bacinello et al. (2011) is quite general and can contain different types of guarantees, taken alone or combined together. Here instead we consider a more specific contract embedding both a GMDB and a GMAB with the same maturity (not a GMWB), so that the most relevant valuation approaches are the first two, static and mixed. In what follows, we fit their general model to our specific case, taking into account, however, that now we are applying state-dependent fees.

3.2 The static approach

Under this approach it is assumed that the policyholder keeps her contract until its natural termination, that is death or maturity, without making any partial or total withdrawal from her policy account value.

The contract value at time $t < T$, on the set $\{\tau > t\}$, is thus given by

$$V_t = E\left[b_\tau^{\mathrm{D}} \mathrm{e}^{-\int_t^\tau r_v \mathrm{d}v} 1_{\{\tau \le T\}} + b_T^{\mathrm{A}} \mathrm{e}^{-\int_t^T r_v \mathrm{d}v} 1_{\{\tau > T\}} \middle| \mathscr{F}_t \right]. \tag{16}$$

Exploiting the structure of the filtration $\mathbb{F}$ and the conditionally Poisson set-up, we can alternatively express V_t, still on the set $\{\tau > t\}$, as[3]

$$V_t = E\left[\int_t^T b_y^{\mathrm{D}} \mathrm{e}^{-\int_t^y (r_v + \mu_v)\mathrm{d}v} \mu_y \mathrm{d}y + b_T^{\mathrm{A}} \mathrm{e}^{-\int_t^T (r_v + \mu_v)\mathrm{d}v} \middle| \mathscr{G}_t \right]. \tag{17}$$

Unlike equation (16), note that in (17) there are no survival indicators, and discounting is made with the mortality-adjusted discount rate $r + \mu$. Moreover, the expectation in (17) is conditional on the elements of the sub-filtration $\mathbb{G}$.

In some situations V_t can be expressed in closed form. This is the case, for example, of the celebrated single premium contract analysed by Brennan and Schwartz (1976) and Boyle and Schwartz (1977). However, if more sophisticated assumptions do not allow to obtain closed-form formulae, a straightforward application of Monte Carlo simulation can be carried out in order to value the expectation in (16) or (17).

3.3 The mixed approach

Under this approach it is assumed that, at any time of contract duration, the policyholder chooses whether or not to exercise the surrender option, and her decision is aimed at maximizing the current value of the contract pay-off.

We denote by λ the time of surrender. Clearly, early termination can take place only if the insured is still alive and the contract is still in force, i.e. $\lambda < \tau \wedge T$. Conventionally, $\lambda \ge \tau \wedge T$ means instead that surrender never takes place. The time λ is in general a stopping time with respect to the filtration $\mathbb{F}$. Given λ, the contract value at time $t < T$, on the set $\{\tau > t,\ \lambda \ge t\}$, can be expressed as

$$\begin{aligned} V_t(\lambda) = E\Bigg[& b_\tau^{\mathrm{D}} \mathrm{e}^{-\int_t^\tau r_v \mathrm{d}v} 1_{\{\tau \le T \wedge \lambda\}} \\ & + b_T^{\mathrm{A}} \mathrm{e}^{-\int_t^T r_v \mathrm{d}v} 1_{\{\tau > T,\, \lambda \ge T\}} \\ & + b_\lambda^{\mathrm{S}} \mathrm{e}^{-\int_t^\lambda r_v \mathrm{d}v} 1_{\{\lambda < \tau \wedge T\}} \Bigg| \mathscr{F}_t \Bigg]. \end{aligned} \tag{18}$$

[3] See also Bacinello et al. (2009) and Bacinello et al. (2010).

Alternatively, exploiting our previous assumptions, we have also

$$V_t(\lambda) = E\left[\left(\int_t^{\lambda} b_y^{\mathrm{D}} \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y \;+\; b_\lambda^{\mathrm{S}} \mathrm{e}^{-\int_t^\lambda (r_v+\mu_v)\mathrm{d}v}\right) 1_{\{\lambda<T\}} \right.$$
$$\left. + \left(\int_t^{T} b_y^{\mathrm{D}} \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y \;+\; b_T^{\mathrm{A}} \mathrm{e}^{-\int_t^T (r_v+\mu_v)\mathrm{d}v}\right) 1_{\{\lambda\geq T\}} \middle| \mathscr{G}_t \right], \quad (19)$$

where now λ is a stopping time with respect to the sub-filtration $\mathbb{G}$.

Finally, the contract value at time $t < T$, on the set $\{\tau > t,\ \lambda \geq t\}$, is obtained by solving the following optimal stopping problem:

$$V_t = \sup_{\lambda\in\mathbb{T}_t} V_t(\lambda), \quad (20)$$

where $\mathbb{T}_t$ is the set of stopping times taking values in the closed interval $[t, +\infty)$, with respect to the filtration $\mathbb{F}$ if $V_t(\lambda)$ is expressed by (18) or to the sub-filtration $\mathbb{G}$ if instead $V_t(\lambda)$ is given by (19).

Note that the contract value V_t can also be expressed as

$$V_t = \max\left\{V_t^{\mathrm{c}},\ b_t^{\mathrm{S}}\right\}, \quad (21)$$

with V_t^{c} denoting the continuation value, given by

$$V_t^{\mathrm{c}} = \sup_{\lambda\in\mathbb{T}_t^{\mathrm{c}}} V_t(\lambda), \quad (22)$$

where $\mathbb{T}_t^{\mathrm{c}}$ is now the set of stopping times taking values in the open interval $(t, +\infty)$, again with respect to the filtration $\mathbb{F}$ if $V_t(\lambda)$ is expressed by (18) or to the sub-filtration $\mathbb{G}$ if instead $V_t(\lambda)$ is given by (19).

Under the assumptions described in Sects. 2 and 3.1 the following result holds:

Theorem 1 *Let $t < T$ and suppose that $\tau > t$ and $\lambda \geq t$. If $A_t \geq \beta$ and the penalty function p is weakly decreasing on $[t, T)$, then $V_t^c \geq b_t^S$, which implies $V_t = V_t^c$. In particular, if $p(t) > 0$, then $V_t^c > b_t^S$.*

Remark Our Theorem 1 is essentially the same as Proposition 2 in MacKay et al. (2017), but it holds in a more general framework than that implied by a deterministic mortality intensity and a geometric Brownian motion for the assets price. The intuition behind this result is clear: when the account value is not below the barrier β, the guarantees at death and maturity are offered for free; hence, there is no incentive for the policyholder to surrender the contract. In particular, if $p(u) = 0$ for any $u \geq t$, i.e. there are no surrender charges, at least from t onwards, then it could also be that the continuation value is equal to the surrender benefit, implying that continuation and surrender decisions are indifferent. In this case, however, for valuation purposes it can be assumed that surrender does not take place.

The proof of Theorem 1 is supplied in 'Appendix'.

The optimal stopping problem (20) needs to be tackled numerically. In particular, in Sect. 4 we apply the least squares Monte Carlo method for the numerical implementation of the valuation model.

We conclude by observing that the contract value obtained in the mixed approach is, of course, not less than the corresponding value under the static approach (American- versus European-style contract).

3.4 Fair contracts

Note that the initial contract value V_0, given by equations (16), (17) in the static approach, and by equation (20) in the mixed one, depends on all the elements of the vector summarizing the contractual terms. In particular, once maturity, single premium, reference fund and roll-up rate are given, it could be seen as a function of all the quantities involved in the recovery of the costs implied by the contract, namely the fee rate φ and the barrier β, along with the penalty function p when acting in the mixed approach. However, for convenience, in almost all our numerical analyses presented in Sect. 4 we will assume that also the threshold β and the penalty function p are given, so that we look at V_0 as a function of (only) the fee rate φ, say $V_0 \doteq V_0(\varphi)$. We state that the contract is fairly priced if and only if V_0 coincides with the initial premium P:

$$V_0(\varphi) = P. \tag{23}$$

Then a fair fee rate, φ^*, is implicitly defined as a solution of equation (23). Of course, this solution is meaningful only if it lies between 0 and 1, and, in the mixed approach, it must be not less than that obtained in the static approach.

4 Numerical analysis

4.1 Assumptions

We consider the variable annuity contract dealt with in the previous sections. We assume that the age of the policyholder at inception is $x = 50$, the contract duration is $T = 15$ (years), and the single premium is $P = 100$.

We adopt the following three-factor model (under the risk-neutral measure Q) for the financial market:

$$\begin{cases} \mathrm{d}r_t = \alpha^{(r)}[\theta_t^{(r)} - r_t]\mathrm{d}t + \sigma^{(r)}\mathrm{d}W_t^{(r)} \\ \mathrm{d}K_t = \alpha^{(K)}[\theta^{(K)} - K_t]\mathrm{d}t + \sigma^{(K)}\sqrt{K_t}\mathrm{d}W_t^{(K)} \\ \mathrm{d}S_t = r_t S_t \mathrm{d}t + \sqrt{K_t} S_t \mathrm{d}W_t^{(S)} \end{cases}, \tag{24}$$

with $r_0 \in \mathbb{R}$, $K_0 > 0$, $S_0 > 0$ given, and $(W^{(r)}, W^{(K)}, W^{(S)})$ a vector of correlated Wiener processes such that $\mathrm{Cov}(\mathrm{d}W_t^{(i)}, \mathrm{d}W_t^{(j)}) = \rho^{(i,j)}\mathrm{d}t$ for $i, j \in \{r, K, S\}$ and

Table 1 Parameters used in the numerical examples

r	K	S	μ
$r_0 = 0.02$	$K_0 = 0.06$	$S_0 = 100$	$\mu_0 = 0.00568$
$\alpha^{(r)} = 0.5$	$\alpha^{(K)} = 0.8$	$\rho^{(S,r)} = 0.2$	$\alpha^{(\mu)} = 0.000004485659$
$\theta_t^{(r)} = 0.02 - 0.0001\mathrm{e}^{-t}$	$\theta^{(K)} = 0.06$	$\rho^{(S,K)} = -0.5$	$\theta^{(\mu)} = 0.06699157$
$\sigma^{(r)} = 0.01$	$\sigma^{(K)} = 0.4$		$\sigma^{(\mu)} = 0.002995216$
$\rho^{(r,K)} = 0$			

$i \neq j$. Hence, we choose the stochastic model by Hull and White (1990)[4] for the instantaneous interest rate, while we describe the evolution of the assets volatility $\sqrt{K}$ through a mean-reverting square-root process, as in the Heston (1993) stochastic volatility model. Note that this specification for the interest rate, a mean-reverting Gaussian process, has long been criticized in the past due to the possibility of producing negative values with strictly positive probability. However, now it has regained a certain popularity because a long-lasting period of very low (if not even negative) interest rates is currently in force. The same three-factor model has been used also by Kang and Ziveyi (2018) to price a variable annuity with a maturity guarantee[5] and a constant fee structure. In particular, they value the contract under the mixed approach by solving numerically the associated free boundary PDE problem.

For the (stochastic) mortality intensity we assume the following non-mean-reverting (as recommended by Cairns et al. 2006, 2008) square-root process:

$$\mathrm{d}\mu_t = \left[\alpha^{(\mu)} + \theta^{(\mu)}\mu_t\right]\mathrm{d}t + \sigma^{(\mu)}\sqrt{\mu_t}\mathrm{d}W_t^{(\mu)}. \tag{25}$$

Here $W^{(\mu)}$ is a Wiener process independent of the vector $(W^{(r)}, W^{(K)}, W^{(S)})$, and $\mu_0 > 0$ is given. Although being very simple, this affine model, that collapses to the well-known deterministic Gompertz force of mortality when $\alpha^{(\mu)} = \sigma^{(\mu)} = 0$, has the desirable property of producing strictly positive paths with probability 1, provided $\mu_0 > 0$ and $2\alpha^{(\mu)} > \sigma^{(\mu)2}$. Moreover, it allows to get closed-form formulae for the survival probability given in equation (14), that converges to 0 when $s - t$ diverges.[6]

The parameters used in our numerical examples are reported in Table 1. In particular, the parameters of the financial market are those calibrated by Kang and Ziveyi (2018), while we have directly calibrated the parameters of the mortality intensity to the data provided by HMD (2016). The period considered in this study goes from 1985 to 2014. We have used the tables of an Italian male aged 50 years in 1985. We have split the 30-year observation interval into a 20-year training period (from 1985 to 2004) and a 10-year test period (from 2005 to 2014). The training data are used to fit the model, while the test data are used to measure its predictive performance. The expected residual lifetime produced by the calibrated process is equal to 32.6406. Note that,

4 Also known as extended Vasicek (1977) model.

5 Without mortality risk.

6 See, for instance, Fung et al. (2014), or Dacorogna and Apicella (2016).

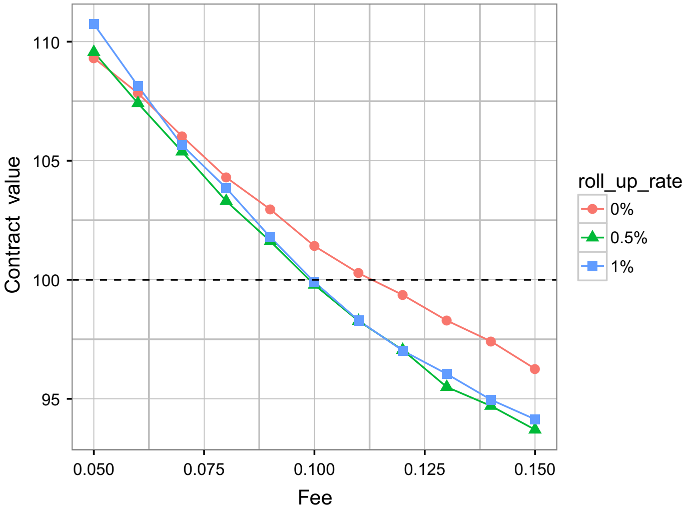

Fig. 1 The initial contract value, under the static approach, versus the state-dependent fee rate, for different roll-up rates δ; single premium $P = 100$, contract maturity $T = 15$, barrier $\beta = Pe^{\delta T}$

although we are acting under the risk-neutral measure, for the mortality we have taken estimates under the historical measure based on data concerning the whole population of males. Then, very likely the resulting survival probabilities will underestimate those of a selected group like that of the purchasers (both males and females) of variable annuity contracts. However, recall that our specific contract provides guarantees in case of early death (or, at the latest, at maturity), so that our assumptions are in line with prudence, i.e. the estimated probabilities contain a safety loading on the mortality (rather than on the longevity) side.

4.2 Results under the static approach

We start by presenting some results under the static approach. To obtain them we resort to Monte Carlo simulation and generate 20000 paths for all stochastic processes involved, with a discretized step of $1/365$ (1 day). We compute the contract value through equation (16), so that we need also to simulate the time of death, through (13).

First of all, we fix a barrier level equal to the minimum amount guaranteed at maturity, $\beta = Pe^{\delta T}$, so that fees are applied only when the maturity guarantee is in-the-money. In Fig. 1 we plot the results obtained in terms of contract value against the fee rate φ for different roll-up rates δ.

Of course, for a given roll-up rate δ the contract value V_0 is decreasing with respect to the fee rate φ. Note that the fair fee rate φ^* is given by the abscissa of the intersection between the contract graph and the dotted horizontal line at level $P = 100$: $\varphi^* = 1130$ basis points (bp) when $\delta = 0$, $\varphi^* = 989$ bp when $\delta = 0.5\%$ and $\varphi^* = 995$ bp when $\delta = 1\%$. It is not a priori clear, instead, which is the effect of the roll-up rate on the

Table 2 The fair fee rate (in bp), under the static approach, for different roll-up rates δ and barrier levels β; single premium $P = 100$, contract maturity $T = 15$

β	$\delta = 0.000$	$\delta = 0.005$	$\delta = 0.010$	$\delta = 0.015$
100	1130	1562	2812	6250
110	703	898	1250	2031
120	527	664	898	1328
130	449	547	742	1055
140	391	488	645	918
150	352	430	566	820

contract value. On the one hand, the higher the δ, the higher the guaranteed amount at death or maturity, but, on the other hand, the higher the δ, the higher the barrier $\beta = Pe^{\delta T}$ so that, for a given fee rate φ, there are more cumulated fees deducted from the account value, that hence results lower. The prevailing effect is not always the same but depends on the level of δ and φ: in particular, the contract value with the intermediate roll-up rate (0.5%) is dominated by that with roll-up 1% for any level of the fees here considered and, apart from the case $\varphi = 500$ bp, it is dominated also by that with a *return-of-premium* guarantee (roll-up 0). Anyway, there is not much difference between contract values with roll-up 0.5% and 1%, while the value of the contract with a *return-of-premium* guarantee, lower than that with $\delta = 1\%$ when $\varphi \leq 600$ bp, is (considerably) more valuable for higher fee levels, resulting in a spread of 135 bp between the corresponding fair rates (see Fig. 1).

Due to the excessively high level of the fair fee rate resulting in the previous examples, we now give up the idea of fixing a barrier equal to the guaranteed amount at maturity and admit the possibility that fees are applied also when the guarantee is (moderately) out-of-the-money. To this end, in Table 2 we report some results concerning the fair fee rate φ^* for different levels of the roll-up rate δ and the barrier β.

As we can see from Table 2, the fair fee rate φ^* increases very fast with the roll-up rate δ when the barrier β is fixed, reaching impractical values (e.g. 1328 bp when $\delta = 1.5\%$ and $\beta = 120$), and decreases with β for fixed levels of δ. For this reason from now on we focus on the case of a *return-of-premium* guarantee, that is also the only case compatible with the current low interest rates environment. In Fig. 2 we compare the contract value obtained with a state-dependent fee structure and a barrier $\beta = 135$ with that implied by constant fees, plotted against different fee rates φ.

From Fig. 2 it is visible that the fair fee rate φ^* required by a state-dependent fee structure is twice as much as that required by constant fees (410 against 205 bp). This is an obvious consequence of two facts: i) differently from constant fees, that are always deducted from the account value, even if the guarantee is deeply out-of-the-money, state-dependent fees are recovered only when the account value is below the barrier, and hence the total time of deducting fees is reduced; ii) the average amount on which fees are applied (account value) is lower (below the barrier), when fees are state dependent. However, it could be difficult to explain these facts to potential customers with a low level of financial literacy, that only perceive the high level of the fees. Then, for commercial reasons, we can try to increase the occupational time of deducting fees

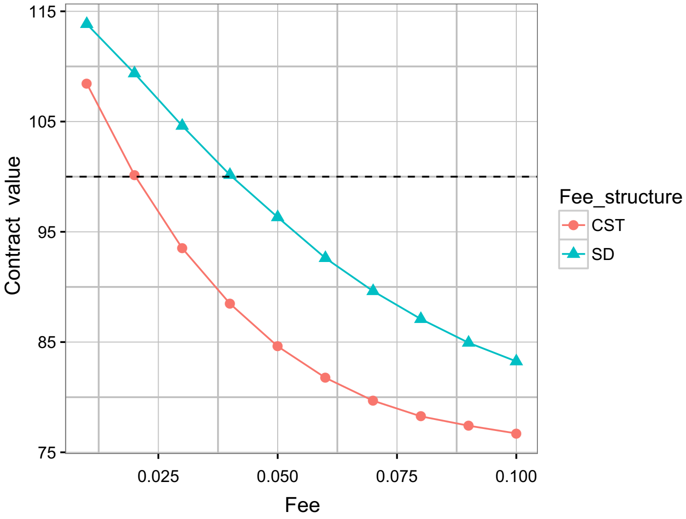

Fig. 2 The initial contract value, under the static approach, versus the fee rate, for both constant and state-dependent fees; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$

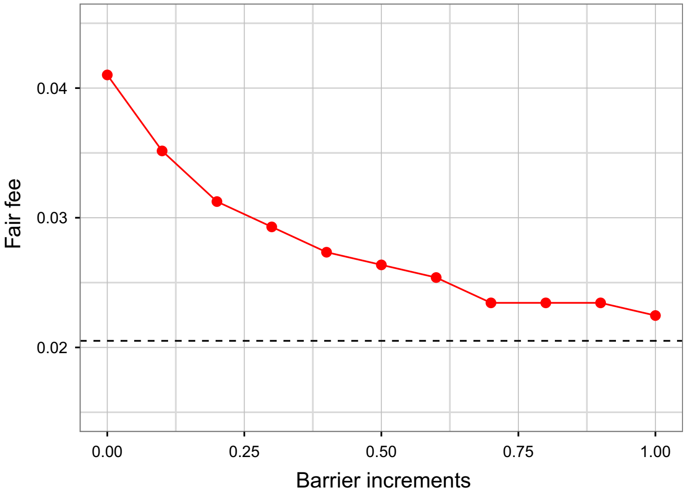

Fig. 3 The fair fee rate φ^*, under the static approach, versus the barrier increment k; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135(1 + k)$

by rebalancing further the barrier β. In particular, we take now $\beta = 135(1 + k)$ and plot, in Fig. 3, the fair fee rate φ^* against the barrier increment k, once again with a *return-of-premium* guarantee. Of course, the fair fee rate is decreasing with the barrier increment k and, when $k \to +\infty$, we recover the fair fee rate obtained with a constant fee structure (see dotted horizontal line).

From Fig. 3 we can argue that, if the insurance company has a target fee rate, e.g. 300 bp, in addition to the initial increase in the barrier from 100 to 135, it should further increase β of about 25%, undermining in this way one of the goals that led to the introduction of state-dependent fees, i.e. the elimination of the misalignment between fees and cost of the guarantees.

4.3 Numerical implementation of the mixed approach

As already mentioned in Sect. 3.3, the optimal stopping problem giving the contract value under the mixed approach needs to be tackled numerically. In particular, in Bernard et al. (2014) it is claimed that the least squares Monte Carlo techniques are unsuitable to solve the problem in the case of state-dependent fees, due to the shape of the surrender region. However, although some drawbacks of Monte Carlo methods are well known,[7] we believe that their intrinsic flexibility, making them practically model independent, constitutes a precious feature. For this reason we tested their application to the solution of our problem. Doing this, we have actually verified that a straightforward application of them is a bit problematic for relatively low levels of the fee rate (usually not over the fair fee rate), i.e. when very likely the surrender incentive has been completely eliminated, leading to a valueless surrender option. In these cases, in fact, the contract value under the static approach turns out to be higher than that under the mixed approach, contradicting the theoretical relation and confirming the claim by Bernard et al. (2014) that numerical errors can be significant. Given this appears to happen only for low levels of the fee, a possible explanation is that the regression tends to underestimate the continuation value, thus inducing surrender even when this is not the optimal decision. This behaviour has been detected by comparing the residuals' plots printed at each regression step for the constant and state-dependent fee cases. While in the constant case the residuals appeared to be balanced between positive and negative values for all regression steps, in the state-dependent case they tended to shift towards positive values in the last few steps. Since the LSMC algorithm proceeds backward, this means that at the very first surrender decision dates the real continuation values were generally much greater than the predicted ones, leading to earlier and sub-optimal terminations of the contract. Therefore, in an attempt to improve the regression, we have tested several methods, such as changing type and number of basis functions, or using different regression techniques (e.g. generalized linear model, ridge regression, the lasso),[8] which, however, have not brought substantial enhancements. In contrast, Theorem 1 has proven to be of some help to this end. From the computational point of view, it allows us to skip the regression step in the LSMC algorithm when $A_t \geq \beta$.[9]

This preliminary analysis was conducted under a simpler framework with constant interest rate, deterministic mortality intensity and asset evolution described through a geometric Brownian motion. That is because in this framework the contract value can

[7] They can be very slow, especially when implying the simulation of various market and biometric quantities over long periods of time, as in the case of variable annuities. Moreover, some stochastic processes are not trivial to simulate efficiently without bias.

[8] For an illustration of these regression techniques see, for instance, Hastie et al. (2009).

[9] A description of the LSMC algorithm for the case of constant fees and more general variable annuity contracts can be found in Bacinello et al. (2011).

Table 3 The initial contract value, under the static and the mixed approaches, for different entry ages x of the policyholder and contract maturities T; single premium $P = 100$, roll-up rate $\delta = 0$, risk-free rate $r = 0.03$, assets volatility $\sigma = 0.165$, surrender penalty $p(t) = 0.05(1 - t/T)^3$ for any $t < T$, barrier $\beta = 150$, mortality intensity $m(y) = \left(1 + 3.5 \cdot 1.075^y\right) \cdot 10^{-4}$

x	T	φ^* (bp)	$V_0^{\text{mixed}**}$	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{static}
50	10	190	99.99	99.65	99.96	99.53
60	10	205	99.99	99.82	99.88	98.43
70	10	237	100.00	99.75	99.84	98.38
50	20	96	99.98	99.25	99.40	99.19
60	20	119	99.98	99.54	99.98	99.38
70	20	163	100.06	99.55	100.03	99.76

Table 4 The initial contract value, under the static and the mixed approaches, for different fee rates φ; single premium $P = 100$, roll-up rate $\delta = 0$, risk-free rate $r = 0.03$, assets volatility $\sigma = 0.165$, surrender penalty $p(t) = 0.05(1 - t/T)^3$ for any $t < T$, barrier $\beta = 150$, age of the insured $x = 50$, contract maturity $T = 10$, mortality intensity $m(y) = \left(1 + 3.5 \cdot 1.075^y\right) \cdot 10^{-4}$

φ (bp)	$V_0^{\text{mixed}**}$	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{static}
50	105.32	104.80	104.86	104.60
100	103.18	102.64	102.82	102.31
150	101.28	100.89	101.05	100.11
200	99.71	99.44	99.60	97.99
250	98.46	98.14	98.22	95.98
300	97.53	97.18	97.31	94.07
350	96.79	96.50	96.51	92.27
400	96.19	95.82	95.94	90.59

be obtained using partial differential equations (PDE) as discussed in MacKay et al. (2017), hence providing us with a benchmark. In what follows we show some results for the contract value under the mixed approach obtained by using PDE ($V_0^{\text{mixed}**}$), LSMC without skipping the regression step when $A_t \geq \beta$ (V_0^{mixed}), and LSMC by skipping this step ($V_0^{\text{mixed}*}$). For comparison, we also report the corresponding values under the static approach (V_0^{static}).

In particular, in Table 3 we reproduce some results obtained by MacKay et al. (2017), that use the PDE approach to compute the contract values and determine also the fair fee rate. More in detail, we compute the contract value when the fee rate is exactly equal to its fair level reported in MacKay et al. (2017) for different policyholder ages at inception x and contract maturities T. The number of simulations is 20000, stochastic processes are simulated with a (forward) discretized step of 1/365 (1 day), for the surrender decision we use a (backward) discretized step of 1/4 (3 months), in the regression we employ Laguerre polynomials up to order 4, and the mortality intensity follows a Makeham law: $\mu_t \doteq m(x + t)$. For additional examples see also Bacinello and Zoccolan (2018).

To produce the results reported in Table 4 we choose the first example of the previous table and analyse the improvement obtained with the LSMC adjustment when different fee rates (both below and above the fair level) are considered. These results seem to indicate a better response for fee rates very close to (and below) their fair level.

Table 5 The initial contract value, under the static and the mixed approaches, for different fee rates φ (in bp) and constant surrender penalties p, along with the resulting fair fee rate φ^* (in bp); single premium $P = 100$, contract maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$

φ	$p = 0.00$		$p = 0.01$		$p = 0.02$		$p = 0.03$		
	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{mixed}	$V_0^{\text{mixed}*}$	V_0^{static}
1	113.26	114.17	112.89	114.14	112.65	113.77	112.42	113.59	114.02
2	107.87	108.57	107.52	108.34	107.51	108.29	107.47	108.28	109.49
3	103.92	104.58	103.48	104.24	103.04	103.80	102.65	103.50	105.15
4	101.81	102.02	101.06	101.46	100.38	100.94	99.71	100.28	100.46
5	100.65	100.91	99.82	100.10	99.08	99.34	98.31	98.56	96.68
6	99.93	100.08	99.06	99.23	98.21	98.38	97.49	97.65	93.00
7	99.51	99.56	98.62	98.70	97.71	97.83	96.82	96.98	90.11
8	99.03	99.10	98.17	98.23	97.30	97.34	96.43	96.46	87.63
φ^*	590	615	485	511	429	450	390	416	410

4.4 Results under the mixed approach

After this preliminary analysis in a very simple framework, we go back to our model described in Sect. 4.1 and present some results under the mixed approach. To obtain them we use again the least squares Monte Carlo technique with 20,000 simulations, we simulate stochastic processes with a (forward) discretized step of $1/365$, for the surrender decision we fix a (backward) discretized step of $1/4$, and the contract value is computed through equations (18) and (20). In the regression, we employ Laguerre polynomials of up to order 3.

In Table 5 we report the results obtained for various fee rates φ and constant surrender penalties $p(t) \equiv p$ for any $t < T$, by assuming a state-dependent fee structure with barrier $\beta = 135$ and roll-up rate $\delta = 0$. In particular, to catch the difference between the contract value under the static and the mixed approaches, as well as the improvement obtained by introducing in the LSMC algorithm the arrangement based on Theorem 1, we show the contract value under the static approach (V_0^{static}) and under the mixed one before and after the adjustment (V_0^{mixed} and $V_0^{\text{mixed}*}$, respectively). In the last row of the table we also display the resulting fair fee rate φ^*.

From Table 5 first of all we can see that the contract value under the mixed approach and, consequently, also the corresponding fair fee rate φ^* are decreasing with the (constant) surrender penalty p, as expected, no matter whether we apply or not the adjustment to the LSMC algorithm. Moreover, as in the static approach, the contract value also decreases with the fee rate φ. As anticipated, the contract value under the mixed approach before the adjustment, V_0^{mixed}, is lower than that under the static approach, V_0^{static}, for (relatively) low levels of the fee. The introduction of the adjustment based on Theorem 1 always improves the results, resulting in $V_0^{\text{mixed}*} > V_0^{\text{mixed}}$ for any level of penalty and fee, although this improvement is more apparent for low levels of the fee, when, very likely, the surrender incentive has been completely eliminated and hence the adjustment turns out to be more helpful in order to stem the

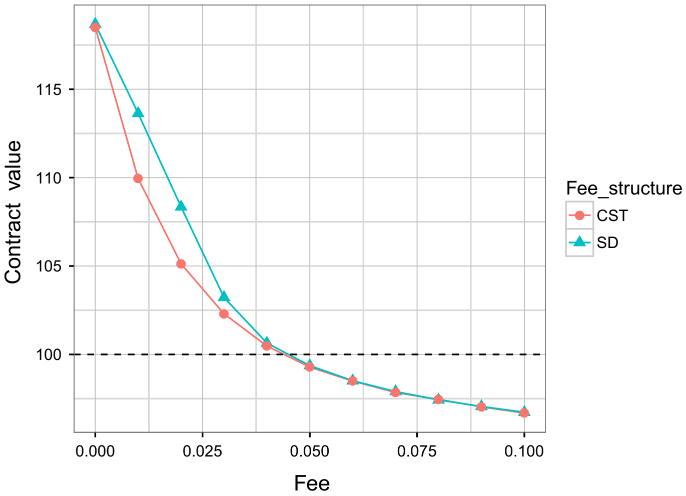

Fig. 4 The initial contract value, under the mixed approach, versus the fee rate, for both constant and state-dependent fees; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$, surrender penalty $p(t) = 2\%$ for any $t < T$

regression error. However, this error is not fully removed even after the adjustment, again for low levels of the fee, especially when combined with high penalties. Anyway, the theoretical relation between the corresponding fair fee rates under static and mixed approach is always preserved after the adjustment, while this is not the case with a surrender penalty of 3% before the adjustment. Finally, if we compare the results between mixed (adjusted) and static approaches in terms of fair fee rates, we have a considerable spread (205 bp)[10] when there are no surrender penalties, and this spread reduces to 101, 40 and 6 bp for surrender penalties of 1%, 2% and 3%, respectively. Then we can clearly perceive the effect of the penalty on the reduction/elimination of the surrender incentive (as happens with the introduction/reduction of the fee threshold β).

Now we only show some results under the mixed approach and use the adjusted algorithm in order to produce them. In particular, in Fig. 4 we plot the contract value against the fee rate φ both in the case of state-dependent fees (again with roll-up 0, barrier 135 and constant penalty 2%) and in that of constant fees.

Unlike in the static approach, from Fig. 4 we can see that the introduction of the surrender option has practically eliminated the difference between the contract value implied by a state-dependent and a constant fee structure, at least for high levels of the fees (over the corresponding fair fee rates). On the other hand, the contract value is exactly the same when no fees are applied (and it coincides with that in the static approach), because in this case guarantees at death and maturity are offered for free and hence it is never convenient to prematurely exit the contract. There remains a perceptible difference between the contract value with state-dependent and constant fees only when the fee rate is low, but this does not seem to influence much the

[10] This spread represents the implied cost of the surrender option.

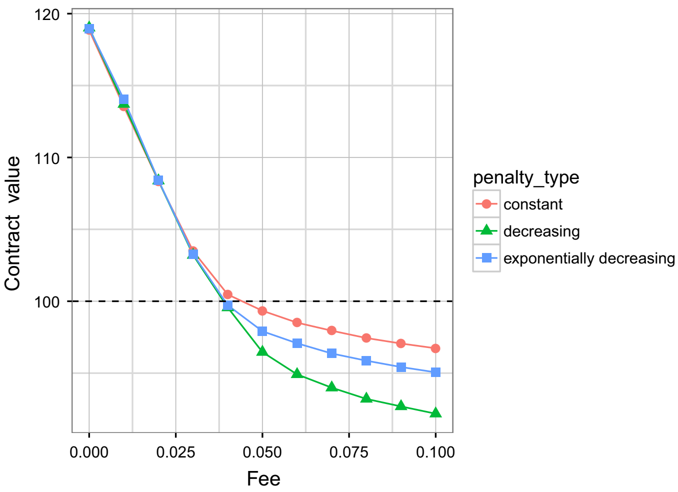

Fig. 5 The initial contract value, under the mixed approach, versus the state-dependent fee rate, for different penalty structures; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$

corresponding fair fee rates that remain very close to each other (430 bp with constant fees and 450 bp with state-dependent ones).

Coming now to another aspect of the contract design, given by the surrender penalty, in Fig. 5 we display the initial contract value versus the (state-dependent) fee rate for three different penalty structures: constant penalty $p(t) = 2\%$, cubically decreasing penalty[11] $p(t) = 0.08(1 - t/T)^3$, and exponentially decreasing penalty $p(t) = 1 - \mathrm{e}^{-0.0405(1-t/T)}$, all for any $t < T$. The last two surrender penalty structures are (special cases of) those employed by MacKay et al. (2017), and their parameters have been fixed in such a way that the average penalty over the 15-year time horizon is 2%. In particular, they both tend to 0 as $t \to T$; the cubically decreasing penalty is four times the constant one when $t = 0$ and rapidly decreases reaching the constant penalty during the sixth year of contract; the level of the exponentially decreasing penalty is almost twice the constant one when $t = 0$ and then decreases more slowly, attaining the constant penalty during the eighth contract year.

From Fig. 5 we can see that for low levels of the fee (below the fair fee rate), all the three graphs are overlapping. This confirms the previous results, obtained in the case of constant penalties: for these fee levels it is never optimal to surrender the contract; hence, the contract value is independent of the penalty structure and is the same as in the static approach. When instead the fee becomes higher, the constant penalty structure becomes less penalizing (high contract values), the exponentially decreasing produces intermediate values of the contract, while the cubically decreasing structure turns out to be more penalizing. The gap between the contract value under the two different decreasing penalty structures becomes higher as the fee rate increases. When surrender is convenient (high levels of the fee), it is then more advantageous for the policyholder to have constant rather than decreasing penalties, thus suggesting that the

[11] In Fig. 4 it is simply referred to as decreasing.

Table 6 Some synthetic statistics on the frequencies of death and surrender (in percentage) and on the distribution of death and surrender times (in years) for fair contracts and different penalty structures $p(t),\ t < T$; single premium $P = 100$, contract maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$

Frequency of exit causes	$p(t) = 2\%$ $\varphi^* = 450$ bp	$p(t) = 1 - e^{-0.0405(1-t/T)}$ $\varphi^* = 392$ bp	$p(t) = 0.08(1 - t/T)^3$ $\varphi^* = 388$ bp
Surrender	65.565	52.570	35.115
Death	5.340	7.450	10.185
Maturity	29.095	39.980	54.700
Average contract duration	5.768	7.996	11.214
Average surrender time	1.502	2.676	6.271
Average death time	7.838	7.947	7.926

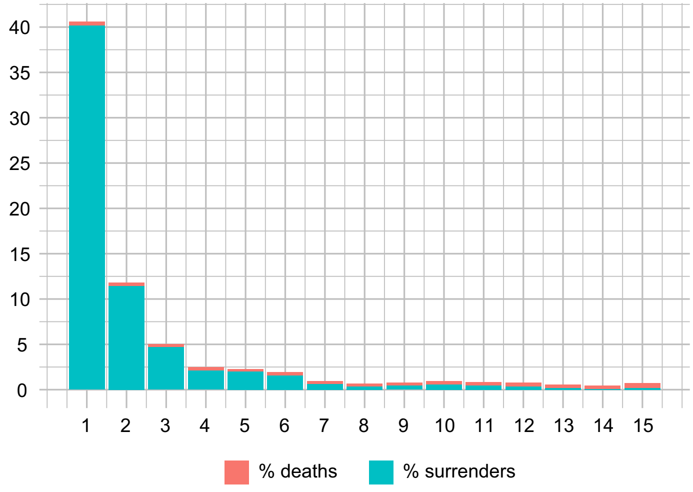

Fig. 6 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$, surrender penalty $p(t) = 2\%$ for any $t < T$, fee rate $\varphi = 450$ bp

convenience to exit the contract is soon enough, before the crossing between constant and decreasing penalties, i.e. while the constant penalty is still the lowest. Finally, the corresponding fair fees under constant, cubically and exponentially decreasing structures are 450, 388 and 392 bp, respectively.

To further investigate the role of different contract configurations (threshold expense structure and surrender penalty structure, in particular) in reducing/eliminating the surrender incentive, as well as to get some information on the profile of expected policyholder behaviour induced by such configurations, in Table 6 we report some synthetic results on the distribution of death and surrender times corresponding to the penalty structures used in Fig. 5, while in Figs. 6, 7 and 8 we display the surrender and death distribution over time corresponding to each penalty structure. This analysis is conducted only for fair contracts. In fact, the insurance company is obviously interested in seeing what happens to contracts offered for sale that should be attractive to both

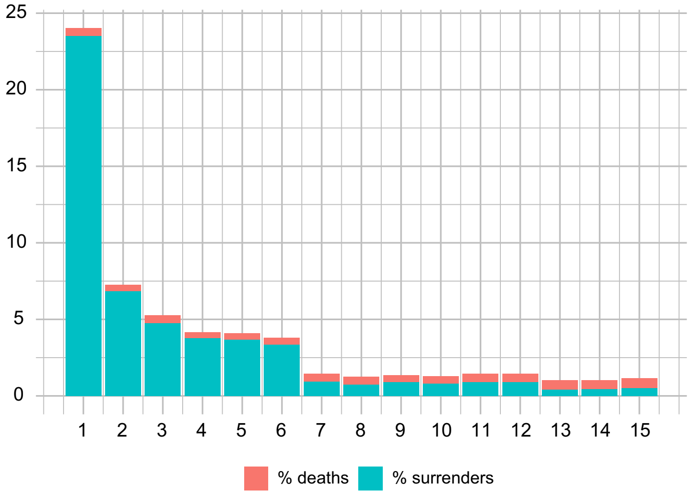

Fig. 7 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$, surrender penalty $p(t) = 1 - e^{-0.0405(1-t/T)}$ for any $t < T$, fee rate $\varphi = 392$ bp

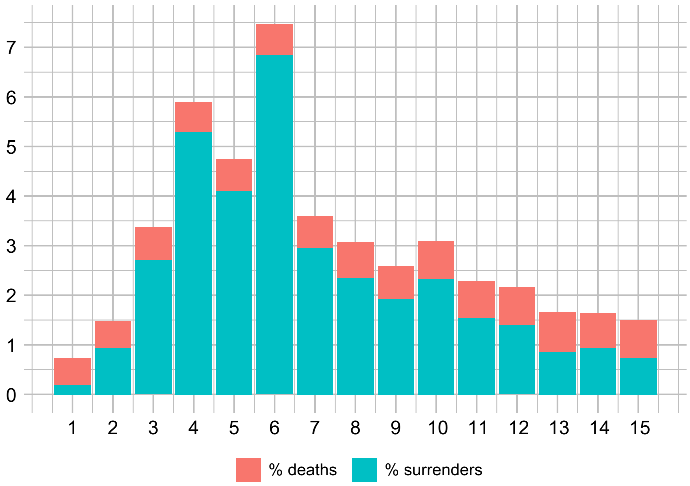

Fig. 8 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = 135$, surrender penalty $p(t) = 0.08(1 - t/T)^3$ for any $t < T$, fee rate $\varphi = 388$ bp

contractors, hence fair. On the other hand, we have already seen, from the previous analysis, that when fees are too low (under the fair fee level) the surrender incentive has been completely eliminated leading to the equality between contract values under static and mixed approaches.

The results reported in Table 6 and Fig. 6 for the case of constant penalties are quite worrying, not only for the high percentage of exits for surrender (more than 65%), but above all for the fact that a high percentage of them is concentrated in the first year of contract. In fact, in around 40% of the simulated paths surrender takes place in the first year, amounting at 61.32% of paths with surrender as exit cause. The resulting average contract duration is only 5.768 years. On the one hand, it could be expected that, with constant penalties, surrender is more likely to take place in the beginning, because the surrender penalty is not high enough to overcome fees that will be applied for all the remaining contract duration (provided the account value is below the barrier). Moreover, in the beginning the surrender incentive is not eliminated because we have fixed a barrier which is 35% higher than the initial account value. On the other hand, a lower level of the barrier, for instance equal to the initial account value, has proven to be completely impractical for the contract marketability due to the high level of the implied fair fee rate (1130 bp in the static approach, see Table 2). Moreover, although the *American-style* contingent-claim approach that we have followed to price the contract provides the worst-case scenario from the insurer's viewpoint, it would be dangerous to offer the contract at a fee rate lower than its fair level because the policyholder has the right to act according to this scenario.

In order to reduce this phenomenon the insurance company could try to modify the penalty structure by adopting a time-decreasing one. From Table 6 and Fig. 7 we can see that the exponentially decreasing penalty is not very effective for this purpose. The percentage of exits for surrender remains still very high (more than 52%), and again around 23.5% of simulated paths imply surrender in the first year of contract (44.75% of paths ending with this exit cause). The average contract duration increases more than 2 years, reaching quota 7.996.

The cubic decreasing penalty, instead, turns out to be more effective. In this case the percentage of surrenders reduces to 35.115%, still high but not so problematic as before because only 0.185% of simulated paths ends with surrender in the first year of contract. The skewness of the surrender distribution is strongly reduced, with a modal value in the sixth year of contract, when the cubic penalty crosses the level of 2% of the constant one. Anyway, in this year the percentage of simulated paths ending with surrender is less than 7%, amounting at 19.52% of exits for this cause. Finally, the average contract duration, 11.214 years, remains completely acceptable (see Table 6 and Fig. 8).

The above analysis shows that the penalty structure may be very effective in order to address the surrender behaviour of the policyholder. Then one way to improve the final result could be to further increase the surrender penalties, particularly in the first year of contract.[12] This result would be reinforced by a corresponding reduction in the fair fee rate. However, some caution is still needed because too high surrender penalties could disappoint potential customers and eventually turn back on the marketability of the insurance product.

The adoption of a threshold expense structure, instead, has not eliminated the surrender incentive, unless the contract is underpriced or the surrender penalty is increased.[13] A question that arises naturally is therefore the following: What would have happened without a barrier? To verify if the introduction of the barrier, combined with the previous penalty functions, has at least reduced (although not eliminated) surrenders, and

[12] There are some variable annuity providers, e.g. Intesa Sanpaolo Life, that do not even admit surrender in the first year of contract. This would be equivalent to fixing $p(t) = 1$ for $t \leq 1$.

[13] See also Table 5, from which it can be inferred that a constant penalty just over 3% allows to eliminate the surrender incentive also for fair contracts.

Table 7 Some synthetic statistics on the frequencies of death and surrender (in percentage) and on the distribution of death and surrender times (in years) for fair contracts and different penalty structures $p(t),\ t < T$; single premium $P = 100$, contract maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = \infty$

Frequency of exit causes	$p(t) = 2\%$ $\varphi^* = 430$ bp	$p(t) = 1 - e^{-0.0405(1-t/T)}$ $\varphi^* = 360$ bp	$p(t) = 0.08(1 - t/T)^3$ $\varphi^* = 320$ bp
Surrender	68.275	59.590	49.695
Death	4.990	6.405	8.035
Maturity	26.735	34.005	42.270
Average contract duration	5.411	7.024	8.931
Average surrender time	1.482	2.380	3.953
Average death time	7.799	7.879	7.796

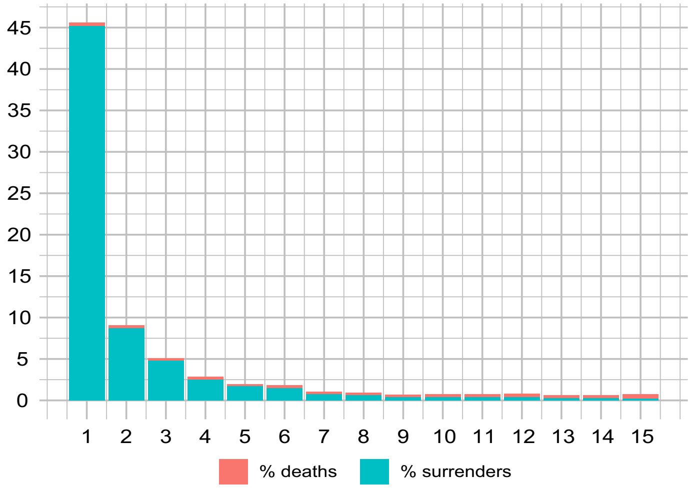

Fig. 9 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = \infty$, surrender penalty $p(t) = 2\%$ for any $t < T$, fee rate $\varphi = 450$ bp

to what extent, in what follows we repeat the analysis as before, but now with constant fees. In particular in Table 7 we report our synthetic results, and in Figs. 9, 10 and 11 we display the surrender and death distributions over time.

Comparing Table 7 with Table 6 we can see that the introduction of the barrier actually reduces the surrender exits and increases the average contract duration for all three penalty structures. However, this effect is very limited in the case of a constant penalty, where the surrender exits decrease of only 2.71%, with an increase in the contracts reaching maturity of 2.36%. The average contract duration increases by just over 4 months, and the average surrender time stretches for about 7 days. In the case of exponentially decreasing penalties this effect is a bit more substantial: the percentage of surrenders reduces of 7.02%, that of contracts reaching maturity increases of 5.975%, the average contract duration increases of almost 1 year and the average surrender time of about 3 and 1/2 months. Finally, the effect is rather substantial in the case

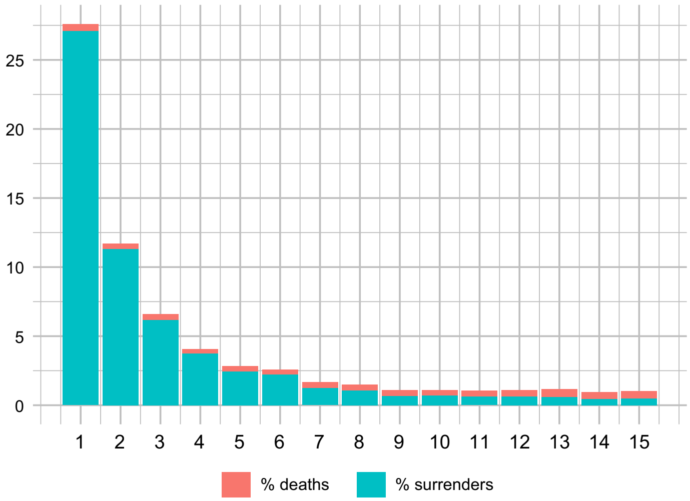

Fig. 10 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = \infty$, surrender penalty $p(t) = 1 - \mathrm{e}^{-0.0405(1-t/T)}$ for any $t < T$, fee rate $\varphi = 392$ bp

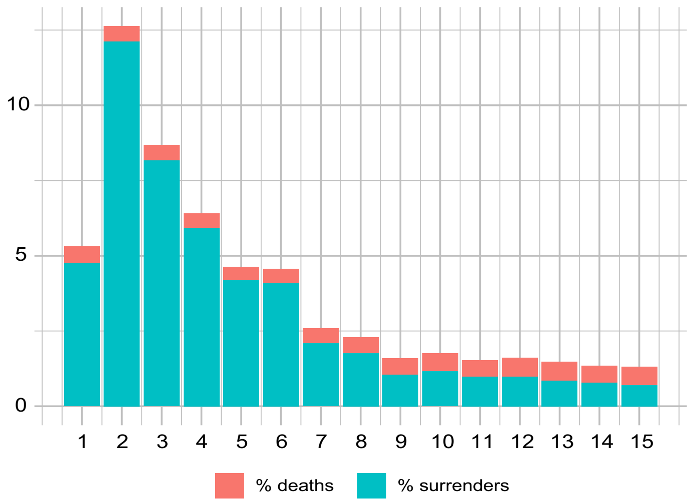

Fig. 11 Annual distribution of surrender and death; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 0$, barrier $\beta = \infty$, surrender penalty $p(t) = 0.08(1 - t/T)^3$ for any $t < T$, fee rate $\varphi = 388$ bp

of cubically decreasing penalties, where the surrender exits decrease of 14.58%, the contracts reaching maturity increase of 12.43%, the average contract duration and the average surrender time increase by more than 2 years and 3 months and almost 2 years and 4 months, respectively. Hence, once again, the adoption of state-dependent fees is much more effective if combined with a suitable penalty structure.

As far as the surrender distribution is concerned, we notice that the introduction of the barrier does not move the modal point in the case of constant and exponentially decreasing penalties, that remains still in the first year of contract. In particular with constant penalties the frequency of simulated paths ending with surrender during the first year of contract passes from 45.225 to 40.205% due to the introduction of the barrier (that is from 66.24 to 61.32% of the exits for surrender),[14] and with exponential penalties from 27.11 to 23.525% of the total paths (respectively, from 45.49 to 44.75% of the paths ending with surrender).[15] In the case of cubically decreasing penalty, instead, the introduction of the barrier allows to move the modal point forward, from the second to the sixth year of contract, with 12.12% and 6.855%, respectively, of simulated paths ending with surrender in these years (24.39% and 19.52%, respectively, of total surrenders).[16]

Another relevant aspect is linked to possible adverse selection introduced by the surrender option. In particular, in our model, one could expect a deterioration in the health conditions of the insureds who do not give up their contracts. In fact, surrender is only driven by endogenous reasons, aimed at maximizing, under the risk-neutral measure, the value of the contract pay-off, no matter if the benefit is directly paid to the policyholder or to her heirs. Then, if the insured is in a very bad health status, it may be convenient to 'wait for death' (rather than surrendering the contract) in order to get the entire account value, or the guaranteed amount if higher (instead of the lower, due to the penalty, surrender benefit). Conversely, if the policyholder feels in excellent health, it may be convenient to suffer the surrender penalty rather than the application of fees for a (supposedly) long period of time.

The adverse selection phenomenon is addressed in the paper by Benedetti and Biffis (2013) that models and analyses its dynamics by representing the conditional survival probabilities of policyholders through a frailty process. Although this analysis is beyond the scope of our paper, and we are well aware that it should be conducted under the physical measure, we conclude this section by giving some information on the frequencies of death, with and without surrender, observed in our simulations (under the risk-neutral measure).

The probability of death within the contract maturity implied by our model is equal to 13.710%. The frequency of death in our simulation results under the static approach (i.e. when surrender is not admitted) is 13.955%. Under the mixed approach, instead, and in the last six examples reported in Tables 6, 7 and Figs. 6, 7, 8, 9, 10, 11, this frequency reaches on average the level of 15.744% (minimum 15.507%; maximum 15.973%) if we, very roughly, measure the risk exposure during the 15 years of contract duration by subtracting from the total number of simulated paths those ending with surrender. Such method, of course, overestimates the real frequency because also paths ending with surrender have been exposed to risk for a while. On the other hand, if we consider also the surrender paths, by weighting them with their time of exposure to risk, we get an average frequency of 12.905% (minimum 12.673%; maximum 13.114%), that very likely provides an underestimation of the real one due to the fact that these paths have been exposed to risk in the first years of contract (i.e. at younger ages), when the probability of death is presumably lower. Anyway, we do not detect any

[14] See Figs. 9 and 6.

[15] See Figs. 10 and 7.

[16] See Figs. 11 and 8.

trend, and in particular any type of monotonicity of these frequencies with respect to the number of surrenders, so that it would be completely hazardous to draw some conclusion from these results, that we show just for information.

5 Summary and conclusions

In this paper we have proposed a quite general Monte Carlo-based valuation model for variable annuities providing guarantees at death and maturity and financed through the application of a state-dependent fee structure of the threshold type. The interaction among fee rates, death/maturity guarantees, fee thresholds and surrender penalties under alternative policyholder behaviours has been extensively analysed from a numerical point of view, letting us to better finalize the contract design. For the numerical implementation of the valuation model we have used Monte Carlo and least squares Monte Carlo methods. In particular, a straightforward application of LSMC techniques has turned out to be a bit problematic for low levels of the fee, due to the shape of the surrender region, but a suitable arrangement of the LSMC algorithm based on a theoretical result first derived in MacKay et al. (2017) and here generalized has allowed us to stem this problem.

Our analysis suggests that the adoption of a threshold expense structure can be effective in order to achieve a first goal, that is to reduce/eliminate the surrender incentive, especially if combined with suitable surrender penalties. However, this may come at the cost of seriously compromising the marketability of the variable annuity product due to the high level of the fee rate required to price it fairly. As we have seen from the numerical examples, this effect can be damped by suitably increasing the barrier level, but in this way the second goal of state-dependent fees, that is to eliminate the misalignment between fees and cost of the guarantees, can be hampered. Hence, the problem of choosing the contract parameters, trying to 'optimize' the trade-off between them, becomes very crucial, particularly in this period of low interest rates. In this sense our numerical analysis can be helpful. An alternative way to partially pursue the above-mentioned goals without charging fees at an excessively high rate could be to keep constant the total (instantaneous) fees once the account value reaches (and exceeds) the barrier β. More precisely, this could be achieved by defining the state-dependent fee rate ψ_t introduced in Sect. 2 as $\psi_t = \varphi \cdot (A_t \wedge \beta)/A_t$, so that it would preserve continuity with respect to the account value, as in the case of constant fees. In particular, ψ_t would still converge to 0 when the account value A_t diverges, without, however, jumping immediately to 0 when A_t reaches β (from below).

Acknowledgements We are very grateful to two anonymous referees for providing us with focused comments and suggestions that led to an enriched version of the paper.

Appendix: Proof of Theorem 1

For convenience, we let

$$b_u^{\mathrm{L}} = b_u^{\mathrm{S}} 1_{\{u<T\}} + b_T^{\mathrm{A}} 1_{\{u \geq T\}}, \tag{26}$$

so that we can rewrite equation (19) in the following, more compact, way:

$$V_t(\lambda) = E\left[\int_t^{\lambda\wedge T} b_y^{\mathrm{D}} \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y \, + \, b_\lambda^{\mathrm{L}} \mathrm{e}^{-\int_t^{\lambda\wedge T} (r_v+\mu_v)\mathrm{d}v} \middle| \mathscr{G}_t\right]. \tag{27}$$

Recall that $A_t \geq \beta$, and define

$$\varepsilon = \inf\{u \in (t, T) : A_u < \beta\} \wedge T. \tag{28}$$

In particular, $\varepsilon \leq T$. Moreover, $A_u \geq \beta \;\; \forall u \in [t, \varepsilon)$, so that, 'before' ε, the fee is not applied and hence the discounted account value behaves as a martingale. Of course, ε is a stopping time with respect to the filtration $\mathbb{G}$, i.e. $\varepsilon \in \mathbb{T}_t^{\mathrm{c}}$. Then the continuation value satisfies

$$V_t^{\mathrm{c}} = \sup_{\lambda\in\mathbb{T}_t^{\mathrm{c}}} V_t(\lambda) \geq V_t(\varepsilon) = E\left[\int_t^{\varepsilon} b_y^{\mathrm{D}} \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y \, + \, b_\varepsilon^{\mathrm{L}} \mathrm{e}^{-\int_t^{\varepsilon} (r_v+\mu_v)\mathrm{d}v} \middle| \mathscr{G}_t\right].$$

Observe that

$$b_y^{\mathrm{D}} = \max\left\{A_y, P\mathrm{e}^{\delta y}\right\} \geq A_y \geq A_y\,[1 - p(y)] \geq A_y\,[1 - p(t)] \quad \forall y \in [t, T],$$

where, for convention, we set $p(T) = 0$. The second inequality is due to the fact that $p(y) \geq 0$ and the third to the (weak) monotonicity of the surrender charge function p. In particular, if $p(t) > 0$ then $b_y^{\mathrm{D}} > A_y\,[1 - p(t)]$. Similarly,

$$b_\varepsilon^{\mathrm{L}} = A_\varepsilon\,[1 - p(\varepsilon)]\, 1_{\{\varepsilon<T\}} + \max\left\{A_\varepsilon, P\mathrm{e}^{\delta\varepsilon}\right\} 1_{\{\varepsilon=T\}} \, \geq \, A_\varepsilon\,[1 - p(t)]\,.$$

All this implies

$$V_t^{\mathrm{c}} \geq [1 - p(t)]\, E\left[\int_t^{\varepsilon} A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y \, + \, A_\varepsilon \mathrm{e}^{-\int_t^{\varepsilon} (r_v+\mu_v)\mathrm{d}v} \middle| \mathscr{G}_t\right], \tag{29}$$

again with a strict inequality if $p(t)>0$. Now we have:

$$\begin{aligned}E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\Big|\mathscr{G}_t\right] &= E\left[E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\Big|\varepsilon, \mathscr{G}_t\right]\Big|\mathscr{G}_t\right]\\ &= E\left[E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon r_v\mathrm{d}v}\Big|\varepsilon, \mathscr{G}_t\right]E\left[\mathrm{e}^{-\int_t^\varepsilon \mu_v\mathrm{d}v}\Big|\varepsilon, \mathscr{G}_t\right]\Big|\mathscr{G}_t\right]\\ &= E\left[A_t Q(\tau > \varepsilon|\tau > t, \varepsilon, \mathscr{G}_t)\Big|\mathscr{G}_t\right]\\ &= A_t E\left[Q(\tau > \varepsilon|\tau > t, \varepsilon, \mathscr{G}_t)\Big|\mathscr{G}_t\right]. \end{aligned} \tag{30}$$

In particular, (30) follows from the law of iterated expectations, the stochastic independence of the mortality intensity μ from the financial variables r and A, and the martingale property of the discounted account value when working 'before' ε. Furthermore, with similar algebraic manipulations we obtain:

$$\begin{aligned}&E\left[\int_t^\varepsilon A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\mathrm{d}y\Big|\mathscr{G}_t\right] = E\left[E\left[\int_t^\varepsilon A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\mathrm{d}y\Big|\varepsilon, \mathscr{G}_t\right]\Big|\mathscr{G}_t\right]\\ &= E\left[\int_t^\varepsilon E\left[A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\Big|\varepsilon, \mathscr{G}_t\right]\mathrm{d}y\Big|\mathscr{G}_t\right]\\ &= E\left[\int_t^\varepsilon E\left[A_y \mathrm{e}^{-\int_t^y r_v\mathrm{d}v}\Big|\varepsilon, \mathscr{G}_t\right]E\left[\mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\Big|\varepsilon, \mathscr{G}_t\right]\mathrm{d}y\Big|\mathscr{G}_t\right]\\ &= E\left[\int_t^\varepsilon A_t E\left[\mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\Big|\varepsilon, \mathscr{G}_t\right]\mathrm{d}y\Big|\mathscr{G}_t\right]\\ &= A_t E\left[E\left[\int_t^\varepsilon \mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\mathrm{d}y\Big|\varepsilon, \mathscr{G}_t\right]\Big|\mathscr{G}_t\right]\\ &= A_t E\left[Q(\tau \le \varepsilon|\tau > t, \varepsilon, \mathscr{G}_t)\Big|\mathscr{G}_t\right]. \end{aligned} \tag{31}$$

Finally, combining (31) and (30) with (29) gives

$$V_t^{\mathrm{c}} \ge [1 - p(t)]\, A_t = b_t^{\mathrm{S}}. \ \square$$

References

Bacinello, A., Zoccolan, I.: Variable annuities with state-dependent fees. In: Corazza, M., Durbán, M., Grané, A., Perna, C., Sibillo, M. (eds.). Springer, pp. 75–80 (2018)

Bacinello, A., Biffis, E., Millossovich, P.: Pricing life insurance contracts with early exercise features. J. Comput. Appl. Math. **233**(1), 27–35 (2009)

Bacinello, A., Biffis, E., Millossovich, P.: Regression-based algorithms for life insurance contracts with surrender guarantees. Quant. Finance **10**(9), 1077–1090 (2010)

Bacinello, A., Millossovich, P., Olivieri, A., Pitacco, E.: Variable annuities: a unifying valuation approach. Insur. Math. Econ. **49**(3), 285–297 (2011)

Bae, T., Ko, B.: On pricing equity-linked investment products with a threshold expense structure. Korean J. Appl. Stat. **23**(4), 621–633 (2010)

Bae, T., Ko, B.: Pricing maturity guarantee under a refracted Brownian motion. Lobachevskii J. Math. **34**(3), 234–247 (2013)

Benedetti, D., Biffis, E.: Insurance contract design and endogenous frailty. Working paper, Imperial College London (2013)

Bernard, C., Hardy, M., MacKay, A.: State-dependent fees for variable annuity guarantees. ASTIN Bull. **44**(3), 559–585 (2014)

Biffis, E.: Affine processes for dynamic mortality and actuarial valuations. Insur. Math. Econ. **37**(3), 443–468 (2005)

Boyle, P., Schwartz, E.: Equilibrium prices of guarantees under equity-linked contracts. J. Risk Insur. **44**(4), 639–660 (1977)

Brennan, M., Schwartz, E.: The pricing of equity-linked life insurance policies with an asset value guarantee. J. Financ. Econ. **3**(3), 195–213 (1976)

Cairns, A., Blake, D., Dowd, K.: Pricing death: framework for the valuation and securitization of mortality risk. ASTIN Bull. **36**(1), 79–120 (2006)

Cairns, A., Blake, D., Dowd, K.: Modelling and management of mortality risk: a review. Scand. Actuar. J. **2**(3), 79–113 (2008)

Dacorogna, N., Apicella, G.: A general framework for modeling mortality to better estimate its relationship to interest rate risks. SCOR Papers 39. www.scor.com/en/file/19926/download?token=tmVNBDDd. Accessed 20 July 2018 (2016)

Delong, L.: Pricing and hedging of variable annuities with state-dependent fees. Insur. Math. Econ. **58**, 24–33 (2014)

Duffie, D.: Dynamic Asset Pricing Theory, 3rd edn. Princeton University Press, Princeton (2001)

Fung, M., Ignatieva, K., Sherris, M.: Systematic mortality risk: an analysis of guaranteed lifetime withdrawal benefits in variable annuities. Insur. Math. Econ. **58**, 103–115 (2014)

Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference and Prediction, 2nd edn. Springer, New York (2009)

Heston, S.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**(2), 327–343 (1993)

HMD Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). www.mortality.org or www.humanmortality.de. Accessed 20 July 2018 (2016)

Hull, J., White, A.: Pricing interest rate derivative securities. Rev. Financ. Stud. **3**(4), 573–592 (1990)

Kang, B., Ziveyi, J.: Optimal surrender of guaranteed minimum maturity benefits under stochastic volatility and interest rates. Insur. Math. Econ. **79**, 43–56 (2018)

MacKay, A., Augustyniak, M., Bernard, C., Hardy, M.: Risk management of policyholder behavior in equity-linked life insurance. J. Risk Insur. **84**(2), 661–690 (2017)

Milevsky, M., Posner, S.: The Titanic option: valuation of the guaranteed minimum death benefit in variable annuities and mutual funds. J. Risk Insur. **68**(1), 93–128 (2001)

Milevsky, M., Salisbury, T.: The real option to lapse a variable annuity: Can surrender charges complete the market? In: Conference Proceedings of the 11th Annual International AFIR Colloquium (2001)

Prudential(UK). Key features of the flexible investment plan (no initial charge and initial charge option)—additional investments. www.pru.co.uk/pdf/FIPK10080.pdf. Accessed 10 July 2018 (2012)

Vasicek, O.: An equilibrium characterization of the term structure. J. Financ. Econ. **5**(2), 177–188 (1977)

Zhou, J., Wu, L.: The time of deducting fees for variable annuities under the state-dependent fee structure. Insur. Math. Econ. **61**, 125–134 (2015)