Dipartimento di scienze economiche, aziendali, matematiche e statistiche “Bruno de Finetti”

Research Paper Series, N. 1, 2017

Variable Annuities with State-Dependent Fees: Valuation, Numerical Implementation, Comparative Static Analysis and Model Risk

Anna Rita Bacinello
Department of Economics, Business, Mathematics and Statistics ‘B. de Finetti’, University of Trieste
Via Università 1, 34123 Trieste, Italy

Ivan Zoccolan
Oracle Italia S.r.l.
V.le Fulvio Testi 136, 20092 Cinisello Balsamo (MI)

UNIVERSITÀ DEGLI STUDI DI TRIESTE

Research Paper Series

Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche
"Bruno de Finetti"
Piazzale Europa 1
34127, Trieste
Tel. +390405587927
Fax +390405587033
http://www.deams.units.it

EUT Edizioni Università di Trieste
Via E.Weiss, 21 - 34128 Trieste
Tel. +390405586183
Fax +390405586185
http://eut.units.it
eut@units.it

ISBN: 978-88-8303-856-3

Research Paper Series, N. 1, 2017

Variable Annuities with State-Dependent Fees: Valuation, Numerical Implementation, Comparative Static Analysis and Model Risk

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

Ivan Zoccolan
Oracle Italia S.r.l.
V.le Fulvio Testi 136, 20092 Cinisello Balsamo (MI)

ABSTRACT[1]

In this paper we consider a variable annuity which provides guarantees at death and maturity financed through the application of a state-dependent fee structure, as defined first in Bae and Ko (2013) and extensively analysed in Bernard et al. (2014) and MacKay et al. (2017). We propose a quite general valuation model for such guarantees, along the lines of Bacinello et al. (2011). We then analyse numerically the interaction between fee rates, death/maturity guarantees, fee thresholds and surrender penalties under alternative model assumptions and policyholder behaviours. This allows us to get also some interesting insights into the model risk. Since the assumptions adopted in the numerical analysis are not at all trivial, we resort to Monte Carlo and Least Squares Monte Carlo methods (LSMC) for the numerical implementation of the valuation model. In particular, special care is needed in the application of LSMC, due to the shape of the surrender

[1]**Corresponding author:** Anna Rita Bacinello, Department of Economics, Business, Mathematics and Statistics 'B. de Finetti', University of Trieste, Via Università 1, 34123 Trieste, Italy, email: bacinel@units.it; tel: (+39) 040 558 7113; fax: (+39) 040 558 7033

region. We are able to stem the numerical errors arising in the regression step by using suitable arrangements of the LSMC valuation algorithm.

KEYWORDS: Variable annuities; State-dependent fees; Surrender option; LSMC; Model risk.

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 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*, or *state-dependent fees*, according to which the fees, still proportional to the account value, are however paid only if this value is below a given threshold, typically 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 (UK)) 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 motions and deterministic mortality intensity. This analysis 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 remarkable 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, since the optimal surrender strategy is no longer based on a simple threshold but on a corridor, that can be very strict. However, the authors do not include a full analysis of optimal surrenders in the complex case of state-dependent fees and, moreover, 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, and with the hope of not having to give up the intrinsic flexibility of Monte Carlo methods, we have tried, first of all, to apply these techniques to the valuation problem. We have verified that a straightforward application of them can actually imply significant numerical errors, for low levels of the fee, since the contract value in presence of the surrender option often turned out to be below that without the option.[2] This fact has then induced us to introduce some arrangements in the LSMC algorithm in order to reduce, and possibly eliminate, the regression error. Beyond optimizing number and type of basis functions, one of the tricks that has allowed us to decidedly 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

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

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 the mortality intensity, if stochastic, is independent of any financially related variable.

Then we have defined a quite general valuation model for variable annuities, with death and survival guarantees and state-dependent fee structure, along the lines of Bacinello et al. (2011), and analysed numerically the interaction between fee rates, death/survival guarantees, fee thresholds and surrender penalties under alternative model assumptions and policyholder behaviours, thus getting also some interesting insights into the model risk.

The paper is structured as follows. In Section 2 we describe the structure of the contract. In Section 3 we present our valuation framework. Section 4 is devoted to the numerical analysis, that is conducted assuming different models for financial and demographic variables and different policyholder behaviours. In particular, in subsection 4c we describe the problems encountered when applying the LSMC method and the arrangements adopted to overcome them. Section 5 concludes the paper. A technical proof is reported in Appendix A, and all figures are collected in Appendix B.

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, T the contract maturity, and 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 the application of a proportional deduction from this account, at a rate denoted by φ (fee rate). However, this deduction is assumed to be made only when the account value is below a given threshold, denoted by β, i.e., we adopt a *state-dependent fee* structure. In particular, if this threshold is equal, e.g., to the minimum amount guaranteed, then the fees are deducted only when the guarantee is *in the money*. As we will see in a moment, the minimum amount guaranteed can change over time, but, for simplicity, we take the barrier β to be constant for all the contract duration. Of course, in the degenerate case of $\beta = \infty$ (no barrier) we recover a *constant fee* structure.

Both death and maturity benefits contain a guarantee of the *roll-up* type, with the same roll-up rate δ. Then the death benefit is given by

$$b_\tau^{\mathrm{D}} = \max\left\{A_\tau, P\mathrm{e}^{\delta\tau}\right\}, \qquad \tau \leq T, \tag{1}$$

while the survival benefit is

$$b_T^{\mathrm{A}} = \max\left\{A_T, P\mathrm{e}^{\delta T}\right\}, \qquad \tau > T. \tag{2}$$

In (1) and (2) we have denoted by τ the residual lifetime of the policyholder, assumed to be aged x years at inception.

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, given by

$$b_t^{\mathrm{S}} = A_t\left(1 - p_t\right), \qquad t < T \wedge \tau, \tag{3}$$

where p_t is a penalty rate, possibly time dependent and such that $0 \leq p_t < 1$ for any t.

3. Valuation framework

a. Assumptions

We fix 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 $\mathbb{F} \doteq (\mathcal{F}_t)_{t\geq 0}$ (satisfying the usual conditions of right continuity and completeness and such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$) 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)). We denote by r_t the instantaneous risk-free rate at time t and by S_t the unit value at time t of the reference fund backing the variable annuity contract.

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 $\mathcal{F}_t$ 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 $\mathcal{G}_0$ 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 arrival of death by setting

$$\tau \doteq \inf\left\{t : \int_0^t \mu_s \mathrm{d}s > \xi\right\}, \tag{4}$$

with μ a $\mathbb{G}^{\mathrm{B}}$-predictable nonnegative process and ξ a unit exponential random variable independent of $\mathcal{G}_\infty$. 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 $\mathcal{G}_t$, is given by

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

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

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

This construction is equivalent to the so called conditionally Poisson setup, which means that τ, conditionally on $\mathcal{G}_\infty$ and under the measure Q, is the first jump time of a Poisson inhomogeneous process with intensity $(\mu_t)_{t\geq 0}$. This setup 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, although in principle partial withdrawals from the account value may be admitted, 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.

b. 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 instantaneous evolution of the account value while the contract is still in force can be formally described as follows:

$$\frac{\mathrm{d}A_t}{A_t} = \frac{\mathrm{d}S_t}{S_t} - \varphi 1_{\{A_t<\beta\}}\mathrm{d}t, \tag{5}$$

with $A_0 = P$ and, again, 1_C denotes the indicator of the event C. Then, the return on the account value is that of the reference fund, adjusted for fees that are applied, according to the fixed rate φ, only when A_t is below the barrier β.

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

$$V_t = E\left[b^{\mathrm{D}}_\tau \mathrm{e}^{-\int_t^\tau r_v \mathrm{d}v} 1_{\{\tau\leq T\}} + b^{\mathrm{A}}_T \mathrm{e}^{-\int_t^T r_v \mathrm{d}v} 1_{\{\tau>T\}} \middle| \mathcal{F}_t \right]. \tag{6}$$

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

$$V_t = E\left[\int_t^T b^{\mathrm{D}}_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y + b^{\mathrm{A}}_T \mathrm{e}^{-\int_t^T (r_v+\mu_v)\mathrm{d}v} \middle| \mathcal{G}_t \right]. \tag{7}$$

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

Unlike equation (6), note that in (7) there are no survival indicators, and discounting is made with the mortality-adjusted discount rate $r + \mu$. Moreover, the expectation in (7) 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, e.g., 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 (6) or (7).

c. 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 payoff. Moreover, the instantaneous evolution of the account value is still described by equation (5).

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, $\lambda \geq \tau \wedge T$ 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 $\{\tau > t,\, \lambda \geq t\}$, can be expressed as

$$\begin{aligned} V_t(\lambda) = E\Big[& b^{\mathrm{D}}_{\tau} \mathrm{e}^{-\int_t^{\tau} r_v \mathrm{d}v} 1_{\{\tau \leq T \wedge \lambda\}} + \\ & + b^{\mathrm{A}}_{T} \mathrm{e}^{-\int_t^{T} r_v \mathrm{d}v} 1_{\{\tau > T,\, \lambda \geq T\}} + \\ & + b^{\mathrm{S}}_{\lambda} \mathrm{e}^{-\int_t^{\lambda} r_v \mathrm{d}v} 1_{\{\lambda < \tau \wedge T\}} \Big| \mathcal{F}_t \Big]. \end{aligned} \tag{8}$$

Alternatively, exploiting our previous assumptions, we have also[4]

$$\begin{aligned} V_t(\lambda) = E\Bigg[& \int_t^T b^{\mathrm{D}}_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y 1_{\{\lambda \geq y\}} \mathrm{d}y + \\ & + b^{\mathrm{A}}_T \mathrm{e}^{-\int_t^T (r_v+\mu_v)\mathrm{d}v} 1_{\{\lambda \geq T\}} + \\ & + b^{\mathrm{S}}_{\lambda} \mathrm{e}^{-\int_t^{\lambda} (r_v+\mu_v)\mathrm{d}v} 1_{\{\lambda < T\}} \Bigg| \mathcal{G}_t \Bigg] \\ = E\Bigg[& \left(\int_t^{\lambda} b^{\mathrm{D}}_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y + b^{\mathrm{S}}_{\lambda} \mathrm{e}^{-\int_t^{\lambda} (r_v+\mu_v)\mathrm{d}v} \right) 1_{\{\lambda < T\}} + \\ & + \left(\int_t^{T} b^{\mathrm{D}}_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v} \mu_y \mathrm{d}y + b^{\mathrm{A}}_{T} \mathrm{e}^{-\int_t^{T} (r_v+\mu_v)\mathrm{d}v} \right) 1_{\{\lambda \geq T\}} \Bigg| \mathcal{G}_t \Bigg], \end{aligned} \tag{9}$$

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

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:

$$V_t = \sup_{\lambda \in \mathcal{T}_t} V_t(\lambda), \tag{10}$$

[4] See Bacinello et al. (2010).

where $\mathcal{T}_t$ is the set of stopping times taking values in $[t, +\infty)$, with respect to the filtration $\mathbb{F}$ if $V_t(\lambda)$ is expressed by (8) or to the subfiltration $\mathbb{G}$ if instead $V_t(\lambda)$ is given by (9).

Note that the contract value V_t can also be expressed as

$$V_t = \max\left\{V_t^{\mathrm{c}},\, b_t^{\mathrm{S}}\right\}, \tag{11}$$

where V_t^{c} denotes the continuation value, given by

$$V_t^{\mathrm{c}} = \sup_{\lambda \in \mathcal{T}_t^{\mathrm{c}}} V_t(\lambda), \tag{12}$$

where $\mathcal{T}_t^{\mathrm{c}}$ is now the set of stopping times taking values in $(t, +\infty)$, again with respect to the filtration $\mathbb{F}$ if $V_t(\lambda)$ is expressed by (8) or to the subfiltration $\mathbb{G}$ if instead $V_t(\lambda)$ is given by (9).

Under the assumptions described in Sections 2 and 3a 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_u is weakly decreasing for any $u \geq t$, then $V_t^{\mathrm{c}} \geq b_t^{\mathrm{S}}$, which implies $V_t = V_t^{\mathrm{c}}$. In particular, if p_u is not constant for any $u \geq t$, or $p_t > 0$, then $V_t^{\mathrm{c}} > b_t^{\mathrm{S}}$.

Remark: Our Theorem 1 is essentially the same as Proposition 2 in MacKay et al. (2017), but it holds in a more general framework and not only when the assets price S follows a geometric Brownian motion and the mortality intensity μ is a deterministic function. 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 exactly 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 A.

The optimal stopping problem (10) 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 this 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,[5] we believe that their intrinsic flexibility, making them practically model-independent, constitutes a very precious plus. Then our aim is to test their application to the solution of our problem, trying to improve them, if necessary. That is what we are going to do in Section 4.

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).

[5]They can be very low, 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.

d. Fair contracts

Note that the initial contract value V_0 given by equations (6), (7), in the static approach, and by equation (10), in the mixed one, turns out to be a function of the fee rate φ on which the account value, and hence the benefits, depend. Then the contract is fairly priced if and only if V_0 coincides with the initial premium P:

$$V_0 \doteq V_0(\varphi) = P. \tag{13}$$

Therefore a fair fee rate, φ^*, is implicitly defined as a solution of equation (13). 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

a. 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 $P = 100$. Let $m(t) = c_1^{-c_2}(x+t)^{c_2-1}$ (with $c_1 > 0$ and $c_2 > 1$) denote the (deterministic) Weibull force of mortality. We fit m to the survival probabilities of a person (male or female) aged 50 years in 2015 implied by the statistics on the population resident in Italy (table ISTAT), getting $c_1 = 88.14778$ and $c_2 = 10.00200$ (with an expected residual lifetime of nearly 34 years).

In order to compare the results obtained under different model assumptions we define the following five models, characterized by an increasing level of complexity.

- **Model 1**:

$$\begin{cases} \mathrm{d}\ln S_t = (r - \sigma_S^2/2)\,\mathrm{d}t + \sigma_S \mathrm{d}W_t^S \\ \mu_t = m(t) \end{cases}, \tag{14}$$

with $S_0 > 0$ given, W^S a standard Brownian motion, $r = 0.03$ and $\sigma_S = 0.2$.

Hence, in this case, we put ourselves in the framework of Black and Scholes (1973), where S follows a geometric Brownian motion, and assume a deterministic mortality intensity.

- **Model 2**:

$$\begin{cases} \mathrm{d}r_t = \alpha_r(\theta_r - r_t)\mathrm{d}t + \sigma_r\sqrt{r_t}\mathrm{d}W_t^r \\ \mathrm{d}\ln S_t = (r_t - \sigma_S^2/2)\,\mathrm{d}t + \sigma_S \mathrm{d}W_t^S \\ \mu_t = m(t) \end{cases}, \tag{15}$$

with $r_0 = \theta_r = \sigma_r = 0.03$, $\alpha_r = 0.6$, W^r a standard Brownian motion independent of W^S, and the remaining variables as in Model 1.

With respect to Model 1, Model 2 introduces stochasticity in the instantaneous market interest rate r, whose dynamics is described by the mean-reverting square-root process of Cox et al. (1985).

- **Model 3**:

$$\begin{cases} \mathrm{d}r_t = \alpha_r(\theta_r - r_t)\mathrm{d}t + \sigma_r\sqrt{r_t}\mathrm{d}W_t^r \\ \mathrm{d}K_t = \alpha_K(\theta_K - K_t)\mathrm{d}t + \sigma_K\sqrt{K_t}\mathrm{d}W_t^K \\ \mathrm{d}\ln S_t = (r_t - K_t/2)\,\mathrm{d}t + \sqrt{K_t}\mathrm{d}W_t^S \\ \mu_t = m(t) \end{cases}, \tag{16}$$

with $K_0 = \theta_K = 0.04$, $\alpha_K = 1.5$, $\sigma_K = 0.4$, W^K a standard Brownian motion independent of W^r and such that $\mathrm{Cov}(\mathrm{d}W_t^K, \mathrm{d}W_t^S) = \rho_{KS}\mathrm{d}t$ with $\rho_{KS} = -0.7$, and all the remaining variables as in Model 2.

With respect to Model 2, Model 3 allows for stochasticity also in the assets volatility $\sqrt{K}$. The variance K is again described through a mean-reverting square-root process, so that we adopt the stochastic volatility model of Heston (1993), extended with stochastic interest rates.

- **Model 4**:

$$\begin{cases} \mathrm{d}r_t = \alpha_r(\theta_r - r_t)\mathrm{d}t + \sigma_r\sqrt{r_t}\mathrm{d}W_t^r \\ \mathrm{d}K_t = \alpha_K(\theta_K - K_t)\mathrm{d}t + \sigma_K\sqrt{K_t}\mathrm{d}W_t^K \\ \mathrm{d}\ln S_t = (r_t - K_t/2)\,\mathrm{d}t + \sqrt{K_t}\mathrm{d}W_t^S \\ \mathrm{d}\mu_t = \alpha_\mu(m(t) - \mu_t)\mathrm{d}t + \sigma_\mu\sqrt{\mu_t}\mathrm{d}W_t^\mu \end{cases}, \tag{17}$$

with $\mu_0 = m(0)$, $\alpha_\mu = 0.5$, $\sigma_\mu = 0.03$, W^μ a standard Brownian motion independent of W^S, W^r, W^K, and all the remaining variables as in Model 3.

With respect to Model 3, Model 4 introduces stochasticity in the mortality intensity, that is modelled through a mean-reverting square-root process with long-term mean given by the deterministic force of mortality m. As already stated in Section 3a, the process μ is independent of the other, financial related, processes S, r and K.

- **Model 5**:

$$\begin{cases} \mathrm{d}r_t = \alpha_r(\theta_r - r_t)\mathrm{d}t + \sigma_r\sqrt{r_t}\mathrm{d}W_t^r \\ \mathrm{d}K_t = \alpha_K(\theta_K - K_t)\mathrm{d}t + \sigma_K\sqrt{K_t}\mathrm{d}W_t^K \\ \mathrm{d}\ln S_t = (r_t - K_t/2 - \lambda_S\gamma_S)\,\mathrm{d}t + \sqrt{K_t}\mathrm{d}W_t^S + \mathrm{d}J_t^S \\ \mathrm{d}\mu_t = \alpha_\mu(m(t) - \mu_t)\mathrm{d}t + \sigma_\mu\sqrt{\mu_t}\mathrm{d}W_t^\mu + \mathrm{d}J_t^\mu \end{cases}, \tag{18}$$

where J^S and J^μ are two independent compound Poisson processes, independent of W^S, W^r, W^K, W^μ. The jump arrival rate of J^S is $\lambda_S = 0.5$ and its i.i.d. jump sizes Δ_S are lognormally distributed, such that $\ln(1 + \Delta_S) \sim N(\gamma_S, \eta_S^2)$, with $\gamma_S = 0$ and $\eta_S = 0.07$. The jump arrival rate of J^μ is $\lambda_\mu = 0.1$ and its i.i.d. jump sizes are exponentially distributed with mean $\gamma_\mu = 0.01$. All the remaining variables are as in Model 4.

While all stochastic processes in Models 1-4 are pure diffusions, hence with (a.s.) continuous paths, Model 5 introduces a jump component both to the assets price process S and to the mortality intensity μ.

b. 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. We compute the contract value through equation (6), so that we need also to simulate the time of death, through (4).

In Figures 1-5 we report the results obtained in the different five models for various roll-up rates δ and fee rates φ (in percentage), by assuming a state-dependent fee structure with barrier equal to the guaranteed amount at maturity, $\beta = Pe^{\delta T}$. 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, in particular, the fair fee rate φ^* is given by the abscissa of the intersection between the contract graph and the dotted horizontal line at level $P = 100$. E.g., $\varphi^* \simeq 5\%$ when $\delta = 0$ in Models 1 and 3, when $\delta = 1\%$ in Model 2 and when $\delta = 2\%$ in Model 4, while $\varphi^* \simeq 4\%$ when $\delta = 0$ in Model 5. It is not a priori clear, instead, which is the effect of the roll-up rate on the contract value. On the one hand, the higher is δ, the higher is the guaranteed amount at death or maturity, but, on the other hand, the higher is also 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 model as well as on the level of δ and φ. For instance, in Models 1 and 2 the highest contract value is achieved when $\delta = 2\%$ for any level of the fee, while the lowest value turns out to be the one in which $\delta = 1\%$ for high levels of the fee (over φ^*), and that with $\delta = 0$ for (relatively) low levels of φ. In Models 3 and 4 the lowest value is achieved when $\delta = 0$ for very low levels of the fee and when $\delta = 1\%$ otherwise, but instead the highest value turns out to be that with a return-of premium guarantee ($\delta = 0$) for (relatively) high levels of the fee (and remains that with $\delta = 2\%$ for relatively low or excessively high levels of φ). Finally, in Model 5 the highest contract value is also achieved when $\delta = 0$ but the lowest one is obtained with $\delta = 2\%$ (except for extremely low levels of the fee, when this situation is completely reversed, or for extremely high levels of φ, when the lowest value becomes again that with $\delta = 1\%$).

In order to get some insights into the model risk, in Figure 6 we plot the contract value V_0 against the (state-dependent) fee rate φ in all the models when the roll-up rate $\delta = 2\%$ and, again, the barrier $\beta = Pe^{\delta T}$. Several interesting, and somewhat unexpected, results are highlighted. First of all, the introduction of stochastic mortality (transition from Model 3 to Model 4) seems to have no effect on the contract value since the corresponding graphs are overlapping. We have almost the same result when passing from Model 1 to Model 2, i.e., when we introduce stochastic interest rates, with practically overlapping graphs and a slight predominance of the contract value in Model 1. When instead we pass from Model 2 to Model 3, i.e., introduce stochastic volatility, the contract value jumps downward, especially for high values of the fee, resulting in an about 170 b.p. lower fair fee rate. The downward jump becomes huge when we pass to Model 5, i.e., introduce jumps to the assets price and to the mortality intensity, with an halved fair fee rate (with respect to Model 3).

The results of this comparison would push us to conclude that more uncertainty there is, the lower the contract value is, but of course they are partly due also to the model parameters. In particular, the uncertainty in the assets price becomes higher with a higher volatility, and this would increase the optional components of the contract. But what can we say if the volatility is stochastic? Model 3 (and following) is characterized by a highly negative correlation between the assets price process S and its stochastic variance K, supported by the empirical evidence according to which slumps in the financial markets are often accompanied by an increase in their volatility, but the comparison between Models 2 and 3 is completely reversed if we fix instead a positive correlation coefficient, say $\rho_{KS} = 0.7$ (see Figure 7, where Model 3' refers to Model 3 adjusted with this new value of ρ_{KS}).

We have also tried to change the order in which the various types of stochastic factors are introduced. E.g., in Figure 8 we report the contract values in Models 1 and 3, along with those in an adjusted Model 2, say Model 2', in which stochastic volatility, instead of stochastic interest rates, is added to Model 1 (with, again, $\rho_{KS} = -0.7$). We can see that there is a substantial difference among the results obtained in the three models: the introduction of stochastic volatility to Model 1 lowers considerably the contract value, but the additional introduction of stochastic interest rates increases it, especially for not excessively high levels of the fee. Anyway, the contract value remains the highest in Model 1 for any level of the fee.

Similarly, we add stochastic mortality to Model 1, letting interest rates and volatility constant. This new model is referred to as Model 2". As can be seen from Figure 9, the introduction of stochastic mortality very slightly lowers the contract value, although there are no substantial differences between the results obtained in Models 1 and 2".

To verify if the above patterns are specific of the particular structure of the fees, in Figure 10 we report the contract values in all the original Models 1-5 when the fees are constant instead of state-dependent. The pattern remains very similar to that observed in Figure 6: no difference between Models 1 and 2 on one hand, and between Models 3 and 4, on the other hand; slightly lower contract values in Models 3 and 4, with respect to Models 1 and 2, and practically coinciding, in all Models 1-4, for extremely high values of the fee; huge decline in the contract value when passing to Model 5, with a gap increasing with respect to the fee rate. From Figures 6 and 10 it is also visible that the fair fee rate with constant fees is much lower than that obtained in the case of state-dependent fees in all models. 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.

To compare the contract values with state-dependent and constant fee structures, with the same fee rate φ, in Figure 11 we plot them against φ, again with a roll-up rate $\delta = 2\%$. Since, in the previous discussion, we have already devoted a lot of time to the model risk, we restrict now ourselves to the results obtained in Model 4. We can see that the fair fee rate φ^* required by a state-dependent fee structure is practically doubled with respect to that required by constant fees. Such a rate is hardly marketable. Then, if we admit that a fee is paid also when the guarantee is moderately out-of-the money, we can try to increase the occupational time of deducting fees by balancing the barrier β. In particular, we take now $\beta = (1 + k)Pe^{\delta T}$ and plot, in Figure 12, the fair fee rate φ^* against the barrier increment k, once again in Model 4, with a roll-up rate $\delta = 2\%$. 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).

c. Results under the mixed approach

We now present some results under the mixed approach. To obtain them we use the Least Squares Monte Carlo technique, and compute the contract value through equation (8). The number of simulations is again 20000 and, for the regression, we use Laguerre polynomials up to order 3. A straightforward application of these techniques is a bit problematic for relatively low levels of the fee (usually not beyond 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 too 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. We have observed this behaviour 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 the attempt to improve the regression, we have tried 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), which however have not brought substantial enhancements. In contrast, Theorem 1 has proven to be extremely helpful to this end. From the computational point of view, it allows us to skip the regression step in the LSMC algorithm when $A_t \geq \beta$.[6]

In Figures 13-17 we report the results obtained in Models 1-5 for various fee rates φ (in percentage), by assuming a state-dependent fee structure with barrier equal to the guaranteed amount at maturity, roll-up rate $\delta = 0.02$ and constant surrender penalties $p_t = 0.02$ for any $t < T$. 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 just described arrangement based on Theorem 1, we display together the contract value under the static approach and under the mixed one before and after the adjustment. As already observed, the contract value under the mixed approach before the adjustment (green line) is substantially lower than that under the static approach (blue line) for low levels of the fee. If instead we introduce the adjustment, for very low levels of the fee the contract value under the mixed approach (red line) practically coincides with that under the static approach, then becomes slightly

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

lower, until the intersection point between red and blue lines. After this point the contract value under the static approach becomes increasingly lower than that under the mixed approach. Then there is a substantial improvement when passing from the green to the red line for low levels of the fee, while for high levels the two graphs completely overlap. It is interesting to notice that all the three graphs have a common intersection in Models 1-4. In particular, in Models 1, 3 and 4 the intersection point coincides with the fair fee rate, equal to around 7% in Model 1 and 5% in Models 3 and 4 (under both static and mixed approaches). This implies that, in these models, the surrender option becomes valuable as soon as the fee rate applied to the contract is higher than its fair level. In Model 2 the (common) intersection point is lower than the corresponding fair fee rates, equal to around 7% under the mixed approach and 6.7% under the static one. In Model 5, instead, each pair of graphs has a different intersection point, although very close to each other: the first is between red and blue lines, the second between green and blue lines, the last between red and green lines. This is the only example in which the fair fee rate under the mixed approach (around 2.7% after the adjustment, see red line) is slightly below that under the static one (around 2.8%, blue line). Summing up: the theoretical relation between the fair fee rates under mixed and static approaches is almost always preserved and, even if the adjusted algorithm still produces some numerical error before the intersection point, the red line is very close to the blue one, so that we can argue that in these cases the surrender incentive has been completely eliminated.

From now on we only show results under the mixed approach, and use the adjusted algorithm in order to produce them. To capture model risk, in Figure 18 we plot the contract value V_0 against the fee rate φ in all models, with the same contract parameters assumed for the previous Figures 13-17. Comparing Figures 6 (static approach) and 18 (mixed approach), we notice that there is the same relation among models. In particular, as just observed, the graphics in the two figures coincide until the intersection point, usually equal to the fair fee rate, and after differ, for (relatively) high levels of the fee. However, it is interesting to point out that, differently from the static approach, when the fee rate becomes higher the difference among models tends to disappear: with a fee rate of 13% we have the same contract value under the first four models (around 97), and a bit lower value under Model 5 (around 95), letting us to conjecture that, in the limit, also this difference tends to vanish. Then the presence of the surrender option allows to 'kill' the adverse consequences arising from an unfavourable model or contract design.

With constant fees we also observe this convergence among models for very high values of the fee rate (see the following Figure 19), but Models 1-4 were much closer to each other and the convergence between them was faster under the static approach (Figure 10), due to the fact that a constant fee structure is much more penalizing than a state-dependent one, implying, on one hand, a lower fair fee rate with respect to the state-dependent case, and, on the other hand, an amplified difference, still in terms of fair fee rate, between static and mixed approach (often null with state-dependent fees). Recall, in fact, that one of the objectives of state-dependent fees is to eliminate (or, at least, reduce) the surrender incentive.

Coming now to another aspect of the contract design, given by the surrender penalty, in the following Figure 20 we display the initial contract value versus the (state-dependent) fee rate for three different penalty structures: constant penalty $p_t = 2\%$, decreasing penalty $p_t = 0.08(1-t/T)^3$, exponentially decreasing penalty $p_t = 1-\mathrm{e}^{-0.08(1-t/T)}$, all for any $t < T$. The last two surrender penalty structures are (special cases of) those employed by MacKay et al. (2017). In particular, they both tend to 0 as $t \to T$; the 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 exponentially decreasing penalty starts from a level of around 7.7% when $t = 0$ and then decreases more slowly, attaining the level of the constant fee soon after the start of the twelfth contract year. We limit ourselves to present the results for Model 4, with a roll-up rate of 2%. 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), while the other two penalty structures lead to the same results, with a slightly lower contract value in the exponentially decreasing case. The gap between the contract value under constant and decreasing penalties becomes higher as the fee rate increases. When surrender is convenient (high levels of the fee), it is then better to have constant rather than decreasing penalties, thus suggesting that the convenience to exit the contract is soon enough, before the crossing between constant and decreasing penalties, i.e., while the constant fee is still the lowest.

As already observed in the examples of the static approach, the fair fee rate implied by a state-dependent fee structure can be significantly higher than that implied by constant fees, leading to the design of unmarketable contracts. To increase marketability, we assume now to fix a target fee level φ and search for a barrier increment k in order that the barrier level $\beta = (1+k)P\mathrm{e}^{\delta T}$ allows to achieve fairness. Taking into account that the contract value is decreasing with k, and converges, as $k \to +\infty$, to that obtained in the case of constant fees, we deduce that our room for maneuver lies between the fair fee rate with $k = 0$ and that with constant fees. With a roll-up rate of 2% this room is rather narrow, while it becomes substantial, e.g., with a return-of-premium guarantee (see Figure 21, sticking to Model 4). Then, as a last 'exercise', we fix a target fee level of 3% and represent, in Figure 22, the contract value V_0 versus the barrier increment k obtained with this target fee rate and a return-of-premium guarantee. Two things are worth highlighting: i) the convergence towards the contract value in the case of constant fees is very fast, attained with a barrier level of about 116% of the single premium P; ii) fairness is also achieved very soon, by taking a barrier level equal to (around) 109% of P.

d. Implementation note

All simulations and graphs in this paper have been implemented in the R language. The code to specify the contract riders, simulate the paths of the financial and biometric processes, perform the static and mixed approach calculations and estimate the fair fee rate has been converted in an R package. We have tested the code for accuracy by reproducing the results found in Bacinello et al. (2011). The package, named valuer, has been released on CRAN `https://CRAN.R-project.org/package=valuer` under the GPL-3 license and the version used for this article is 1.1.1.

5. Summary and conclusions

In this paper we have proposed a quite general valuation model for variable annuities providing guarantees at death and maturity and financed through the application of a state-dependent fee. The interaction among fee rates, death/maturity guarantees, fee thresholds and surrender penalties under alternative model assumptions and policyholder behaviours has been extensively analysed from a numerical point of view, letting us, on one side, to capture the implications of model risk and, on the other, 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 LMSC techniques has turned out to be a bit problematic for low levels of the fee, due to the shape of the surrender region. 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 overcome this problem. As far as model risk is concerned, our conclusion is that, at least for the parameters here assumed, more sophisticated models are not necessarily more prudential for pricing purposes. However, since models calibration was beyond the scope of the paper, it would be interesting to further deepen this aspect.

APPENDIX A

Proof of Theorem 1

For convenience, we let

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

so that we can rewrite equation (9) in the following, more compact, way:

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

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

$$\varepsilon = \begin{cases} T & \text{if } A_u \geq \beta \quad \forall u \in (t,T) \\ \inf\{u \in (t,T) : A_u < \beta\} & \text{otherwise} \end{cases}.$$

Of course ε is a stopping time with respect to the filtration $\mathbb{G}$, hence $\varepsilon \in \mathcal{T}_t^{\mathrm{c}}$. Then the continuation value satisfies

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

Observe that

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

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 $y = T$ it is understood that $p_y = 0$. If $p_t > 0$, or $p_y < p_t$, then the inequality between b_y^{D} and $A_y(1-p_t)$ is strict. 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

$$V_t^{\mathrm{c}} \geq (1-p_t)E\left[\int_t^\varepsilon A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y \mathrm{d}y \,+\, A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\middle|\mathcal{G}_t\right] \doteq (1-p_t)L_t, \qquad \text{(A1)}$$

again with a strict inequality if $p_t{>}0$ or p not constant over $[t,T]$. Now we split the expectation in (A1) into the sum of two expectations, and work out them starting from the second, $E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\middle|\mathcal{G}_t\right]$, that we further condition on ε:

$$\begin{aligned}
E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\middle|\mathcal{G}_t\right] &= E\left[E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon (r_v+\mu_v)\mathrm{d}v}\middle|\varepsilon,\mathcal{G}_t\right]\middle|\mathcal{G}_t\right]\\
&= E\left[E\left[A_\varepsilon \mathrm{e}^{-\int_t^\varepsilon r_v\mathrm{d}v}\middle|\varepsilon,\mathcal{G}_t\right]E\left[\mathrm{e}^{-\int_t^\varepsilon \mu_v\mathrm{d}v}\middle|\varepsilon,\mathcal{G}_t\right]\middle|\mathcal{G}_t\right]\\
&= E\left[A_t Q(\tau>\varepsilon|\tau>t,\varepsilon,\mathcal{G}_t)\middle|\mathcal{G}_t\right]\\
&= A_t E\left[Q(\tau>\varepsilon|\tau>t,\varepsilon,\mathcal{G}_t)\middle|\mathcal{G}_t\right]. \qquad \text{(A2)}
\end{aligned}$$

The first equality in (A2) follows from the law of iterated expectations, the second from the stochastic independence between the mortality intensity μ and any financial related variable, in particular r and A, the third from the martingality of the discounted account value. Recall, in fact, that, 'before' ε the fee is not applied, hence the instantaneous return on the account value A is the same as the instantaneous return on the assets price S. The last equality is a consequence of the fact that A is a $\mathbb{G}$-adapted process, i.e., A_t is $\mathcal{G}_t$-measurable. As far as the first expectation in (A1) (after splitting) is concerned, with similar algebraic manipulations we obtain:

$$\begin{aligned}
E\left[\int_t^\varepsilon A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\mathrm{d}y\middle|\mathcal{G}_t\right] &= E\left[E\left[\int_t^\varepsilon A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\mathrm{d}y\middle|\varepsilon,\mathcal{G}_t\right]\middle|\mathcal{G}_t\right]\\
&= E\left[\int_t^\varepsilon E\left[A_y \mathrm{e}^{-\int_t^y (r_v+\mu_v)\mathrm{d}v}\mu_y\middle|\varepsilon,\mathcal{G}_t\right]\mathrm{d}y\middle|\mathcal{G}_t\right]\\
&= E\left[\int_t^\varepsilon E\left[A_y \mathrm{e}^{-\int_t^y r_v\mathrm{d}v}\middle|\varepsilon,\mathcal{G}_t\right]E\left[\mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\middle|\varepsilon,\mathcal{G}_t\right]\mathrm{d}y\middle|\mathcal{G}_t\right]\\
&= E\left[\int_t^\varepsilon A_t E\left[\mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\middle|\varepsilon,\mathcal{G}_t\right]\mathrm{d}y\middle|\mathcal{G}_t\right]\\
&= A_t E\left[E\left[\int_t^\varepsilon \mathrm{e}^{-\int_t^y \mu_v\mathrm{d}v}\mu_y\mathrm{d}y\middle|\varepsilon,\mathcal{G}_t\right]\middle|\mathcal{G}_t\right]\\
&= A_t E\left[Q(\tau\leq\varepsilon|\tau>t,\varepsilon,\mathcal{G}_t)\middle|\mathcal{G}_t\right]. \qquad \text{(A3)}
\end{aligned}$$

Finally, summing up both expectations in (A1) gives

$$V_t^{\mathrm{c}} \geq (1-p_t)L_t = (1-p_t)A_t = b_t^{\mathrm{S}}. \ \square$$

APPENDIX B

Figures

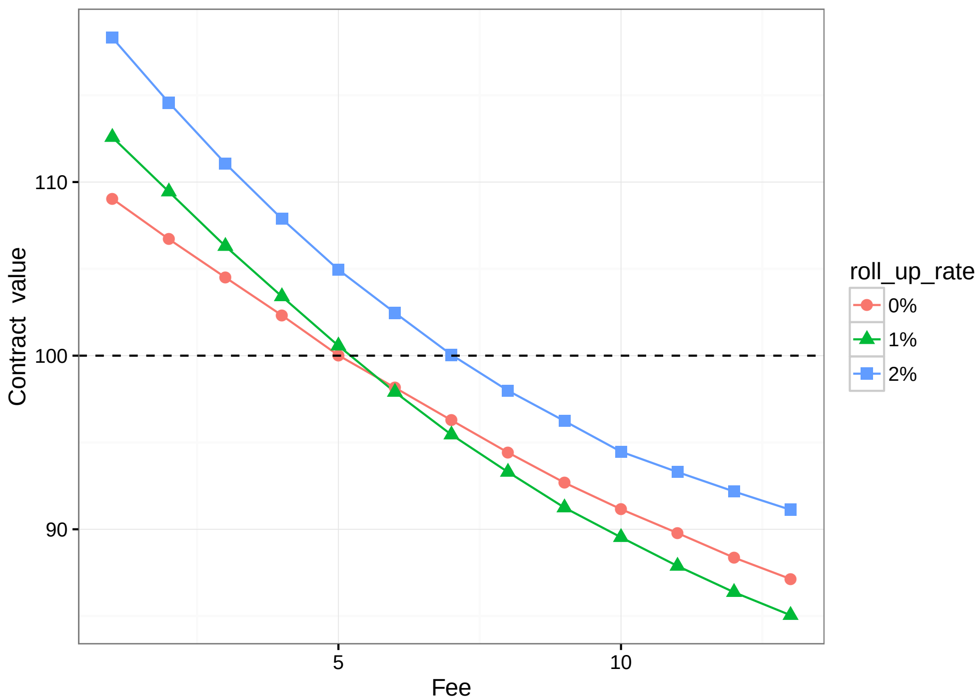

Figure 1: **Model 1**: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different roll-up rates δ; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

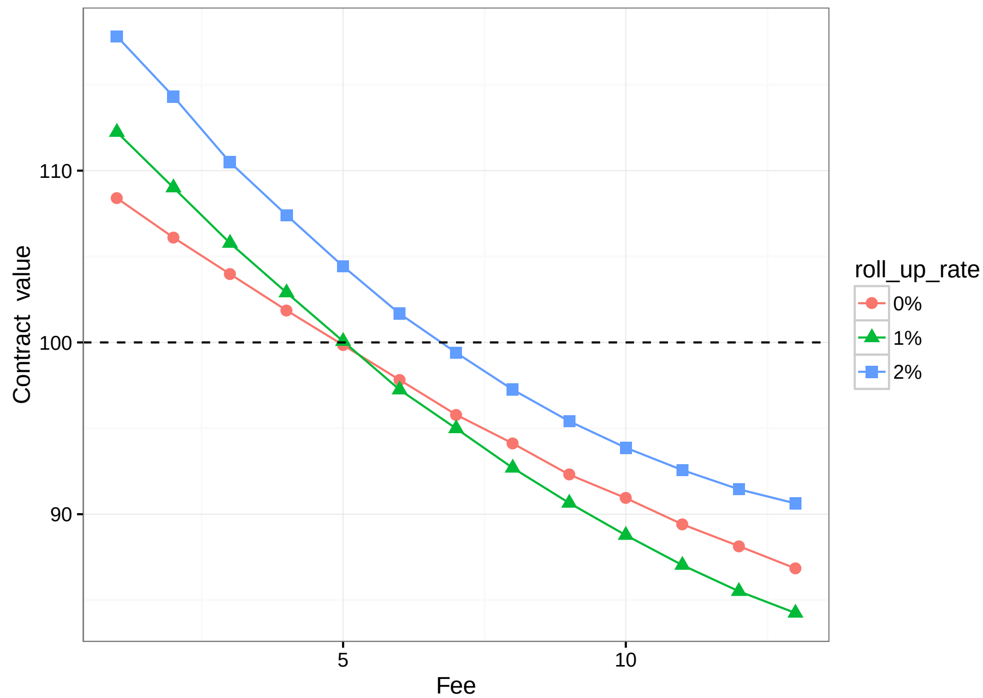

Figure 2: **Model 2**: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different roll-up rates δ; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

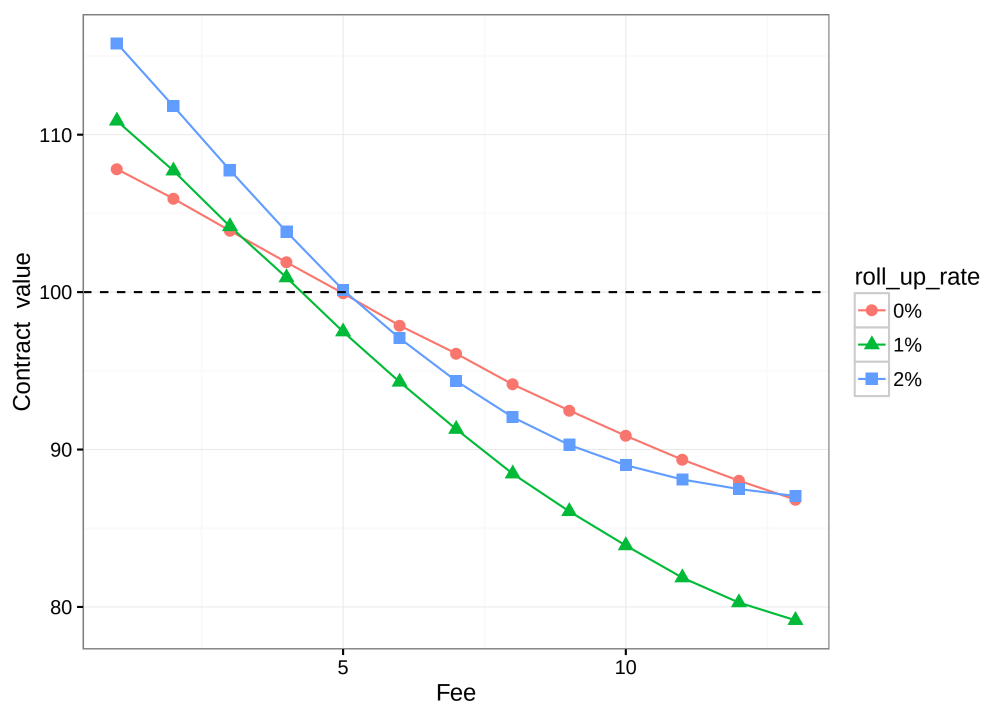

Figure 3: **Model 3**: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different roll-up rates δ; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

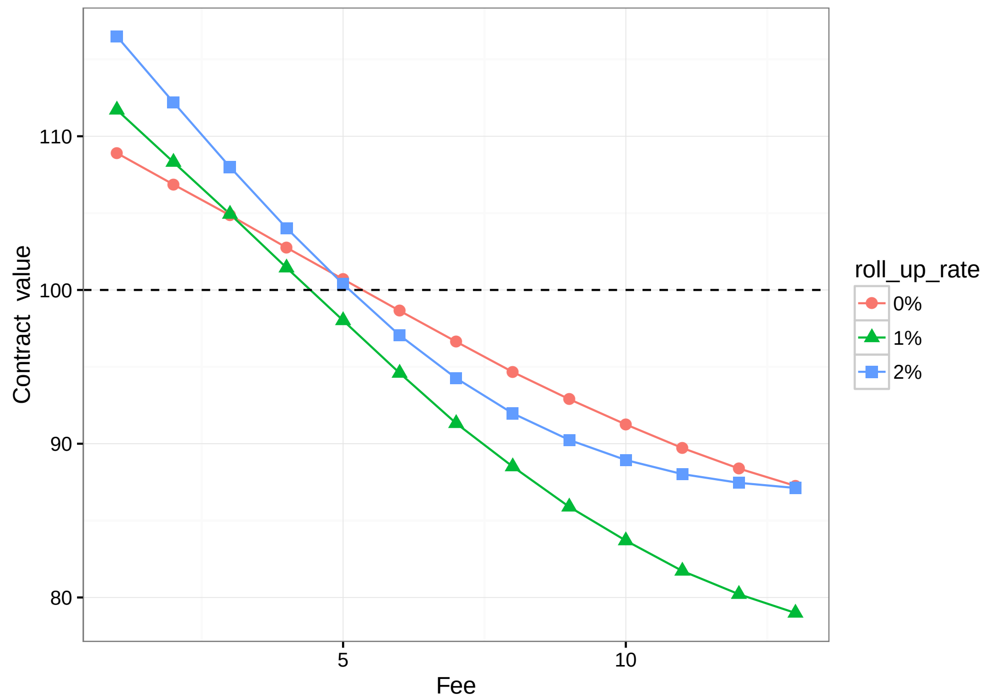

Figure 4: **Model 4**: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different roll-up rates δ; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

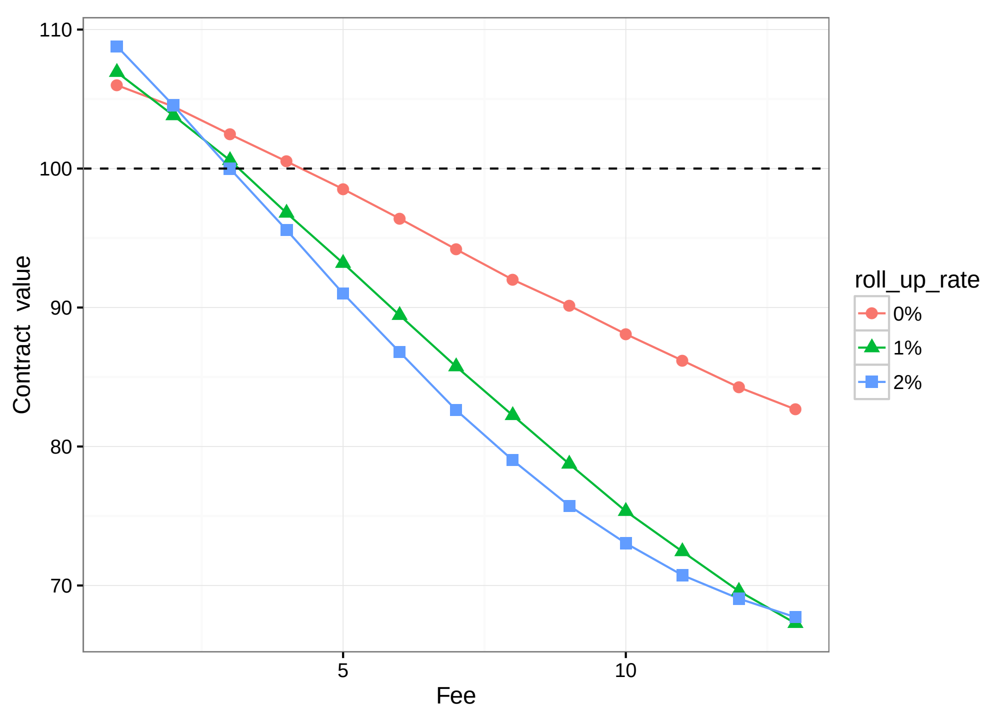

Figure 5: **Model 5**: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different roll-up rates δ; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

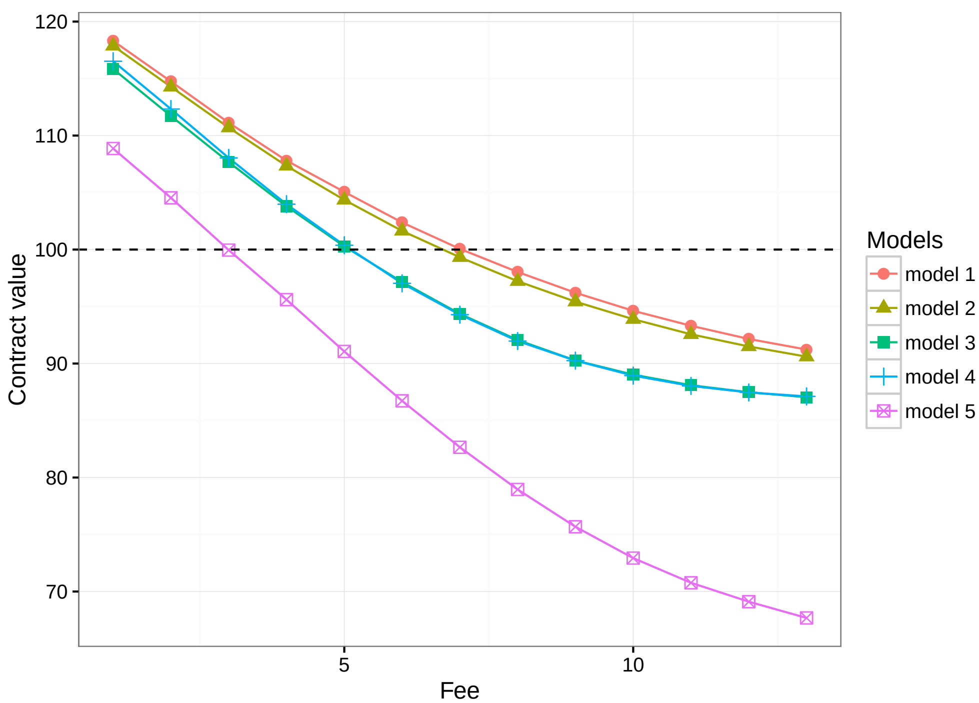

Figure 6: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

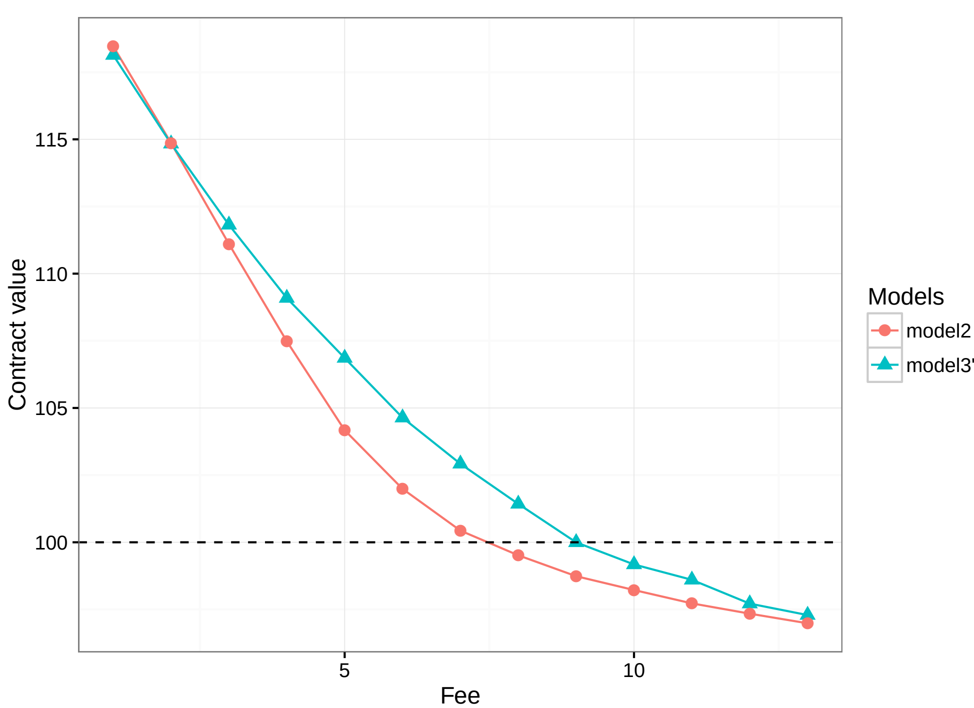

Figure 7: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

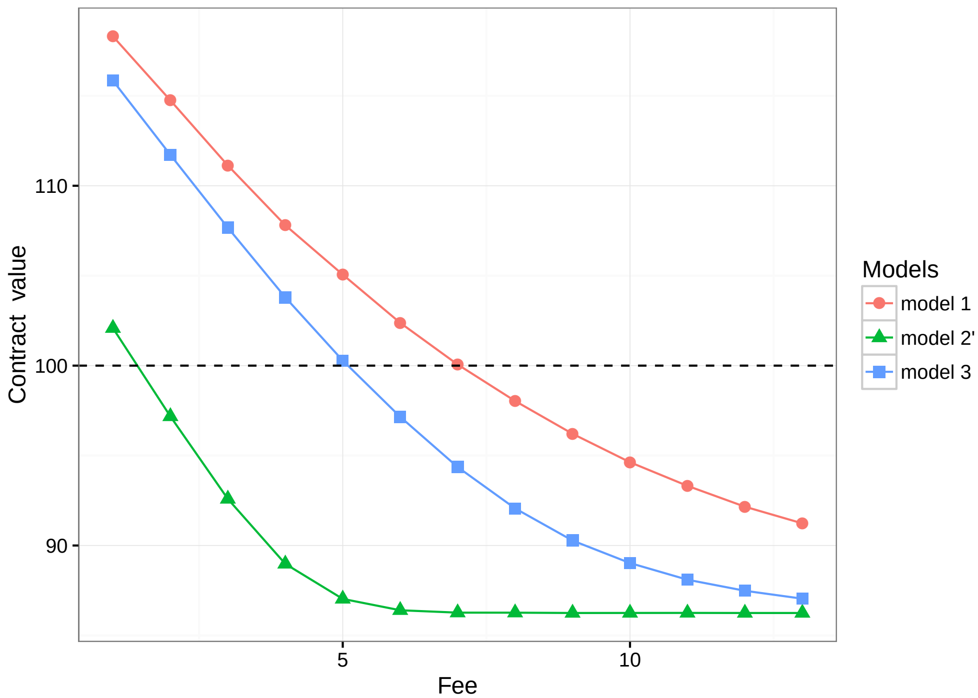

Figure 8: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

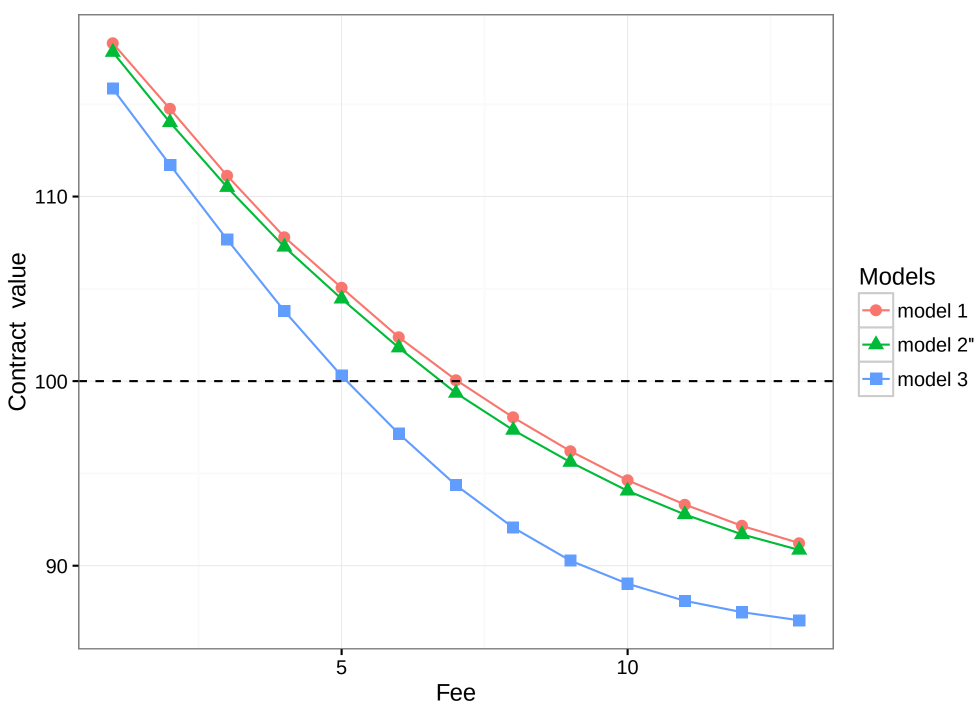

Figure 9: The initial contract value, under the **static** approach, versus the **state-dependent** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

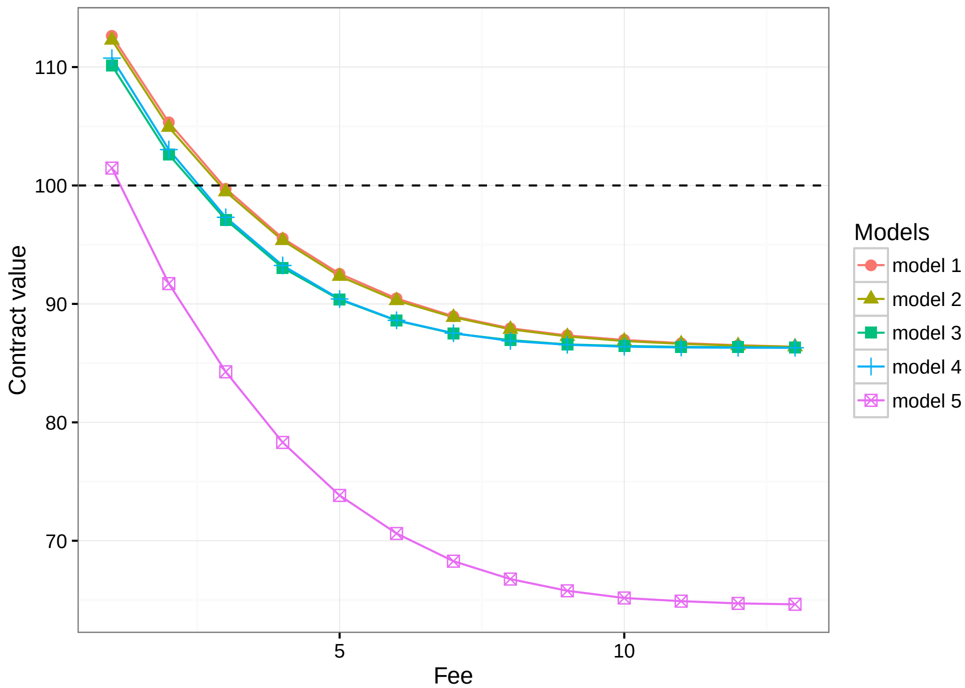

Figure 10: The initial contract value, under the **static** approach, versus the **constant** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$.

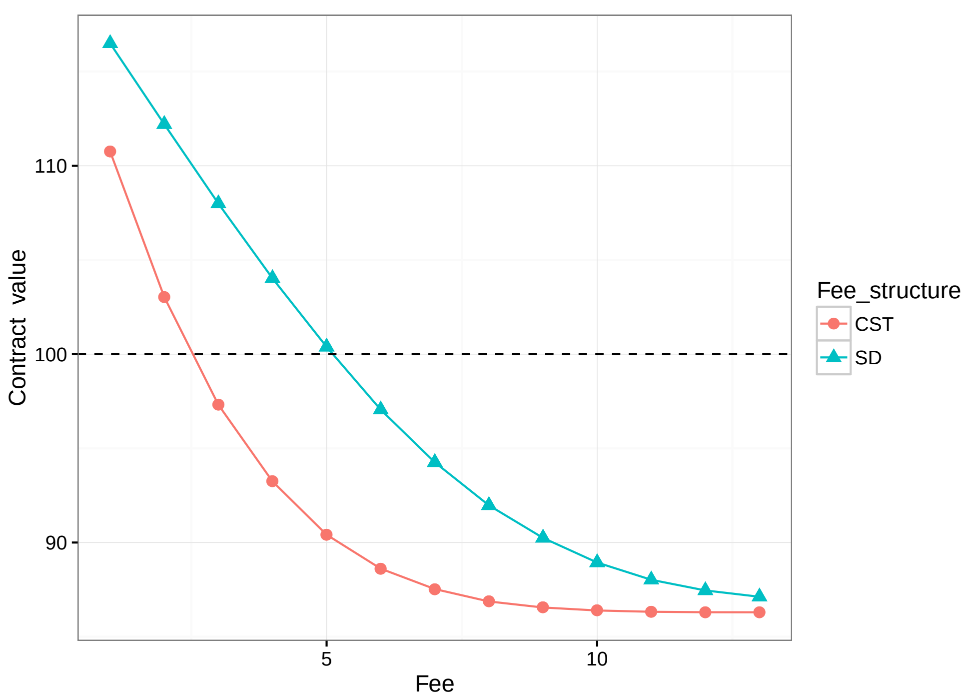

Figure 11: **Model 4**: The initial contract value, under the **static** approach, versus the fee rate, for both constant and state-dependent fees; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

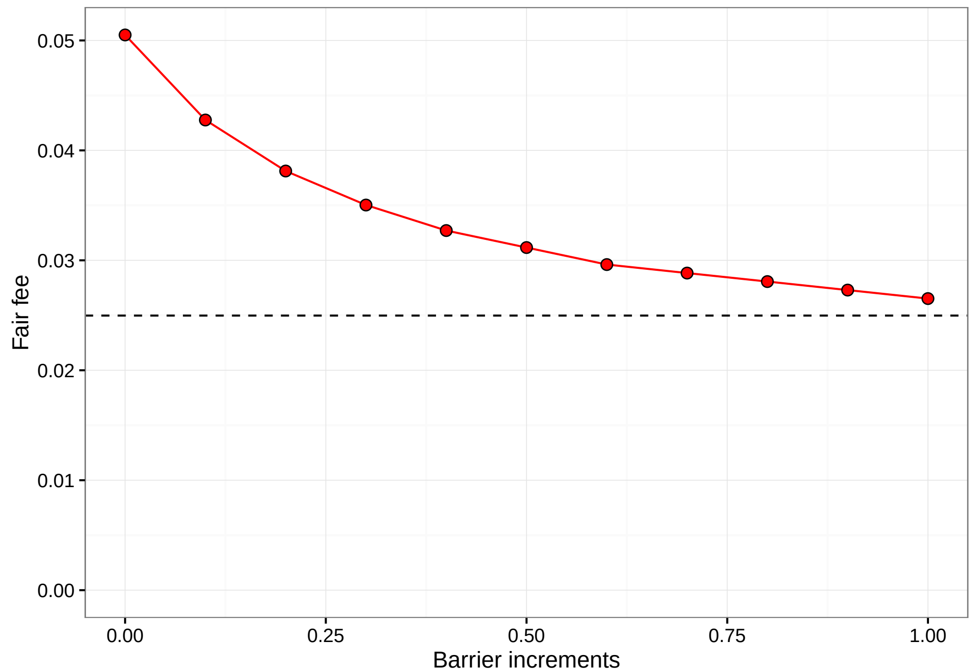

Figure 12: **Model 4**: The **state-dependent** fair fee rate φ^*, under the **static** approach, versus the barrier increment k; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = (1 + k)Pe^{\delta T}$.

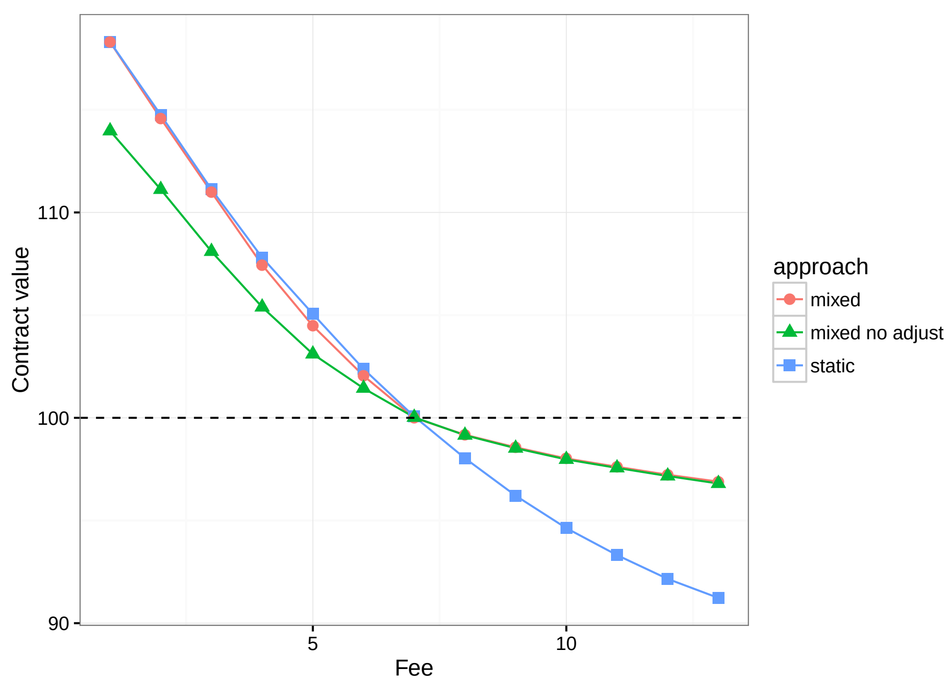

Figure 13: **Model 1**: The initial contract value, under the static and the mixed approaches, versus the **state-dependent** fee rate; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

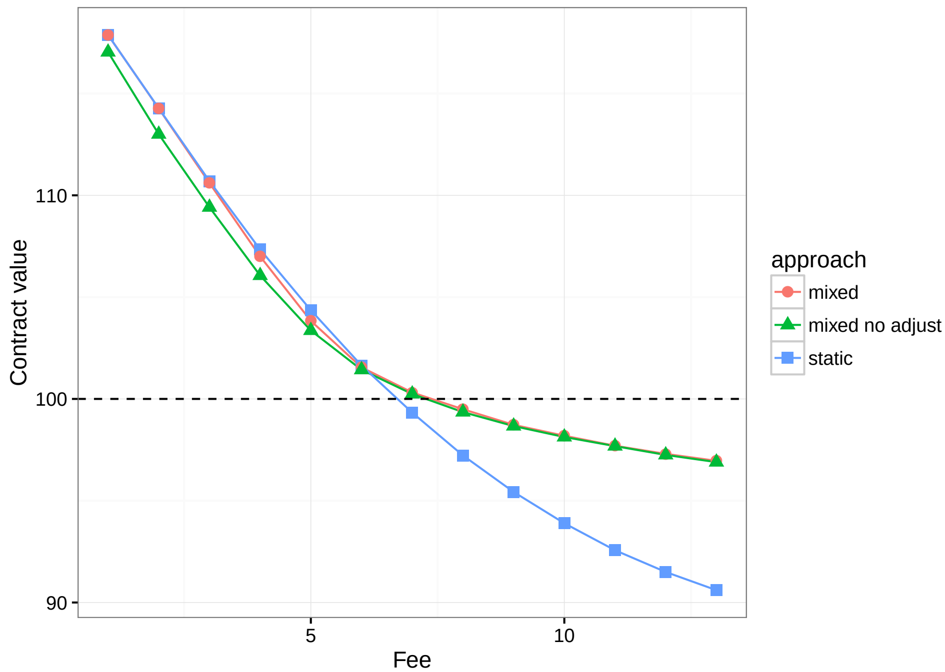

Figure 14: **Model 2**: The initial contract value, under the static and the mixed approaches, versus the **state-dependent** fee rate; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

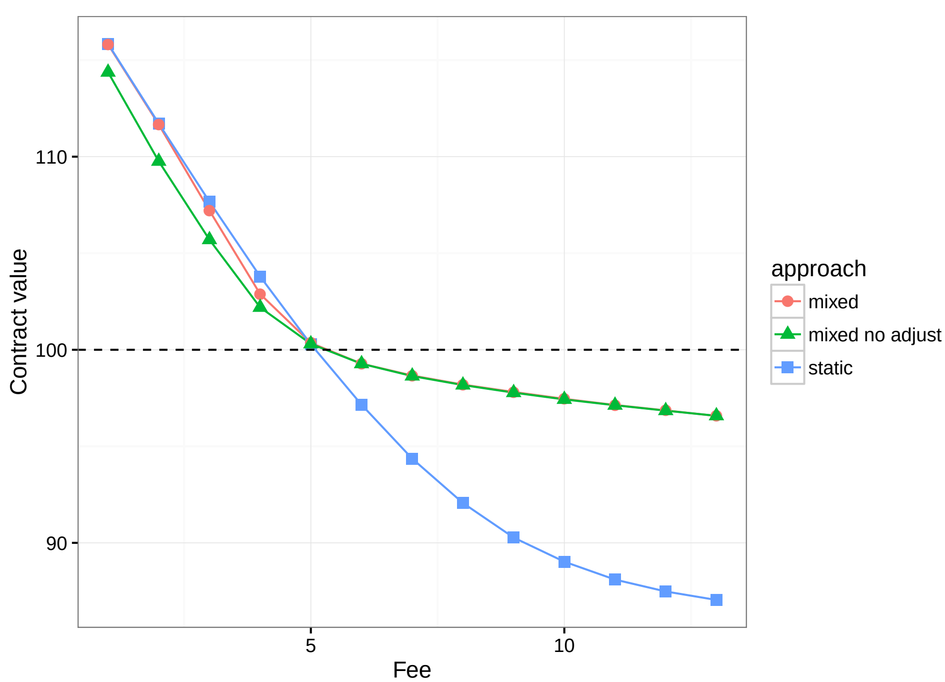

Figure 15: **Model 3**: The initial contract value, under the static and the mixed approaches, versus the **state-dependent** fee rate; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

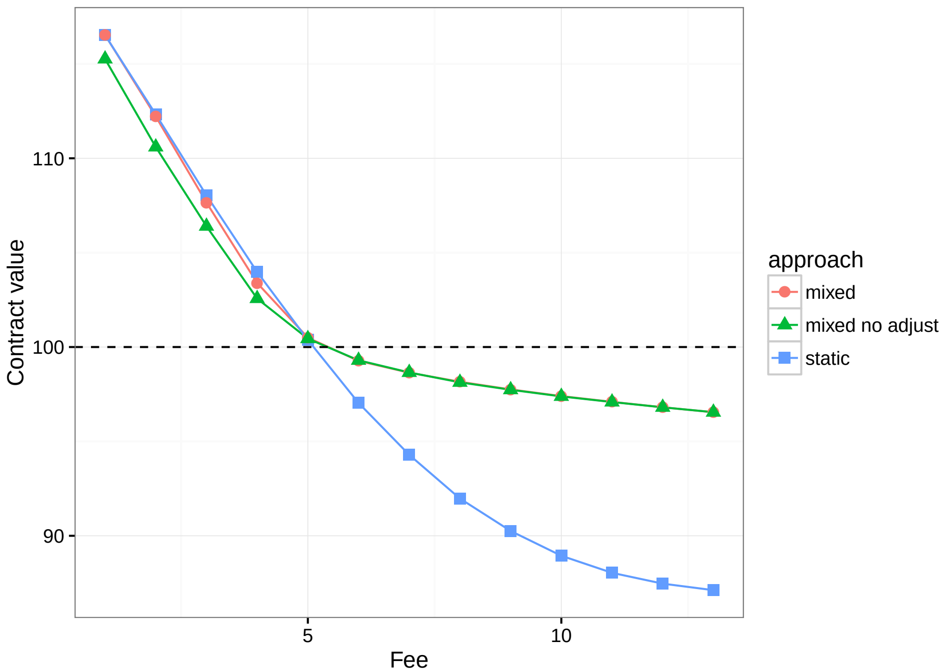

Figure 16: **Model 4**: The initial contract value, under the static and the mixed approaches, versus the **state-dependent** fee rate; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

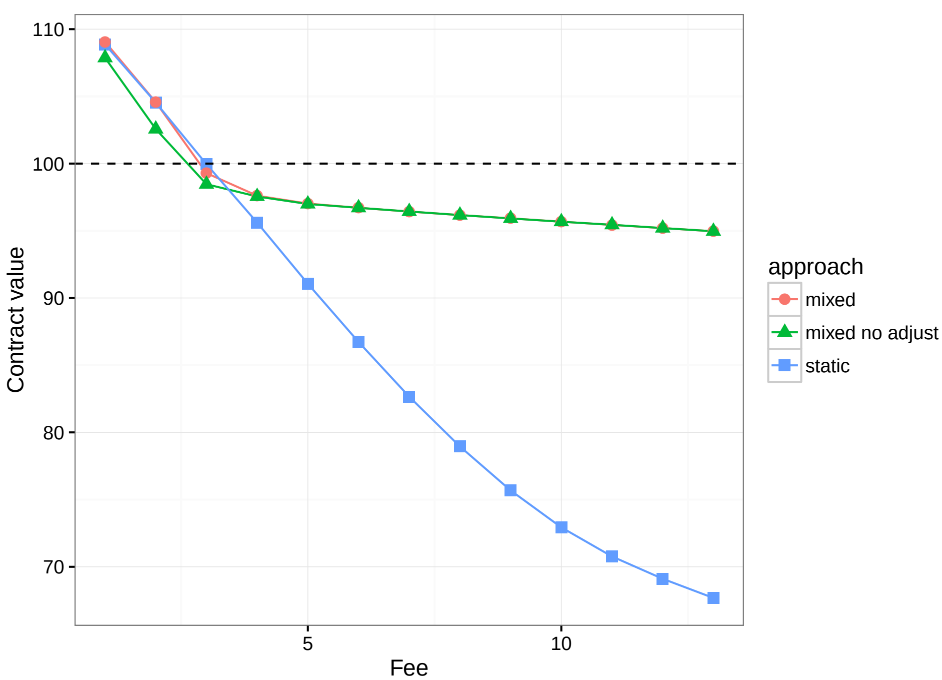

Figure 17: **Model 5**: The initial contract value, under the static and the mixed approaches, versus the **state-dependent** fee rate; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

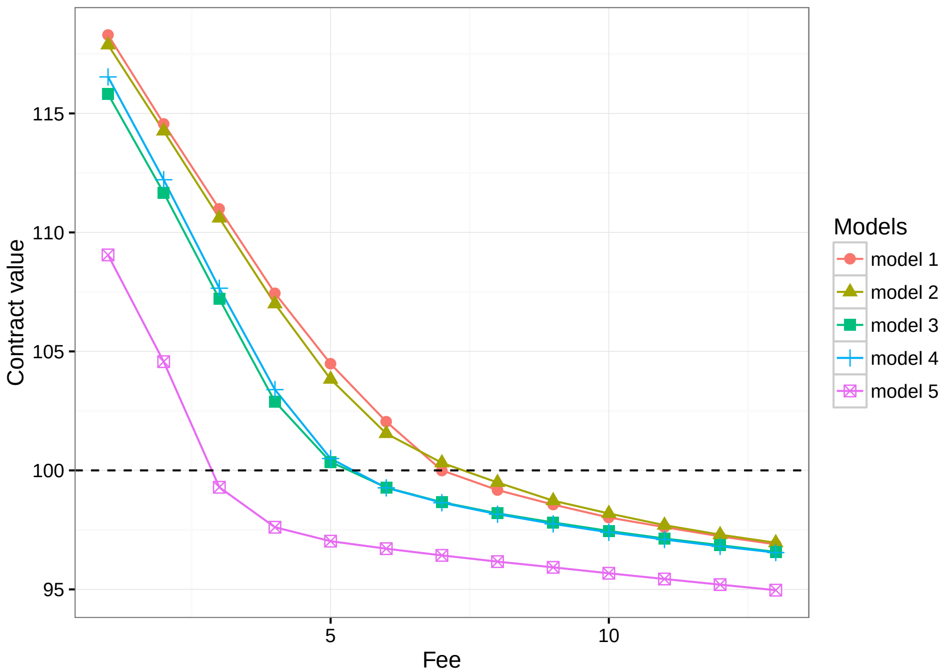

Figure 18: The initial contract value, under the **mixed** approach, versus the **state-dependent** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = P\mathrm{e}^{\delta T}$, surrender penalty $p_t = 2\%$ for any $t < T$.

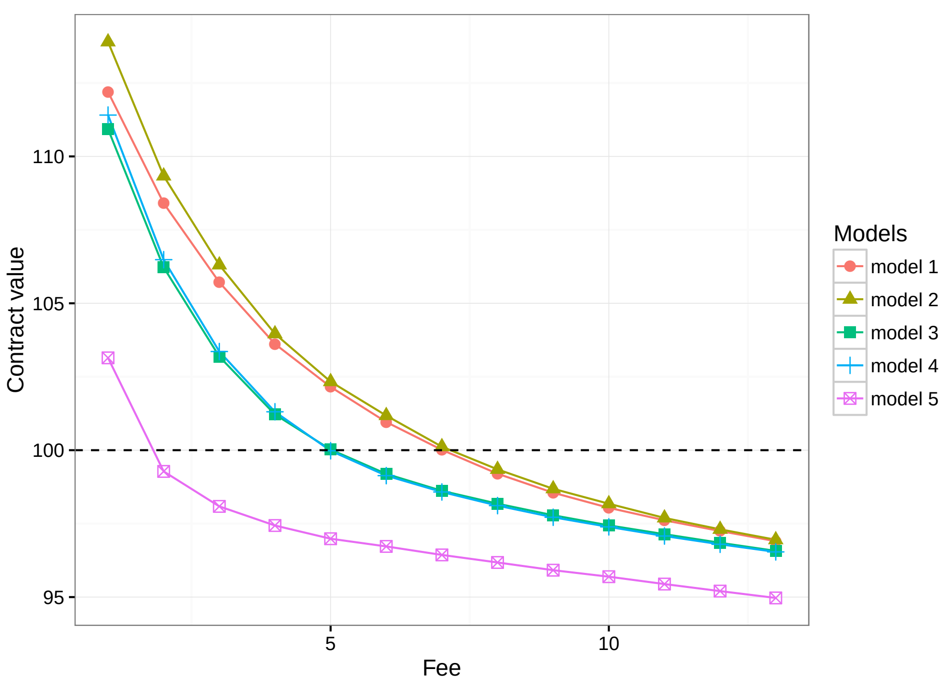

Figure 19: The initial contract value, under the **mixed** approach, versus the **constant** fee rate, for different models; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, surrender penalty $p_t = 2\%$ for any $t < T$.

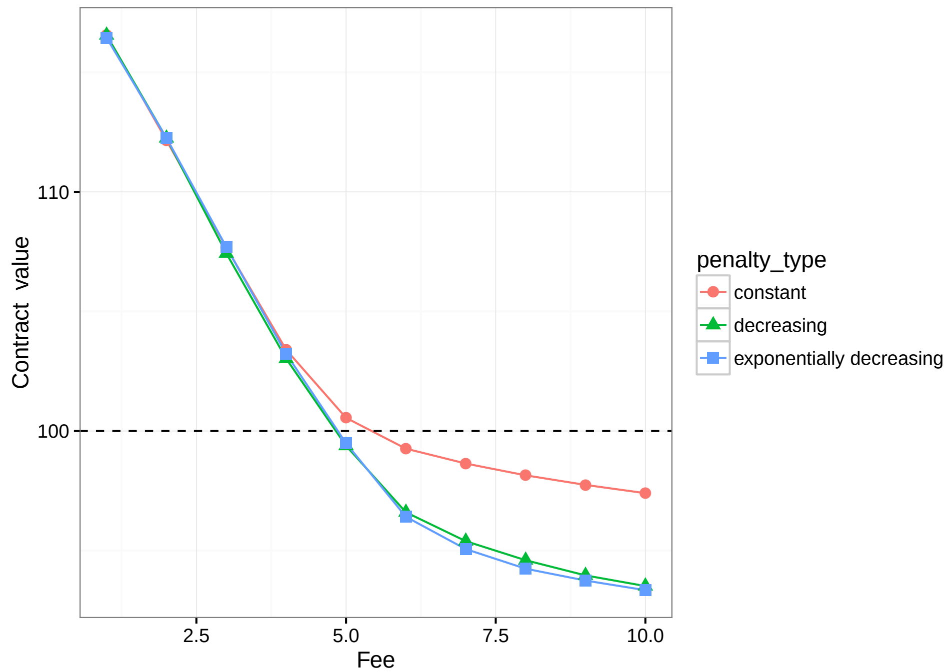

Figure 20: **Model 4**: The initial contract value, under the **mixed** approach, versus the **state-dependent** fee rate, for different penalty structures; single premium $P = 100$, maturity $T = 15$, roll-up rate $\delta = 2\%$, barrier $\beta = Pe^{\delta T}$.

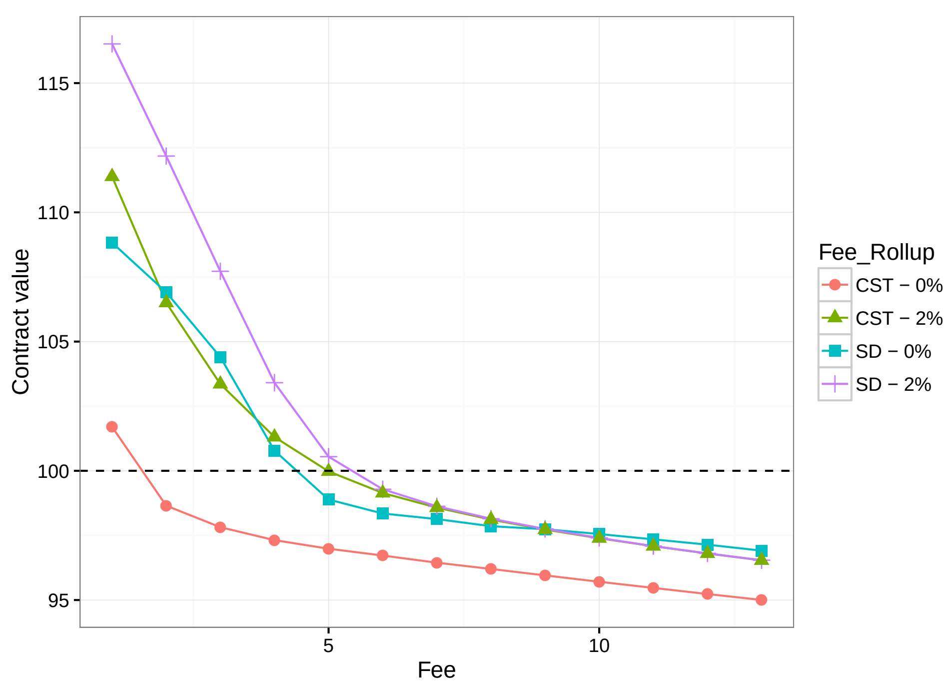

Figure 21: **Model 4**: The initial contract value, under the **mixed** approach, versus the fee rate, for both constant and state-dependent fees and different roll-up rates; single premium $P = 100$, maturity $T = 15$, barrier $\beta = Pe^{\delta T}$.

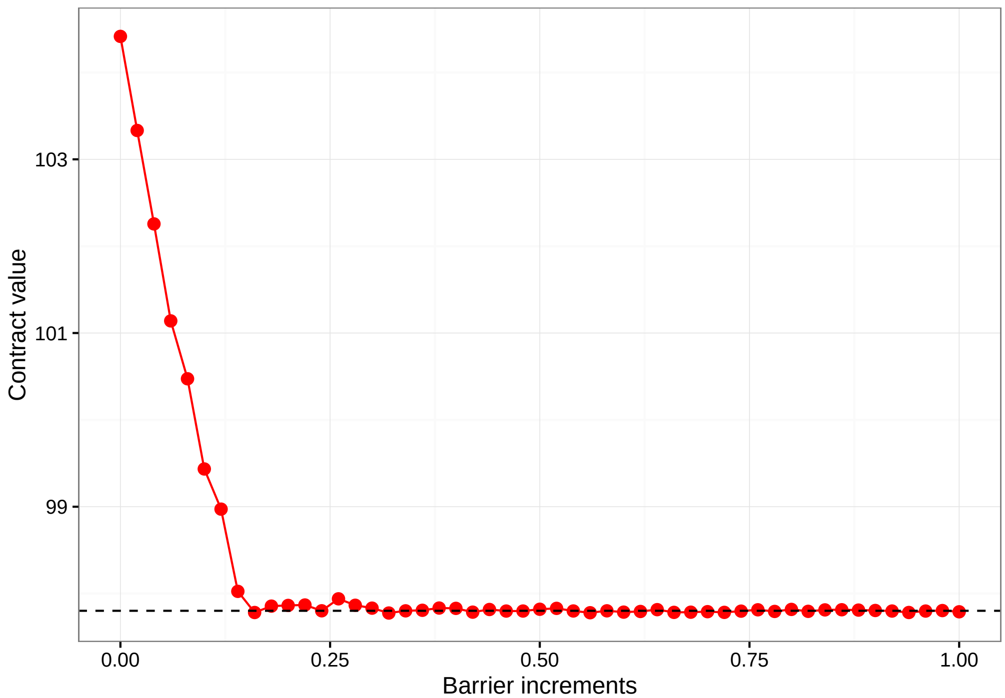

Figure 22: **Model 4**: The initial contract value, under the **mixed** approach, versus the barrier increment k; single premium $P = 100$, maturity $T = 15$, **state-dependent** fee rate $\varphi = 3\%$, roll-up rate $\delta = 0$, barrier $\beta = (1 + k)P$.

Acknowledgments.

This work has been done in the context of the research project on 'Metodi e modelli matematici e statistici per la valutazione e gestione del rischio in ambito finanziario e assicurativo', financed by the University of Trieste.

References

A.R. Bacinello, E. Biffis, and P. Millossovich. Pricing life insurance contracts with early exercise features. *Journal of Computational and Applied Mathematics*, 233(1):27–35, 2009.

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

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

T. Bae and B. Ko. On pricing equity-linked investment products with a threshold expense structure. *Korean Journal of Applied Statistics*, 23(4):621–633, 2010.

T. Bae and B. Ko. Pricing maturity guarantee under a refracted brownian motion. *Lobachevskii Journal of Mathematics*, 34(3):234–247, 2013.

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

E. Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468, 2005.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

P.P. Boyle and E.S. Schwartz. Equilibrium prices of guarantees under equity-linked contracts. *The Journal of Risk and Insurance*, 44(4):639–660, 1977.

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

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

L. Delong. Pricing and hedging of variable annuities with state-dependent fees. *Insurance: Mathematics and Economics*, 58:24–33, 2014.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, third edition, 2001.

S.L. Heston. 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, 1993.

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

M.A. Milevsky and S.E. Posner. The titanic option: Valuation of the guaranteed minimum death benefit in variable annuities and mutual funds. *The Journal of Risk and Insurance*, 68(1):93–128, 2001.

M.A. Milevsky and T.S. Salisbury. The real option to lapse a variable annuity: Can surrender charges complete the market? 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. Available on line at `www.pru.co.uk/pdf/FIPK10080.pdf` (accessed on 14 November 2016), 2016.

J. Zhou and L. Wu. The time of deducting fees for variable annuities under the state-dependent fee structure. *Insurance: Mathematics and Economics*, 61:125–134, 2015.