Risk Management of Policyholder Behavior in Equity-Linked Life Insurance

Anne MacKay*, Maciej Augustyniak†, Carole Bernard‡ and Mary R. Hardy§ ¶

April 9, 2015

Abstract

The financial guarantees embedded in variable annuity (VA) contracts expose insurers to a wide range of risks, lapse risk being one of them. When policyholders' lapse behavior differs from the assumptions used to hedge VA contracts, the effectiveness of dynamic hedging strategies can be significantly impaired. By studying how the fee structure and surrender charges affect surrender incentives, we obtain new theoretical results on the optimal surrender region and use them to design a marketable contract that is never optimal to lapse.

Keywords: Variable annuities, pricing, GMMB, dynamic hedging, surrender behavior.

*Corresponding author: Anne MacKay is with the department of Mathematics at ETH Zürich, Email anne.mackay@math.ethz.ch.

†M. Augustyniak is with the department of Mathematics and Statistics at the University of Montreal, Email augusty@dms.umontreal.ca.

‡C. Bernard is with the department of Finance at Grenoble Ecole de Management, Email carole.bernard@grenoble-em.com.

§M.R. Hardy is with the department of Statistics and Actuarial Science at the University of Waterloo, Email mrhardy@uwaterloo.ca.

¶All authors acknowledge support from NSERC. C. Bernard and M.R. Hardy acknowledge support from research grants awarded by the Global Risk Institute and the Society of Actuaries CAE (Centers of Actuarial Excellence). A. MacKay also acknowledges the support of the Hickman scholarship of the Society of Actuaries. The authors would like to thank P. Forsyth, B. Li, A. Kolkiewicz and D. Saunders for helpful discussions.

1 Introduction

Variable annuities (VAs) and other types of equity-linked insurance products have grown in popularity over the last 20 years. These products combine a participation in equity performance with insurance features and are therefore normally sold by insurance companies. For example, a policyholder entering into a VA contract pays an initial premium, and chooses to invest it in one or more mutual funds.[1] However, unlike mutual funds, VAs offer various financial guarantees upon death of the policyholder or at maturity of the contract that protect the policyholder's capital against market downturns. Some VAs also guarantee minimum periodic withdrawals or income amounts for a fixed or varying period of time (see Hardy, 2003 for more details).

Financial options embedded in VAs have payoff structures that are similar to those of options traded on stock exchanges, but include three distinctive features that complicate their pricing and risk management. First, they include guarantees that are triggered upon death of the policyholder or after a long-term maturity, which generally exceeds five years. Second, they are financed by a fee, which is typically paid as a fixed proportion of the account value, as opposed to being paid upfront. Third, the policyholder has the flexibility to withdraw the account value and cancel the contract before maturity by paying a penalty (which may be zero after an initial period) to the insurer. These nonstandard features give rise to risks that are difficult to mitigate, such as mortality risk, equity risk, interest rate risk and policyholder behavior risk. As many insurance companies suffered large losses on their VA business during the recent financial crisis, it is vital to examine original ways to manage these risks.

This paper focuses on surrender or lapse risk[2] and studies how an insurer can mitigate this risk with product design. Surrender risk refers to the risk that policyholders terminate their contract prior to maturity and it is a specific type of policyholder behavior risk. Surrenders may impair the insurer's business for several reasons. First, the insurance company may

[1] In some cases, additional premium amounts can be deposited in the VA account at regular intervals.

[2] Some authors distinguish between "lapse" and "surrender". A lapse may refer to the policyholder not taking advantage of the possibility to renew an expiring policy (by ceasing premium payments and without receiving any payout from the insurer), or any voluntary cessation without a surrender payment to the policyholder (Dickson, Hardy, and Waters (2013)). The "surrender" (or cancellation) refers to the specific action by a policyholder during the policy term to terminate the contract and recover the surrender value. See for instance Eling and Kochanski (2013) for details. Throughout this paper, as in Eling and Kochanski (2013), we will not make this distinction given that we do not consider the renewing option but only the surrender option.

not be able to fully recover the initial expenses and upfront costs of acquiring new business (Pinquet, Guillén, and Ayuso, 2011), as well as setting up a hedge for the guarantees in the contract. Second, surrenders may cause liquidity issues (and loss of future profits) (Kuo, Tsai, and Chen,2003) as there is potential for large cash demands in very short time frames. Finally, surrenders may give rise to an adverse selection problem because policyholders with insurability issues tend not to lapse their policies (on this topic, see Benedetti and Biffis, 2013). Due to these negative consequences, insurers typically include surrender charges in their VA products to discourage lapsation (Milevsky and Salisbury, 2001). These charges are generally high in the first few years of the contract to provide a way for the insurer to recover acquisition expenses. Although surrender charges act as a disincentive for policyholders to lapse, there are many situations where surrendering can be advantageous for the policyholder, even after accounting for surrender penalties.

Various methods have been proposed to model policyholder lapse behavior. They range from simple, deterministic lapse rates to sophisticated models such as De Giovanni (2010)'s rational expectation and Li and Szimayer (2014)'s limited rationality. Knoller, Kraut, and Schoenmaekers (2015) show that the moneyness of the embedded option plays a role in surrender behavior. Kuo, Tsai, and Chen (2003) and Tsai (2012) study the relationship between lapse rates of life insurance policies and the level of the interest rate, and find a link between both rates over a long-term horizon. A recent empirical study by Eling and Kiesenbauer (2014) based on the German life market shows that product design and policyholder characteristics have a statistically significant impact on lapse rates, but finds that unit-linked contracts are not surrendered more often than traditional life insurance policies, which do not incorporate any equity-linked insurance components. Pinquet, Guillén, and Ayuso (2011) claim that sub-optimal lapses are caused by insufficient knowledge of insurance products. Eling and Kochanski (2013) provide an extensive review of the recent literature on lapses and highlight the challenges faced by insurers in terms of modeling lapse behavior and mitigating this risk. Overall, identifying the appropriate lapse model is challenging because of lack of data and of the difficulty to identify the real motivation behind a policyholder's decision to surrender the contract by observing data on lapses.

Given that modeling lapses is complex and far from being well-understood, the insurer can adopt a conservative approach to price and hedge the VA contract by assuming that all policyholders are rational and surrender their policies optimally (from a financial engineering standpoint) (see for example Milevsky and Salisbury, 2001; Bauer, Kling, and Russ, 2008; Bernard, MacKay, and Muehlbeyer, 2014).[3] Nevertheless, assuming the worst case scenario for policyholder behavior can result in a product that is too expensive to be marketable, and that is very complicated to manage and hedge. To simplify risk management, the insurer may decide to ignore lapses when implementing a hedging strategy. However, Kling, Ruez, and Ruß (2014) conclude that the effectiveness of hedging strategies can be highly compromised when lapse experience deviates from the VA issuer's assumptions.

The main contribution of this article is to show how the insurer can develop a VA design for fees and surrender charges that eliminates the surrender incentive, while keeping the contract marketable and attractive to the policyholder (simple design, low fees and low surrender charges). In the proposed design, the worst case scenario for policyholder behavior corresponds to not surrendering the policy. Therefore, the insurer is not required to model surrender behavior for pricing and hedging purposes because by assuming that the policyholder will remain in the VA contract until death or maturity, the insurer will conservatively price and hedge this contract.[4] As a result, the risk of having an inappropriate lapse model is mitigated and risk management of the VA contract is simplified.

We start from a standard VA contract combining guaranteed minimum death benefit (GMDB) and guaranteed minimum accumulation benefit (GMAB) riders financed by a fee paid continuously as a constant percentage of the account value. We then derive an explicit closed-form expression for a model-free minimal surrender charge schedule that eliminates the surrender incentive for all account values throughout the whole term of the contract. However, we find that these minimal surrender penalties are generally too high to lead to a marketable VA product. To address this issue, we consider the state-dependent fee structure introduced by Bernard, Hardy, and MacKay (2013), where the fee is paid to the insurer only when the account value is below a certain threshold (see also Delong, 2014; Moenig and Zhu, 2014 and Zhou and Wu, 2015).[5] The motivation is the following: since the state-dependent fee structure reduces the incentive to lapse the VA contract, it can

[3]The optimal surrender decision is an optimal stopping problem that can be technically challenging. This is analogous to dealing with an American option, and pricing the maturity guarantee while ignoring optimal lapses is similar to pricing a European option.

[4]This hedging strategy is in reality a superhedge. In fact, if the policyholder keeps her policy until maturity, this strategy is exactly replicating the liabilities for a large portfolio. In case of early surrender, our contract design is such that the payout to the policyholder is always less than the value of the liabilities. Since the hedging portfolio of the insurer is established to replicate these liabilities, the difference between the value of this portfolio and the payout represents a profit for the insurer. This idea is further developed in Bauer, Bergmann, and Kiesel (2010).

[5]Similar fee structures were also studied independently by Bae and Ko (2010, 2013).

be combined with smaller surrender charges to achieve our goal of removing the surrender incentive. We show how to solve for the minimal surrender charge function that attains this objective and conclude that it is possible to design an attractive and marketable VA contract that eliminates the surrender incentive by combining a state-dependent fee with surrender charges.

This article is organized as follows. Section 2 introduces the market model, the VA contract, and the partial differential equation (PDE) approach used for pricing. Section 3 presents an analysis of the optimal surrender incentive when the fee is paid as a constant percentage of the fund throughout the term of the contract, and shows how this incentive can be eliminated. Section 4 performs a similar analysis in the state-dependent fee case, and also presents an example of a contract design that eliminates the surrender incentive. Section 5 concludes.

2 Pricing the VA contract

2.1 Market model and notation

We consider a VA contract with maturity T and underlying account value at time t denoted by F_t, $t \in [0,T]$. Suppose that the initial premium F_0 is fully invested in an index whose value process $\{S_t\}_{0\leqslant t\leqslant T}$ has real-world ($\mathbb{P}$-measure) dynamics

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^{\mathbb{P}},$$

where $W_t^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion.[6] Suppose also that the usual assumptions of the Black-Scholes model are satisfied. Therefore, the market is complete and there exists a unique risk-neutral measure $\mathbb{Q}$ under which the index S_t follows a geometric Brownian motion with drift equal to the risk-free rate r, so that

$$\frac{dS_t}{S_t} = r dt + \sigma dW_t^{\mathbb{Q}}.$$

[6]We work on a filtered probability space $(\Omega, \mathcal{F}^M, \{\mathcal{F}_t^M\}_{0\leqslant t\leqslant T}, \mathbb{P})$ where $(\Omega, \mathcal{F}^M)$ is a measurable space, $\{\mathcal{F}_t^M\}_{0\leqslant t\leqslant T}$ is the natural filtration generated by the Brownian motion (with $\mathcal{F}_t^M = \sigma(\{W_s^{\mathbb{P}}\}_{0\leqslant s\leqslant t})$) and $\mathbb{P}$ is the real-world measure. We assume that the probability space is complete ($\mathcal{F}_0$ contains the $\mathbb{P}$-null sets) and that the filtration $\{\mathcal{F}_t\}_{0\leqslant t\leqslant T}$ is right-continuous.

We assume that the financial guarantee embedded in the VA contract is financed by a fee paid continuously as a constant percentage c of the account value. This characteristic differentiates the VA from options traded on stock exchanges, since these are generally financed by a charge paid up-front. This constant fee structure, which is typical for VAs, is problematic because it gives rise to a mismatch between the liability of the insurance company (the financial guarantee) and its income (the future fees that will be collected before maturity or surrender). For instance, when the value of the underlying fund increases, the liability of the insurer decreases while the expected value of the future fee income moves in the opposite direction, i.e., increases. This mismatch creates an important surrender incentive since the policyholder is paying a high price for a guarantee that has little value (for more details, see Milevsky and Salisbury (2001) and Bernard, MacKay, and Muehlbeyer (2014)).

To address this problem, we modify the VA design and allow for a state-dependent fee, as first proposed by Bernard, Hardy, and MacKay (2013). Under a state-dependent fee structure, the insurer only charges the fee when the account value is below a given level β, called the fee barrier threshold. Since the evolution of the account value depends on this threshold β, we add the superscript (β) to the symbol F_t. The $\mathbb{P}$-dynamics of the account value are given by

$$\frac{dF_t^{(\beta)}}{F_t^{(\beta)}} = (\mu - c\mathbb{1}_{\{F_t^{(\beta)}<\beta\}})dt + \sigma dW_t^{\mathbb{P}}, \tag{1}$$

where $\mathbb{1}_A$ is the indicator function of the set A. Without loss of generality, we assume that $F_0^{(\beta)} = S_0$. When $\beta = \infty$, the fee is paid throughout the term of the contract regardless of the account value, and equation (1) simplifies to

$$\frac{dF_t^{(\infty)}}{F_t^{(\infty)}} = (\mu - c)dt + \sigma dW_t^{\mathbb{P}}.$$

The case $\beta = \infty$ corresponds to the typical fee structure considered for VA contracts in the literature, and will be referred to as the "constant fee case" in this paper.

2.2 Guaranteed benefits and surrender charges

We consider a VA contract with a T-year GMAB rider guaranteeing a payoff of

$$\max(G_T, F_T^{(\beta)}),$$

to the policyholder at maturity T, where G_T denotes a pre-determined guarantee at time T. In addition, we include a roll-up GMDB rider that pays out a benefit of

$$\max(G_t, F_t^{(\beta)}),$$

to the policyholder if she dies before the contract matures at time $0 < t < T$. We assume that

$$G_t = e^{gt} F_0^{(\beta)}, \quad 0 < t \leqslant T,$$

where $0 \leqslant g < r$ represents the guaranteed roll-up rate.

The policyholder is allowed to surrender the contract at any time $0 < t < T$, in which case she receives $(1 - \kappa_t)F_t^{(\beta)}$: the account value $F_t^{(\beta)}$ diminished by the surrender charge $\kappa_t F_t^{(\beta)}$, where $0 \leqslant \kappa_t \leqslant 1$. Typically, the surrender charge κ_t is a decreasing function of time to discourage policyholders from lapsing early on in the contract.[7] Since the contract cannot be surrendered at maturity, we set $\kappa_T = 0$.

We consider two decreasing surrender charge functions, in addition to the case $\kappa_t = 0$ for all t. First, we employ a *vanishing* surrender charge function, $\kappa_t = \kappa(1-t/T)^3$. This function mimics surrender penalties found on the market, which are typically high in the first years of the contract, and drop rapidly to add liquidity to the VA investment. Second, we use the function $\kappa_t = 1 - e^{-\kappa(T_1 - t\wedge T_1)}$, where $0 < T_1 \leqslant T$, which results in approximately linear surrender charges from $t = 0$ to T_1. Such a surrender charge schedule is similar to the one included in New York Life's Premier Variable Annuity (New York Life, 2014) which starts at 8% and decreases by 1% each year thereafter until zero.[8]

[7]Early surrenders have a detrimental impact on the insurer's business because VA contracts have front-loaded expenses that are recouped from fees collected during the first few years of the contract.

[8]We could have also considered a piecewise constant surrender charge function in our analysis. However, we preferred to use smooth surrender charge schedules because they lead to smooth surrender regions.

2.3 Mortality model

Since the VA contract that we consider includes death benefits, we must incorporate mortality risk. Therefore, we introduce the filtration $\{\mathcal{F}_t^B\}_{0\leqslant t\leqslant T}$ that carries binary information on the mortality of the policyholder (alive or dead). We let $\mathcal{F}_t = \mathcal{F}_t^M \vee \mathcal{F}_t^B$ for all $0 \leqslant t \leqslant T$. By adding mortality risk, the market becomes incomplete and, as a result, there exists infinitely many equivalent risk-neutral measures (see Dhaene, Kukush, Luciano, Schoutens, and Stassen (2013) for more details). Similarly to Bauer, Kling, and Russ (2008) and Hardy (2000), we assume risk-neutrality of the insurer with respect to mortality risk and choose to price under the measure that preserves independence between market and mortality risks. The first assumption implies that death probabilities are identical under real-world and risk-neutral measures, and that the insurer does not require to earn excess returns for assuming mortality risk. This can be justified on the basis that mortality risk is fully diversifiable. The second assumption implies that the risk-neutral measure for both market and mortality risks is the product measure of $\mathbb{Q}$ and the measure for mortality risk. For simplicity, we will also denote this product measure by $\mathbb{Q}$.

The future lifetime of a policyholder age x is represented by the random variable $\rho > 0$ and the force of mortality at age x is denoted by λ_x (we diverge from the usual actuarial notation, which uses μ_x for the force of mortality, to avoid confusion with μ used in the asset price model). The probability that a policyholder age $x+t$ survives to age $x+u$, $0 \leqslant t < u$, is given by

$$_{u-t}p_{x+t} = e^{-\int_t^u \lambda_{x+s} ds}.$$

In numerical illustrations, we assume that mortality follows Makeham's law, i.e., the force of mortality at age x is modeled as

$$\lambda_x = A + Bc^x, \quad x > 0,$$

with $A = 0.0001$, $B = 0.00035$, and $c = 1.075$. This mortality model yields life expectancies of 21.7, 15.1 and 9.8 years at ages 50, 60 and 70, respectively.

2.4 Valuation of the VA contract

We assume that the VA contract is issued to a policyholder age x and that death benefits are paid out at the time of death. We let $V(t, F_t^{(\beta)})$ denote the value of the contract at time

$t \in [0,T]$ given that the policyholder is alive at age $x+t$. Since the VA contract can be surrendered at any time before maturity, its pricing becomes an optimal stopping problem. To define this problem, we must introduce further notation. Denote by $\mathcal{T}_t$ the set of all stopping times τ greater than or equal to t and bounded by T. Then, for a policyholder alive at time t, define the *continuation value* of the VA contract as

$$\begin{aligned} V^*(t, F_t^{(\beta)}) &= \sup_{\tau \in \mathcal{T}_t} E_{\mathbb{Q}} \left[e^{-r(\tau - t)} \psi(\tau, F_\tau^{(\beta)}) \mathbb{1}_{\{\rho > \tau\}} + e^{-r(\rho - t)} \max(G_\rho, F_\rho^{(\beta)}) \mathbb{1}_{\{\rho \leqslant \tau\}} \mid \mathcal{F}_t \right] \\ &= \sup_{\tau \in \mathcal{T}_t} E_{\mathbb{Q}} \left[E_{\mathbb{Q}} \left[e^{-r(\tau - t)} \psi(\tau, F_\tau^{(\beta)}) \mathbb{1}_{\{\rho > \tau\}} + e^{-r(\rho - t)} \max(G_\rho, F_\rho^{(\beta)}) \mathbb{1}_{\{\rho \leqslant \tau\}} \mid \mathcal{F}_t, \tau, F_\tau^{(\beta)} \right] \mid \mathcal{F}_t \right] \\ &= \sup_{\tau \in \mathcal{T}_t} E_{\mathbb{Q}} \left[e^{-r(\tau - t)} \psi(\tau, F_\tau^{(\beta)})_{\tau - t} p_{x+t} + \int_t^\tau e^{-r(u-t)} \max(G_u, F_u^{(\beta)})_{u-t} p_{x+t} \lambda_{x+u} du \,\middle|\, \mathcal{F}_t \right], \end{aligned} \tag{2}$$

where

$$\psi(t, F_t^{(\beta)}) = \begin{cases} (1 - \kappa_t) F_t^{(\beta)}, & \text{if } t \in [0, T), \\ \max(G_T, F_T^{(\beta)}), & \text{if } t = T, \end{cases}$$

is the payoff of the contract at surrender or maturity. The expression inside the expectation of equation (2) reflects the assumption that mortality risk is fully diversifiable for the insurance company. Similar expressions were obtained by Bauer, Kling, and Russ (2008) and Milevsky and Salisbury (2001).

The price of the VA contract at time $t \in [0,T]$ is then given by

$$V(t, F_t^{(\beta)}) = \begin{cases} V^*(t, F_t^{(\beta)}), & \text{if } F_t^{(\beta)} \in \mathcal{C}_t, \\ \psi(t, F_t^{(\beta)}), & \text{if } F_t^{(\beta)} \in \mathcal{S}_t, \end{cases}$$

where $\mathcal{C}_t$ and $\mathcal{S}_t$ are two disjoint sets of fund values satisfying $\mathcal{C}_t \cup \mathcal{S}_t = [0, \infty)$, and representing, respectively, the continuation and optimal surrender regions. We define $\mathcal{C}_t$ as the set of fund values for which the continuation value of the VA contract is strictly greater than the surrender benefit:

$$\mathcal{C}_t = \{F_t^{(\beta)} < \infty : V^*(t, F_t^{(\beta)}) > \psi(t, F_t^{(\beta)})\}.$$

Because the continuation value of the contract can never be strictly less than the surrender benefit, the optimal surrender region is given by

$$\mathcal{S}_t = \{F_t^{(\beta)} < \infty : V^*(t, F_t^{(\beta)}) = \psi(t, F_t^{(\beta)})\}.$$

When the VA fee is paid regardless of the account value ($\beta = \infty$), the optimal surrender region at time t, if it exists, is of threshold type, that is $\mathcal{S}_t = \{F_t^{(\infty)} \geqslant B_t\}$, with or without surrender penalties (see Bernard, MacKay, and Muehlbeyer (2014) for more details). The symbol B_t represents the fund threshold which induces a rational policyholder to surrender her VA contract at time t. The curve $\{(t, B_t) : t \in (0, T)\}$ is usually referred to as the optimal surrender boundary. Our analysis in Section 4 shows that the optimal surrender region in the case of a state-dependent fee is not necessarily of threshold type.

Unless otherwise indicated, we assume that VA contracts are fairly priced. The fair fee is defined as the fee rate c^* satisfying

$$F_0^{(\beta)} = V(0, F_0^{(\beta)}; c^*),$$

where $V(0, F_0^{(\beta)}; c^*)$ is the price of the contract evaluated at the fee rate c^*.

2.5 PDE representation of the VA contract price

Using the Feynman-Kac representation formula (see, for example, pp. 73–75 of Björk (2004)), we see that (2) is the solution to a partial differential equation (PDE). Therefore, we can obtain the price and optimal surrender region of the VA contract by numerically solving this PDE. In the Black-Scholes framework, under the usual no-arbitrage assumptions, $V(t, F_t^{(\beta)})$ must satisfy the following PDE in the continuation region $\mathcal{C}_t$,[9]

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial {F_t^{(\beta)}}^2}{F_t^{(\beta)}}^2\sigma^2 + \frac{\partial V}{\partial F_t^{(\beta)}}F_t^{(\beta)}\left(r - c\mathbb{1}_{\{F_t^{(\beta)}<\beta\}}\right) - (r + \lambda_{x+t})V + \lambda_{x+t}\max(G_t, F_t) = 0, \tag{3}$$

for $0 \leqslant t \leqslant T$ (and $F_t^{(\beta)} \in \mathcal{C}_t$). In the optimal surrender region $\mathcal{S}_t$, we have

$$V(t, F_t^{(\beta)}) = \psi(t, F_t^{(\beta)}),$$

[9] A similar PDE representation for the constant fee case is provided by Ulm (2006).

for $0 \leqslant t \leqslant T$ (and $F_t^{(\beta)} \in \mathcal{S}_t$).

We impose the following boundary conditions:

$$\begin{aligned} &V(T, F_T^{(\beta)}) = \max(G_T, F_T^{(\beta)}), \\ &\lim_{F_t^{(\beta)} \to 0} V(t, F_t^{(\beta)}) = V(t, 0) = G_T e^{-r(T-t)} {}_{T-t}p_{x+t} + \int_t^T G_u e^{-r(u-t)} {}_{u-t}p_{x+t} \lambda_{x+u} du. \end{aligned} \tag{4}$$

The first boundary condition is simply the VA payoff at maturity. The second boundary condition corresponds to a lower bound for the contract value which assumes $F_t^{(\beta)} = 0$ and, therefore, that it is certain that the guarantee is in the money at death or maturity.

To solve the PDE in (3), we must also specify an upper boundary condition. However, the behavior of the contract price for high account values depends on the fee structure, and it is generally not possible to specify this boundary exactly for a finite value of $F_t^{(\beta)}$.

2.5.1 Upper boundary in the constant fee case

In the constant fee case, when surrender penalties are not large enough to eliminate the surrender incentive, an upper boundary can be specified exactly because the price of the VA contract corresponds to the surrender benefit for sufficiently high fund values.

On the other hand, Section 3.2 shows that there exists a minimal surrender charge function such that the optimal strategy corresponds to not surrendering the policy at all times. The asymptotic behavior of the contract price in this case can be calculated by choosing $\tau = T$ in expression (2), so that

$$\lim_{F_t^{(\infty)} \to \infty} \frac{V(t, F_t^{(\infty)})}{F_t^{(\infty)}} = e^{-c(T-t)} {}_{T-t}p_{x+t} + \int_t^T e^{-c(u-t)} {}_{u-t}p_{x+t} \lambda_{x+u} du.$$

However, finite difference methods can be avoided to solve the problem in this specific case, because the price of the VA contract has an analytical representation that includes an integral averaging Black-Scholes European put prices over the policyholder's future lifetime distribution (see, for example, Hardy (2003); Ulm (2008)).

2.5.2 Upper boundary in the state-dependent fee case

In Appendix A, we show that the following asymptotic behavior holds in the state-dependent fee case regardless of the form assumed for the surrender charge function:

$$\lim_{F_t^{(\beta)}\to\infty} \frac{V(t, F_t^{(\beta)})}{F_t^{(\beta)}} = 1. \tag{5}$$

This result implies that we may use $F_t^{(\beta)}$ as an upper boundary for $V(t, F_t^{(\beta)})$ when solving the PDE in (3) numerically. This limiting behavior can be justified intuitively as follows. When $F_t^{(\beta)} \gg \beta$, the maturity benefit has almost no value, and the policyholder does not expect to pay any more fees, implying that

$$E_{\mathbb{Q}}\left(e^{-r(u-t)} \max(G_u, F_u^{(\beta)}) \mid \mathcal{F}_t\right) \approx E_{\mathbb{Q}}\left(e^{-r(u-t)} F_u^{(\beta)} \mid \mathcal{F}_t\right) \approx F_t^{(\beta)}, \quad t \leqslant u \leqslant T.$$

Assuming no fees will be paid until death or maturity eliminates the surrender incentive and allows us to value the contract as

$$\begin{aligned}
&E_{\mathbb{Q}}\left[e^{-r(T-t)} \max(G_T, F_T^{(\beta)})\,{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)} \max(G_u, F_u^{(\beta)})\,{}_{u-t}p_{x+t}\lambda_{x+u} du \,\middle|\, \mathcal{F}_t\right] \\
&\quad \approx F_t^{(\beta)} \left({}_{T-t}p_{x+t} + \int_t^T {}_{u-t}p_{x+t}\lambda_{x+u} du\right) \\
&\quad = F_t^{(\beta)}.
\end{aligned}$$

Under a state-dependent fee, Bernard, Hardy, and MacKay (2013) derived integral representations for the prices of GMAB and GMDB riders assuming that the VA contract cannot be surrendered. Since we will not make this assumption in the following sections, we will use the PDE approach to value VA contracts as it allows us to treat all possible surrender assumptions. Appendix B provides additional details on the numerical implementation of this approach.

3 VA contract in the constant fee case

This section analyzes how surrender penalties impact the fair fee charge and the surrender incentive in the constant fee case. Throughout this section, $\beta = \infty$, so we omit the superscript on the symbol $F_t^{(\beta)}$ and write $F_t = F_t^{(\infty)}$. We explain how to design a VA contract without any surrender incentive, and show that the surrender penalties needed to remove this incentive in the constant fee case lead to contracts which may be difficult to market. Understanding the interplay between the fee rate, surrender charges and the surrender incentive in this case is the first step towards designing a contract that eliminates the surrender incentive while offering reasonable fee rates and surrender charges.

3.1 Fair fee and optimal surrender region

We study VA contracts with an initial investment of $F_0 = 100$, $T = 10$ and 20-year maturities and a guaranteed roll-up rate of $g = 0$ issued to policyholders age $x = 50$, 60 or 70. The volatility parameter of the market model was fitted to a data set of weekly percentage log-returns on the S&P500 from October 28, 1987 to October 31, 2012, from which we obtained $\sigma = 0.165$. We further assume $r = 0.03$.

Table 1 presents fair fees calculated for the constant fee case under different assumptions for the surrender charge function. We consider two examples of the realistic surrender charge schedules introduced in Section 2.2, in addition to the extreme cases of $\kappa_t = 0$ and $\kappa_t = 1$ for all t. Note that the case $\kappa_t = 1$ for all t directly eliminates any surrender incentive, and simplifies the valuation of the VA contract to a European option pricing problem.

Maturity	$T = 10$			$T = 20$		
Age at issue	50	60	70	50	60	70
$\kappa_t = 0$	0.0393	0.0442	0.0549	0.0195	0.0266	0.0415
$\kappa_t = 0.05(1 - t/T)^3$	0.0184	0.0200	0.0234	0.0078	0.0102	0.0152
$\kappa_t = 1 - e^{-0.008(10-t\wedge 10)}$	0.0127	0.0139	0.0164	0.0073	0.0090	0.0127
$\kappa_t = 1$ (no surrender)	0.0115	0.0126	0.0148	0.0050	0.0065	0.0099

Table 1: Fair fee in the constant fee case ($\beta = \infty$).

When the insurer does not charge a penalty for early surrender, the fee income represents its only revenue. This income compensates the insurer for both the guarantee offered and early surrender risk (the risk of not being able to collect future fees on the account value). Table 1 shows that to fully mitigate lapse risk in the absence of surrender charges ($\kappa_t = 0$), the insurer must charge a fee which may be too high to be marketable.[10] When $T = 20$ and $x = 50$ or 60, the resulting fair fee may actually lie in the higher end of a marketable range. However, it must be noted that the fair fees we compute represent lower bounds for the fee rates an insurer would charge. In practice, the insurer must be compensated for administrative expenses as well as make a profit, and may want to include a buffer in the fee to account for model risk.

One way to reduce the fair fee and still fully mitigate lapse risk is to introduce surrender penalties in the product design. These penalties represent an additional revenue for the insurer, enabling it to decrease the constant fee charge, and also work as a disincentive to lapse. Table 1 shows that the two surrender charge schedules we consider help reduce the fair fee by more than 50%.

With a lower fair fee, the fund threshold at which the guarantee becomes less valuable than the expected future fees increases. In addition, surrender charges reduce the amount received at lapsation, further increasing this threshold. The combination of these two effects therefore results in an upward shift of the optimal surrender boundary. Figures 1 and 2 plot these boundaries for a VA product issued to a 60-year-old policyholder with maturity $T = 10$ and 20, respectively. Clearly, the two realistic surrender charge schedules considered in Table 1 are unable to eliminate the surrender incentive completely. This conclusion could also have been made by noting that the fair fees associated with these schedules are higher than those for the case $\kappa_t = 1$ for all t (see Table 1). As a portion of these higher fees must compensate the insurer for lapse risk, a surrender incentive must exist.

It is interesting to note that the exponential surrender charge function we consider is sufficient to eliminate the surrender incentive in the first years of the contract for the case $T = 20$, but not for $T = 10$. This is due to the lower fee rate associated with the 20-year contract. In general, a VA product with a lower fee rate can be combined with smaller surrender charges to eliminate the surrender incentive. In the next section, we investigate this observation in greater depth.

[10] Milevsky and Salisbury (2001) argued that VA contracts without surrender charges are not marketable for an additional reason: a rational policyholder would lapse the contract if the account value rises sharply right after contract inception. This is simply because the optimal surrender boundary must intersect F_0 at $t = 0$ due to the constraint imposed by the fair fee, i.e., $F_0 = V(0, F_0)$.

3.2 Minimal surrender charge to eliminate surrender incentives

Figures 1 and 2 illustrate that the two realistic surrender charge schedules examined in Table 1 do not include high enough penalties to remove the surrender incentive, and are only successful at reducing it. We now determine the minimal surrender charge schedule that completely eliminates this incentive. The main motivations for such a contract design are (i) to reduce the fair fee to a minimum so that this fee need not offer a compensation for lapsation risk and (ii) to simplify the implementation of the optimal hedging strategy for the insurer. A VA contract without any surrender incentive avoids the need to model lapse behavior in the hedging strategy, because the optimal strategy is limited to hedging maturity and death benefits only. As lapses are sub-optimal in this context, they must necessarily result in a profit for the insurer.

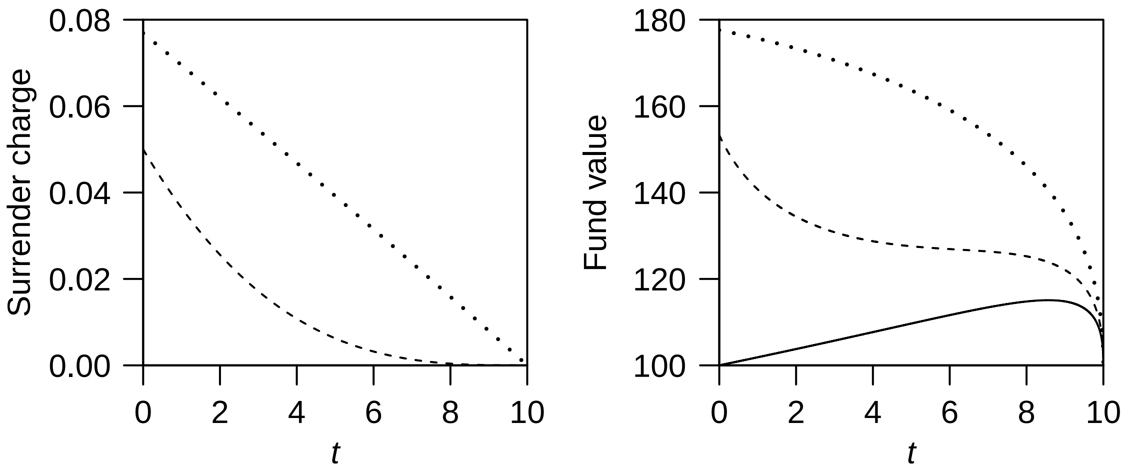

Figure 1: Fairly priced VA products with maturity $T = 10$. *Left panel*: Surrender charge function. *Right panel*: Optimal surrender boundary. *Legend*: solid line: $\kappa_t = 0$; dashed line: $\kappa_t = 0.05(1 - t/T)^3$; dotted line: $\kappa_t = 1 - e^{-0.008(10-t\wedge 10)}$.

To design such a policy, we look for a surrender charge function κ_t satisfying,

$$V^*(t, F_t) > (1 - \kappa_t)F_t, \quad \forall F_t \geqslant 0, \text{ and } \forall t \in [0, T). \tag{6}$$

That is, surrender penalties must be sufficiently high so that the continuation value of the contract, $V^*(t, F_t)$, is larger than the surrender benefit, $(1 - \kappa_t)F_t$, for any given time t. Proposition 3.1 provides the minimal surrender charge function that is needed to satisfy condition (6). Its proof is given in Appendix C.

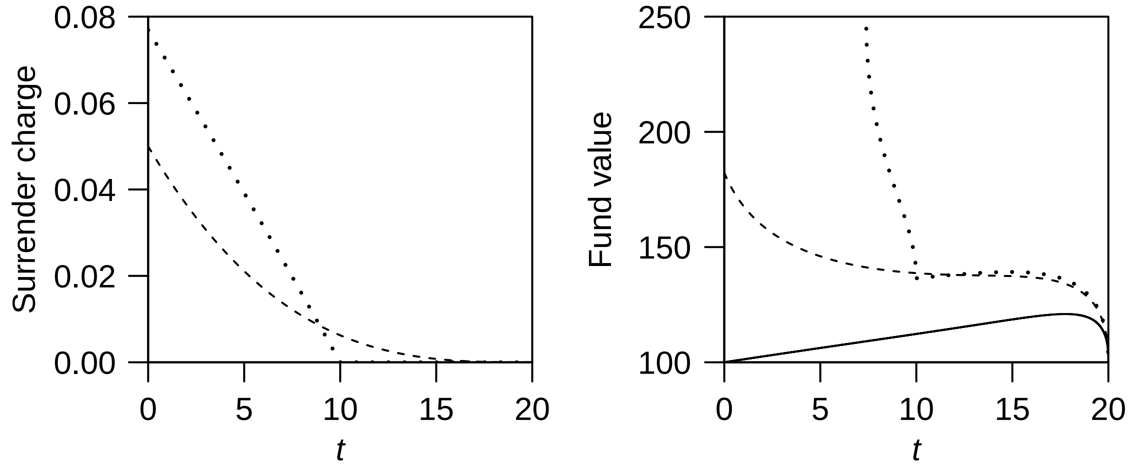

Figure 2: Fairly priced VA products with maturity $T = 20$. *Left panel*: Surrender charge function. *Right panel*: Optimal surrender boundary. *Legend*: solid line: $\kappa_t = 0$; dashed line: $\kappa_t = 0.05(1 - t/T)^3$; dotted line: $\kappa_t = 1 - e^{-0.008(10-t\wedge 10)}$.

Proposition 3.1. *Using notation from Section 2 in the constant fee case ($\beta = \infty$), the minimal value of κ_t at each time $t \in [s, T)$ such that it is never optimal to surrender the policy in the time interval $[s, T)$ is equal to*

$$\kappa_t^\star = 1 - e^{-c(T-t)}{}_{T-t}p_{x+t} - \int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du, \quad s \leqslant t < T.$$

The result given by Proposition 3.1 holds for any arbitrage-free complete market model (not just for the Black-Scholes model), as long as the pricing measure assumes risk-neutrality with respect to mortality risk and independence between mortality and financial risks. It shows that if $\kappa_t \geqslant \kappa_t^\star$ for all future $t < T$, then the contract will not be surrendered by a rational policyholder.

Remark 3.1. The condition $\kappa_t \geqslant \kappa_t^\star$ is also sufficient to guarantee that it is not optimal to lapse at time t, regardless of the form assumed for κ_u, $t < u < T$. To understand why, observe that

$$\begin{aligned}
\frac{V^*(t,F_t)}{F_t(1-\kappa_t)} &\geqslant \frac{\text{fair value of maturity and death benefits}}{\text{value of surrender benefit}} \\
&= \frac{E_{\mathbb{Q}}\left[e^{-r(T-t)}\max(G_T,F_T)_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}\max(G_u,F_u)_{u-t}p_{x+t}\lambda_{x+u}du \,\middle|\, \mathcal{F}_t\right]}{F_t(1-\kappa_t)} \\
&> \frac{e^{-c(T-t)}{}_{T-t}p_{x+t} + \int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du}{1-\kappa_t}, \quad \forall F_t > 0.
\end{aligned}$$

It is clear that, for any value of F_t at time t, whenever $\kappa_t \geqslant \kappa_t^\star$, the continuation value, $V^*(t,F_t)$, must be strictly greater than the surrender value, $F_t(1-\kappa_t)$, which means that surrender is sub-optimal. The converse of this result does not necessarily hold: if surrendering is not optimal at time t, $\forall F_t \geqslant 0$, then it is not necessarily true that $\kappa_t \geqslant \kappa_t^\star$. In other words, it is possible that surrender is not optimal at time t even if $\kappa_t \in [0,\kappa_t^\star)$. Such a situation can only occur if the condition $\kappa_u \geqslant \kappa_u^\star$ is not always satisfied $\forall u \in (t,T)$. In sum, a surrender charge of $\kappa_t^\star$ at time t is always high enough to eliminate the surrender incentive at time t. However, it will correspond to the minimal surrender charge achieving this objective only if it is never optimal to surrender the contract after time t.

The first step to design a fairly priced VA product which does not give rise to a surrender incentive, but with minimal surrender charges, is to compute the fair fee assuming that the contract cannot be surrendered, or equivalently, $\kappa_t = 1$, $\forall t$. This step corresponds to a European option pricing problem and only involves computing the fair value of maturity and death benefits. The second step is then to compute the minimal surrender charge schedule by applying Proposition 3.1.

In Figure 3, we plot the resulting minimal surrender charge schedules associated with 10 and 20-year VA contracts issued to a policyholder age 60. These surrender charge functions, along with the corresponding fair fees (see last line of the caption of Figure 3), generate fairly priced VA product designs that do not give rise to a surrender incentive, i.e., rational policyholders will never choose to lapse these contracts. Unfortunately, these designs include high surrender penalties that could significantly impact the marketability of the VA product. In fact, surrender charges start at a value above 8% and drop below 5% only halfway through the contract. Such penalties are slightly higher and their decrease is slower than the charges observed on the market, especially in the case of contracts with longer maturities. For example, surrender charges in New York Life's Premier Variable

Annuity (New York Life (2014)) start at 8% and decrease by 1% each year thereafter until zero. This is less than the surrender charges we obtain, and our fee structure does not include any security margin. Thus, it is reasonable to conclude that surrender charges designed to completely eliminate the optimal surrender incentive would be higher than those currently offered on the market.

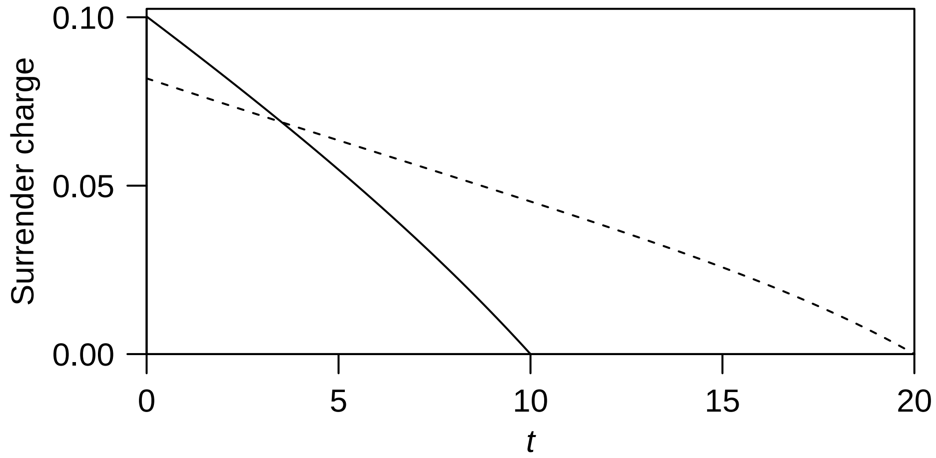

Figure 3: Minimal surrender charge schedule that completely eliminates the surrender incentive in the constant fee case for a fairly priced VA contract issued to a policyholder age 60. *Legend*: solid line: $T = 10$ and $c = 0.0126$; dashed line: $T = 20$ and $c = 0.0065$.

For marketing reasons, it may be desirable for an insurer to offer VA products with low surrender charges after a few years. In fact, VAs compete with mutual funds, which have little to no surrender charges, and the relatively lower liquidity of VAs can be used as an argument against them. For this reason, insurers might prefer VA designs that offer a protection against lapse risk with lower surrender charges. In the next section, we show how the introduction of a state-dependent fee into the VA product design can help achieve this objective and lead to product designs that appear interesting for both the insurer and the policyholder.

4 VA contract in the state-dependent fee case

In Section 3, we have shown that in the constant fee case (the typical fee structure in the industry) the surrender penalties required to eliminate the surrender incentive can be very high under reasonable market assumptions. We now explore how a state-dependent fee structure can help solve this problem by examining its impact on the optimal surrender region. The main motivation for considering this fee structure is the following. Since fees are paid only when the fund value is below a certain threshold, the incentive to surrender the contract is significantly reduced at high account values. Therefore, this fee structure can be combined with smaller surrender charges to discourage lapsation.

The VA contract and assumptions considered in this section are identical to those in Section 3, with the exception that fees are now paid only when the account value is below a certain threshold $\beta < \infty$. Bernard, Hardy, and MacKay (2013) studied the case $\beta = F_0^{(\beta)}$ in the absence of surrender charges. As this design leads to an unrealistically high fair fee, we choose to consider a threshold of $\beta = 150$ in our numerical illustrations because it leads to practical contract designs.

4.1 Fair fee and optimal surrender region

Table 2 presents fair fees calculated in the state-dependent fee case with $\beta = 150$ under different assumptions for the surrender charge function. In general, these fair fees are higher than those computed in the constant fee case (see Table 1), because the expected income of the insurer is reduced in the presence of a fee barrier threshold. However, note that the fair fee is the same for $\beta = 150$ and ∞ when $\kappa_t = 0$. This observation suggests that the state-dependent fee structure may not always lead to a decrease in the surrender incentive.

Maturity	$T = 10$			$T = 20$		
Age at issue	50	60	70	50	60	70
$\kappa_t = 0$	0.0393	0.0442	0.0549	0.0195	0.0266	0.0415
$\kappa_t = 0.05(1 - t/T)^3$	0.0190	0.0205	0.0237	0.0096	0.0119	0.0163
$\kappa_t = 1 - e^{-0.008(10-t\wedge 10)}$	0.0167	0.0179	0.0204	0.0098	0.0120	0.0165
$\kappa_t = 1$ (no surrender)	0.0166	0.0177	0.0202	0.0093	0.0114	0.0155

Table 2: Fair fee in the state-dependent fee case with a fee barrier threshold of $\beta = 150$.

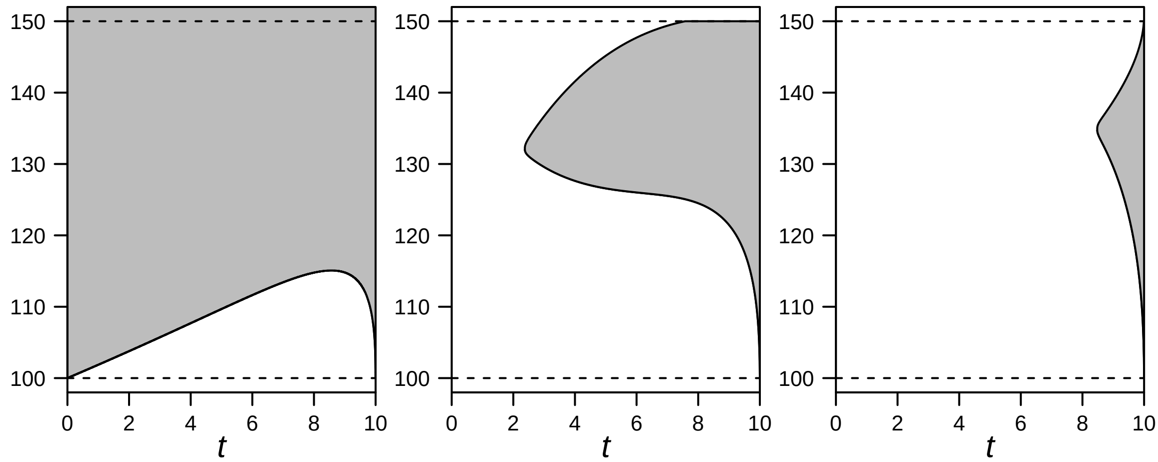

Figure 4: Optimal surrender region in the state-dependent fee case with $\beta = 150$ for a 10-year VA contract issued to a policyholder age 60. *Left panel*: $\kappa_t = 0$ and $c = 0.0442$. *Middle panel*: $\kappa_t = 0.05(1 - t/T)^3$ and $c = 0.0205$. *Right panel*: $\kappa_t = 1 - e^{-0.008(10 - t\wedge 10)}$ and $c = 0.0179$.

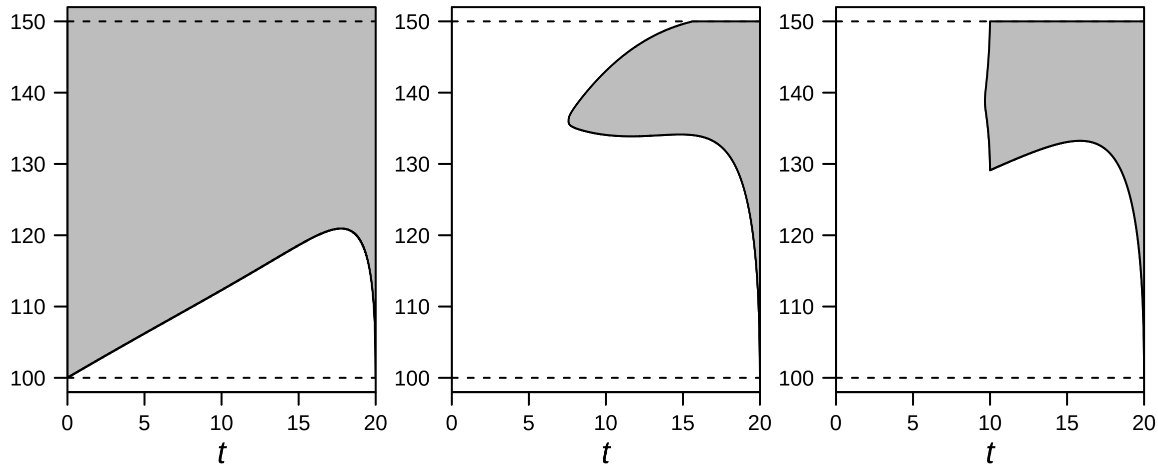

Figure 5: Optimal surrender region in the state-dependent fee case with $\beta = 150$ for a 20-year VA contract issued to a policyholder age 60. *Left panel*: $\kappa_t = 0$ and $c = 0.0266$. *Middle panel*: $\kappa_t = 0.05(1 - t/T)^3$ and $c = 0.0119$. *Right panel*: $\kappa_t = 1 - e^{-0.008(10 - t\wedge 10)}$ and $c = 0.0120$.

Figures 4 and 5 illustrate the optimal surrender regions for a VA product issued to a policyholder age 60 with maturity $T = 10$ and 20, respectively. The left panel in these figures is associated with $\kappa_t = 0$, and clearly shows that it is always optimal to lapse before the account value reaches the fee barrier threshold β. Therefore, a rational policyholder will never wait for the fund to reach the fee barrier threshold of $\beta = 150$, and from her perspective, constant or state-dependent fee designs are equivalent. This explains why the fair fee and the optimal surrender regions are the same when $\beta = 150$ and ∞.

Nevertheless, when surrender charges are introduced, the middle and right panels of Figures 4 and 5 reveal that the state-dependent fee structure can significantly impact the optimal surrender region (for example, compare Figures 1 and 2 to Figures 4 and 5). Four important observations can be made.

(i) The optimal surrender strategy is no longer of threshold type.

(ii) It is not optimal for the policyholder to lapse the contract when the account value is close to or above the fee barrier threshold β.

(iii) The surrender incentive is eliminated in the early years of the contract term. The high surrender penalties early in the contract incentivizes the policyholder to wait for the account value to grow above the fee barrier threshold.

(iv) In general, the state-dependent fee structure reduces the surrender incentive, at the cost of increasing the fair fee.

Observation (ii) can be proven to hold in general and is formalized in Proposition 4.1 (the proof is given in Appendix D).

Proposition 4.1. *Let $F_t^{(\beta)}$ and κ_t be defined as in Section 2 and assume $\beta < \infty$. Then, for any $t \in [0, T]$,*

$$V^*(t, F_t^{(\beta)}) \geqslant F_t^{(\beta)}(1 - \kappa_t), \quad \forall\, F_t^{(\beta)} \geqslant \beta. \tag{7}$$

If $\kappa_t > 0$ at time t, the inequality in (7) *is strict.*

Proposition 4.1 simply states that when the fee is state-dependent and the account value is above the fee barrier threshold β, the contract is always worth at least as much as the surrender benefit $F_t^{(\beta)}(1 - \kappa_t)$. The intuition for this result is the following. If the account value is above β, the policyholder does not have a clear incentive to surrender because the VA product offers her a maturity guarantee and a death benefit for which she is not required to pay for at the moment (the guarantees are offered for *free*). As a result, a rational policyholder will wait for the account value to reach β, before even considering surrendering the contract.

Remark 4.1. When the surrender charge at time t is greater than 0, that is, $\kappa_t > 0$, Proposition 4.1 implies that the optimal surrender region cannot include fund values above or equal to β. However, as explained in the proof of Proposition 4.1 (see Appendix D), when $\kappa_t = 0$, we could have,

$$V^*(t, F_t^{(\beta)}) = F_t^{(\beta)}, \quad \forall F_t^{(\beta)} \geqslant \beta. \tag{8}$$

By the definition of the optimal surrender region given in Section 2.2 (the policyholder is assumed to lapse when the continuation value of the contract is equal to the surrender benefit), the equality in equation (8) actually induces surrender. This argument explains why the optimal surrender region includes fund values above β in the left panels of Figures 4 and 5. However, strictly speaking, a rational policyholder would actually be indifferent to lapse in this situation because the surrender benefit is exactly equal to the continuation value of the contract assuming a specific surrender strategy (see Appendix D for more details).

4.2 Minimal surrender charge to eliminate surrender incentives

In Section 3, we derived a closed-form expression for the minimal surrender charge function that eliminates the surrender incentive in the constant fee case ($\beta = \infty$), and showed that the resulting surrender penalties may be too high to be of practical value. We now explain how to determine the corresponding minimal surrender charge function in the state-dependent fee case ($\beta < \infty$) and show an example where the resulting contract design appears marketable under reasonable assumptions. For ease of notation, we denote the total value of maturity and death benefits (that is, the value of the VA contract if surrenders are not allowed) by $U(t, F_t^{(\beta)})$, where

$$U(t, F_t^{(\beta)}) = E_{\mathbb{Q}}\left[e^{-r(T-t)} \max(G_T, F_T^{(\beta)})_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)} \max(G_u, F_u^{(\beta)})_{u-t}p_{x+t}\lambda_{x+u} du \,\middle|\, \mathcal{F}_t \right].$$

Obtaining a simple closed-form expression for the minimal surrender charge function when $\beta < \infty$ is challenging due to the following reasons. First, although there exists an integral representation for the value of the maturity benefit (see, Bernard, Hardy, and MacKay (2013)), or for the discounted expectation of the account value at future times, the expressions involved depend on the current account value $F_t^{(\beta)}$, typically in a complex way.

Second, the function $U(t, F_t^{(\beta)})/F_t^{(\beta)}$ is generally not monotone in $F_t^{(\beta)}$ because the expected future fees are themselves not necessarily monotone in $F_t^{(\beta)}$ when a state-dependent fee is considered, i.e., the fee income at t increases with $F_t^{(\beta)}$ while $F_t^{(\beta)} < \beta$, but drops to 0 as soon as $F_t^{(\beta)} \geqslant \beta$. Therefore, we must resort to a numerical procedure to solve for the minimal surrender schedule that eliminates the surrender incentive in the state-dependent fee case. This procedure is outlined in the following steps.

Step 1: Find the fair value of the fee rate c assuming that the contract cannot be surrendered, or equivalently, $\kappa_t = 1$, $\forall t$.

Step 2: For each $t \in [0, T)$, numerically obtain the account value $F_t^\star$ at which the ratio of the value of maturity and death benefits to the fund value is the smallest:[11]

$$F_t^\star = \arg\inf_{F_t^{(\beta)} \geqslant 0} \left\{ \frac{U(t, F_t^{(\beta)})}{F_t^{(\beta)}} \right\}.$$

Step 3: Finally, set

$$\kappa_t = \max\left(1 - \frac{U(t, F_t^\star)}{F_t^\star}, 0\right).$$

The fair fee in Step 1 can be computed with a PDE approach, or via Proposition 3.1 in Bernard, Hardy, and MacKay (2013). The infima $F_t^\star$ in Step 2 are easily obtained after setting up a finite difference grid for the value of the European contract because this grid provides us with values of $U(t, F_t^{(\beta)})$ at discrete values of t and $F_t^{(\beta)}$ in the grid.

Note that to obtain the minimal surrender charge schedule, it is not necessary to solve an optimal stopping problem. The algorithm only requires us to calculate the present value of maturity and death benefits at different points in time and for different account values $F_t^{(\beta)}$. When these present values cannot be obtained semi-analytically (as in Bernard, Hardy, and MacKay (2013)) or are difficult to implement with PDE methods (due to product or model complexity), they can generally be easily computed using Monte Carlo simulations.

[11] This step implicitly assumes that the continuation value of the VA contract is equal to the value of the contract if surrenders are not allowed, i.e.,

$$V^*(t, F_t^{(\beta)}) = U(t, F_t^{(\beta)}),$$

which is the correct assumption to make because κ_t is chosen in such a way that lapsation is not optimal during the entire length of the contract.

4.3 Marketable VA product with no surrender incentive

We apply the numerical procedure just introduced to 10 and 20-year VA contracts with a fee barrier threshold of $\beta = 150$ issued to a policyholder age 60. The fair fees for these contracts when there is no surrender incentive corresponds to $c = 0.0177$ and 0.0114 for contract maturities of $T = 10$ and 20 years, respectively (see first line of Table 2). The left panels in Figures 6 and 7 illustrate the minimal surrender charge schedule that eliminates the surrender incentive for each of the two contracts considered. The values $F_t^\star$ used in the calculation of the surrender charge are given in the right panels of these figures. Observe that $V(t, F_t^\star) = (1 - \kappa_t)F_t^\star$ along this boundary, and that surrender is never optimal on either side of the boundary. It is therefore never clearly optimal for the policyholder to lapse her contract.

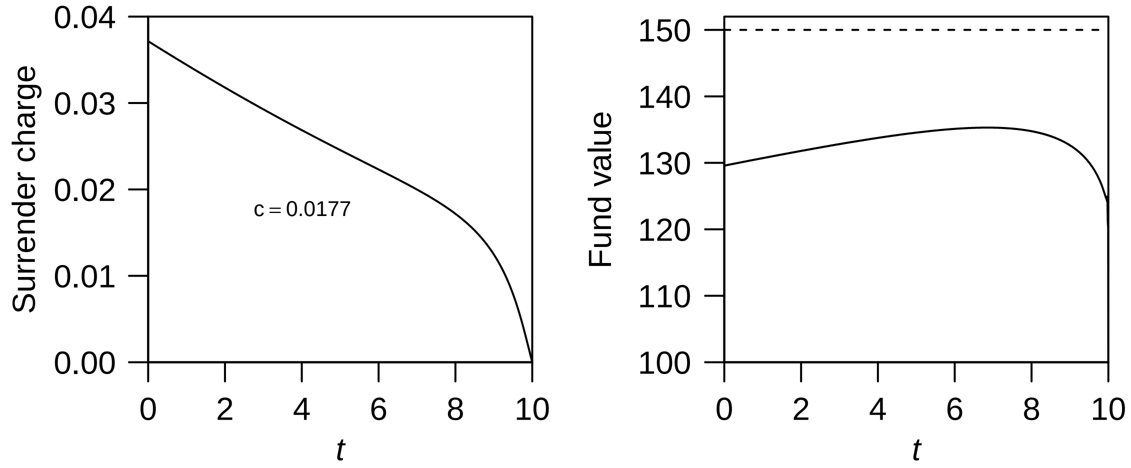

Figure 6: VA contract issued to a policyholder age 60 with maturity $T = 10$, fair fee rate $c = 0.0177$ and fee barrier threshold $\beta = 150$. *Left panel*: Minimal surrender charge function not giving rise to an optimal lapsation boundary. *Right panel*: Values of $F_t^\star$.

These contract designs include surrender charges and fees that appear marketable. For example, surrender penalties are below 3% and 2% during most of the contract term for maturities $T = 10$ and 20 years, respectively. This is significantly lower than the minimal surrender charge schedule required to eliminate the surrender incentive in the constant fee case (see Section 3.2).

We can further compare these two product designs to a typical constant fee product design with $\beta = \infty$ and the exact same schedule of surrender charges. It turns out that when pricing this contract under optimal policyholder behavior, the fair value of c when $\beta = \infty$ is identical to the one when $\beta = 150$, and the optimal lapsation boundaries for contract maturities $T = 10$ and 20 correspond to the curves in the right panel of Figures 6 and 7, respectively.

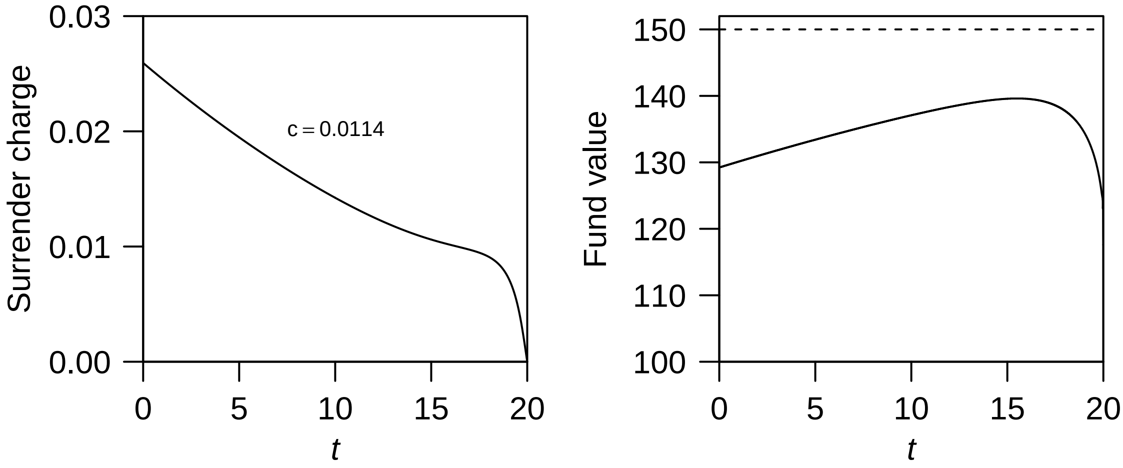

Figure 7: VA contract issued to a policyholder age 60 with maturity $T = 20$, fair fee rate $c = 0.0114$ and fee barrier threshold $\beta = 150$. *Left panel*: Minimal surrender charge function not giving rise to an optimal lapsation boundary. *Right panel*: Values of $F_t^\star$.

This result may seem surprising at first, but has an intuitive explanation. First, note that for a given surrender charge function, and assuming that the policyholder lapses optimally, the state-dependent fair fee is an upper bound the constant fair fee. This is due to the fact that under the state-dependent fee design, the fee might be paid over a period of time shorter than in the constant fee case. Second, consider a policyholder who lapses as soon as the account value hits the curve $\{(t, F_t^\star) : t \in (0, T)\}$ and assume that this curve always lies below β (see the right panel of Figures 6 and 7). Since along this curve, the surrender benefit is exactly equal to the continuation value of the contract when $\beta = 150$, this surrender strategy allows the policyholder to extract the maximum value from that contract and is thus optimal when $\beta = 150$ (as is not surrendering). Moreover, note that a policyholder following this surrender strategy would have received the exact same payoff and paid the exact same fees if the product design had $\beta = \infty$. Consequently, the fair value of c computed for $\beta = 150$ must be a lower bound for the fair fee when $\beta = \infty$, as it is the fair c under *one* possible surrender strategy. Therefore, the fair value of c cannot change when we modify the fee barrier threshold from $\beta = 150$ to $\beta = \infty$, because we have shown that this value is both an upper and lower bound for the fair fee when $\beta = \infty$.

Since contracts with $\beta = \infty$ and 150 charge the same fee rate and have the same surrender penalties, designs with $\beta = 150$ are more attractive to a policyholder than the ones with $\beta = \infty$ due to the presence of a threshold above which the fee is not paid. However, from a risk management standpoint, it can be argued that these designs are also preferable for the insurer because they do not give rise to an optimal surrender region. As a consequence, the VA product can be managed assuming no surrenders (as this is the optimal behavior when $\beta = 150$), which simplifies the construction of the hedging portfolio and reduces the importance of modeling lapses for pricing and hedging purposes.

5 Concluding Remarks

This paper provided some insights to answer a very practical question: How can an insurer use product design to mitigate lapse risk and simplify risk management (hedging) in VAs? To answer this question, we examined the interplay between the fee structure of a VA contract combining GMDB and GMAB riders and the schedule of surrender charges. We showed that by considering a state-dependent fee structure and relatively low surrender charges, an insurer can design a contract that will never be optimal to lapse, while still being marketable. We also argued that such an outcome can generally not be achieved with a constant fee structure.

The risks embedded in this new contract design are easier for the insurer to manage because only the (European) value of maturity and death benefits offered need to be considered to construct the most effective hedging strategy. In fact, it is no longer necessary to model surrender behavior for pricing and hedging purposes as surrenders are sub-optimal in this context. In other words, if the insurer's hedging portfolio replicates the value of death and maturity benefits, this portfolio will always be sufficient to pay for the surrender benefit if the policyholder decides to lapse before maturity. The difference between the value of this portfolio and the surrender benefit represents an instant profit to the insurer that can offer a compensation for the liquidity strain created by premature lapsation.

Moreover, not having to consider an optimal stopping problem when establishing the hedging strategy is advantageous to the insurer for at least two reasons. First, it is very impractical for the insurer to have to account for all of the different optimal surrender regions associated with the VA products sold. When these products do not give rise to a surrender incentive, the hedging strategy can be implemented in a uniform manner across the portfolio of VAs. Second, the insurer can use straightforward Monte Carlo simulations, rather than PDE methods or computationally demanding backward procedures, to value the hedging portfolio and calculate its Greeks. This is especially important because VA products sold in the market have many complex features, and Monte Carlo simulations are often the only feasible alternative to value these products.

Our focus on optimal surrender behavior can also be justified by the possibility of secondary markets for equity-linked life insurance. In fact, Gatzert, Hoermann, and Schmeiser (2009) explain how both consumers and insurers can benefit from a secondary market for life insurance contracts. On the one hand, consumers get a better price (lower surrender charges) and on the other hand, it makes the life insurance market more attractive and thus potentially increases the demand for the insurer's products. Hilpert, Li, and Szimayer (2014) further discuss the impact of the surrender option and the existence of a secondary market for equity-linked life insurance, and show that the introduction of sophisticated investors by the secondary market may lead to higher premiums that account for a higher proportion of optimal surrenders. This result is somewhat consistent with Gatzert, Hoermann, and Schmeiser (2009) who explain that life insurers need to abandon lapse-supported pricing (i.e. pricing under the assumption of sub-optimal lapses that benefit the insurer). In other words, as the existence of a secondary market mechanism should suppress contract arbitrages, policyholders having access to this market should tend to act optimally.

Further research should investigate the robustness of product design and dynamic hedging strategies under various market models. The product design analysis could also be extended to VA contracts offering other types of financial guarantees.

References

Bae, T., and B. Ko (2010): "On Pricing Equity-Linked Investment Products with a Threshold Expense Structure," *Korean Journal of Applied Statistics*, 23(4), 621–633.

——— (2013): "Pricing maturity guarantee under a refracted Brownian motion," *Lobachevskii Journal of Mathematics*, 34(3), 234–247.

Bauer, D., D. Bergmann, and R. Kiesel (2010): "On the risk-neutral valuation of life insurance contracts with numerical methods in view," *Astin Bulletin*, 40(1), 65–95.

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

Benedetti, D., and E. Biffis (2013): "Insurance Contract Design and Endogenous Frailty," *Working Paper*.

Bernard, C., M. Hardy, and A. MacKay (2013): "State-dependent fees for variable annuity guarantees," *ASTIN Bulletin*, 44(3), 1–27.

Bernard, C., A. MacKay, and M. Muehlbeyer (2014): "Optimal surrender policy for variable annuity guarantees," *Insurance: Mathematics and Economics*, 55, 116–128.

Björk, T. (2004): *Arbitrage Theory in Continuous Time*. Oxford University Press.

De Giovanni, D. (2010): "Lapse rate modeling: a rational expectation approach," *Scandinavian Actuarial Journal*, 2010(1), 56–67.

Delong, L. (2014): "Pricing and Hedging of Variable Annuities with State-Dependent Fees," *Insurance: Mathematics and Economics*, 58, 24–33.

Dhaene, J., A. Kukush, E. Luciano, W. Schoutens, and B. Stassen (2013): "On the (in-) dependence between financial and actuarial risks," *Insurance: Mathematics and Economics*, 52(3), 522–531.

Dickson, D. C., M. R. Hardy, and H. R. Waters (2013): *Actuarial Mathematics for Life Contingent Risks*. Cambridge University Press., Cambridge, England, 2nd edn.

Eling, M., and D. Kiesenbauer (2014): "What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market," *Journal of Risk and Insurance*, 81(2), 241–269.

Eling, M., and M. Kochanski (2013): "Research on Lapse in Life Insurance - What Has Been Done and What Needs to Be Done?," *Journal of Risk Finance*, 14(4), 5–5.

Gatzert, N., G. Hoermann, and H. Schmeiser (2009): "The Impact of the Secondary Market on Life Insurers' Surrender Profits," *Journal of Risk and Insurance*, 76(4), 887–908.

Hardy, M. R. (2000): "Hedging and reserving for single-premium segregated fund contracts," *North American Actuarial Journal*, 4(2), 63–74.

Hardy, M. R. (2003): *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Hilpert, C., J. Li, and A. Szimayer (2014): "The Effect of Secondary Markets on Equity-Linked Life Insurance With Surrender Guarantees," *Journal of Risk and Insurance*, 81(4), 943–968.

Kim, I. J., and G. G. Yu (1996): "An alternative approach to the valuation of American options and applications," *Review of Derivatives Research*, 1(1), 61–85.

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

Knoller, C., G. Kraut, and P. Schoenmaekers (2015): "On the Propensity to Surrender a Variable Annuity Contract-An Empirical Analysis of Dynamic Policyholder Behavior," *Journal of Risk and Insurance, forthcoming, DOI: 10.1111/jori.12076.*

Kuo, W., C. Tsai, and W.-K. Chen (2003): "An Empirical Study on the Lapse Rate: The Cointegration Approach," *Journal of Risk and Insurance*, 70(3), 489–508.

Li, J., and A. Szimayer (2014): "The effect of policyholders' rationality on unit-linked life insurance contracts with surrender guarantees," *Quantitative Finance*, 14(2), 327–342.

Milevsky, M. A., and T. S. Salisbury (2001): "The Real Option to Lapse a Variable Annuity: Can Surrender Charges Complete the Market," *Conference Proceedings of the 11th Annual International AFIR Colloquium*.

Moenig, T., and N. Zhu (2014): "Lapse and Reentry in Variable Annuities," *Available at SSRN 2503995*.

New York Life (2014): "Premier Variable Annuity," http://www.newyorklife.com/nyl-internet/file-types/NYL-Premier-VA-Fact-Sheet-No-CA-NY.pdf; Retrieved May 10, 2014.

Pinquet, J., M. Guillén, and M. Ayuso (2011): "Commitment and Lapse Behavior in Long-Term Insurance: A Case Study," *Journal of Risk and Insurance*, 78(4), 983–1002.

Tsai, C. (2012): "The Impacts of Surrender Options on Reserve Durations," *Risk Management and Insurance Review*, 15(2), 165–184.

Ulm, E. R. (2006): "The effect of the real option to transfer on the value of guaranteed minimum death benefits," *Journal of Risk and Insurance*, 73(1), 43–69.

——— (2008): "Analytic solution for return of premium and rollup guaranteed minimum death benefit options under some simple mortality laws," *ASTIN Bulletin*, 38(2), 543.

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

A Proof of Equation (5)

To prove (5) we need the two following lemmas.

Lemma A.1. *Let $F_t^{(\beta)}$, $0 \leqslant t \leqslant T$, be as defined in Section 2 and let $\beta < \infty$. Then,*

$$\lim_{y\to\infty} \frac{E_{\mathbb{Q}}[e^{-r(T-t)}F_T^{(\beta)}{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du | F_t^{(\beta)} = y]}{y} = 1. \tag{9}$$

Proof. We first show that for any $u \in (t, T]$,

$$\lim_{y\to\infty} \frac{E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)} | F_t^{(\beta)} = y]}{y} = 1. \tag{10}$$

Let $m_F(t,u) = \inf_{t\leqslant s\leqslant u} F_s^{(\beta)}$ and $m_S(t,u) = \inf_{t\leqslant s\leqslant u} S_s$ be the minimum values attained by the account and the index, respectively, between times t and u. Then,

$$\begin{aligned} &\frac{E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)} | F_t^{(\beta)} = y]}{y} = \\ &\quad \frac{E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}\mathbb{1}_{\{m_F(t,u)>\beta\}} | F_t^{(\beta)} = y]}{y} + \frac{E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}\mathbb{1}_{\{m_F(t,u)\leqslant\beta\}} | F_t^{(\beta)} = y]}{y}. \end{aligned} \tag{11}$$

To prove that $\lim_{y\to\infty} \frac{E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}|F_t^{(\beta)}=y]}{y} = 1$, we show that the first term of expression (11) goes to 1 as $y \to \infty$, and then show that the second term goes to 0 as $y \to \infty$.

Let $C_t = e^{-c\int_0^t \mathbb{1}_{\{F_s^{(\beta)}<\beta\}}ds}$ and note that C_t is $\mathcal{F}_t$-measurable. Observe that if $F_t^{(\beta)} = C_tS_t > \beta$, then

$$F_u^{(\beta)}\mathbb{1}_{\{m_F(t,u)>\beta\}} = C_tS_u\mathbb{1}_{\{m_S(t,u)>\frac{\beta}{C_t}\}}, \text{ a.s. for } t < u \leqslant T, \tag{12}$$

since the fee is not paid when the account value is above β. It follows that

$$E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}\mathbb{1}_{\{m_F(t,u)>\beta\}} | F_t^{(\beta)} = y] = C_tE_{\mathbb{Q}}\left[e^{-r(u-t)}S_u\mathbb{1}_{\{m_S(t,u)>\frac{\beta}{C_t}\}} \mid S_t = \frac{y}{C_t}\right]. \tag{13}$$

The expectation on the right-hand side of equation (13) is the price of a down-and-out contract on the underlying stock with barrier β/C_t and maturity u. Under the Black-Scholes model, the price of this option has a closed-form solution (see, for example, Chapter

18 of Björk (2004)), and we can write

$$C_t E_{\mathbb{Q}}\left[e^{-r(u-t)} S_u \mathbb{1}_{\{m_S(t,u) > \frac{\beta}{C_t}\}} | S_t = \frac{y}{C_t}\right] =$$

$$y\mathcal{N}\left(\frac{\ln\frac{y}{\beta} + \left(r + \frac{\sigma^2}{2}\right)(u-t)}{\sigma\sqrt{u-t}}\right) - \beta\left(\frac{\beta}{y}\right)^{\frac{2r}{\sigma^2}} \mathcal{N}\left(\frac{\ln\frac{\beta}{y} + \left(r + \frac{\sigma^2}{2}\right)(u-t)}{\sigma\sqrt{u-t}}\right),$$

where $\mathcal{N}(\cdot)$ denotes the standard normal cumulative distribution function. Thus,

$$\frac{C_t E_{\mathbb{Q}}[e^{-r(u-t)} F_u^{(\beta)} \mathbb{1}_{\{m_F(t,u) > \beta\}} | F_t^{(\beta)} = y]}{y}$$
$$= \mathcal{N}\left(\frac{\ln\left(\frac{y}{\beta}\right) + \left(r + \frac{\sigma^2}{2}\right)(u-t)}{\sigma\sqrt{u-t}}\right) - \left(\frac{\beta}{y}\right)^{\frac{2r}{\sigma^2}+1} \mathcal{N}\left(\frac{\ln\left(\frac{\beta}{y}\right) + \left(r + \frac{\sigma^2}{2}\right)(u-t)}{\sigma\sqrt{u-t}}\right).$$

The result follows since $\lim_{z\to\infty}\mathcal{N}(z) = 1$ and $\lim_{z\to-\infty}\mathcal{N}(z) = 0$.

To show that the second term on the right-hand side of (11) vanishes for large values of y, we first note that

$$\frac{E_{\mathbb{Q}}\left[e^{-r(u-t)} F_u^{(\beta)} \mathbb{1}_{\{m_F(t,u) \leqslant \beta\}} | F_t^{(\beta)} = y\right]}{y} \leqslant \frac{E_{\mathbb{Q}}\left[e^{-r(u-t)} S_u \mathbb{1}_{\{m_S(t,u) \leqslant \frac{\beta}{C_t}\}} | S_t = \frac{y}{C_t}\right]}{y}, \tag{14}$$

since for any $0 \leqslant t \leqslant T$, $F_t = S_t C_t \leqslant S_t$, a.s. The right-hand side of equation (14) is the price of a down-and-in contract on the underlying stock with barrier β/C_t. The price of this contract also has a closed-form solution (again, see Chapter 18 of Björk (2004)), which allows us to write

$$\frac{E_{\mathbb{Q}}[e^{-r(u-t)} F_u^{(\beta)} \mathbb{1}_{\{m_F(t,u) \leqslant \beta\}} | F_t^{(\beta)} = y]}{y} \leqslant$$
$$\frac{1}{C_t}\left\{\mathcal{N}\left(\frac{\ln\frac{\beta}{y} - (\widetilde{r} + \frac{\sigma^2}{2})(u-t)}{\sigma\sqrt{u-t}}\right) + \left(\frac{\beta}{y}\right)^{\frac{2r}{\sigma^2}+1} \mathcal{N}\left(\frac{\ln\frac{\beta}{y} + (\widetilde{r} + \frac{\sigma^2}{2})(u-t)}{\sigma\sqrt{u-t}}\right)\right\}.$$

Since $\lim_{y\to-\infty}\mathcal{N}(y) = 0$, $\lim_{y\to\infty} \frac{E_{\mathbb{Q}}[e^{-r(T-t)} F_T^{(\beta)} \mathbb{1}_{\{m_F(t,T) \leqslant \beta\}} | F_t^{(\beta)} = y]}{y} = 0$.

Then, we observe that

$$E_{\mathbb{Q}}\left[e^{-r(T-t)}F_T^{(\beta)}{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du|F_t^{(\beta)} = y\right]$$
$$= {}_{T-t}p_{x+t}E_{\mathbb{Q}}[e^{-r(T-t)}F_T^{(\beta)}|F_t^{(\beta)} = x] + \int_t^T {}_{u-t}p_{x+t}\lambda_{x+u}E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}|F_t^{(\beta)} = y]du.$$

Taking the limit as $y \to \infty$, we have

$$\lim_{y\to\infty}\left\{\frac{{}_{T-t}p_{x+t}E_{\mathbb{Q}}[e^{-r(T-t)}F_T^{(\beta)}|F_t^{(\beta)} = y] + \int_t^T {}_{u-t}p_{x+t}\lambda_{x+u}E_{\mathbb{Q}}[e^{-r(u-t)}F_u^{(\beta)}|F_t^{(\beta)} = y]du}{y}\right\}$$
$$= {}_{T-t}p_{x+t} + \int_t^T {}_{u-t}p_{x+t}\lambda_{x+u}du = {}_{T-t}p_{x+t} + {}_{T-t}q_{x+t} = 1.$$

$\square$

Lemma A.2. *Let $F_t^{(\beta)}$, $0 \leqslant t \leqslant T$, be as defined in Section 2. Then,*

$$\lim_{y\to\infty}\frac{y + \sup_{\tau\in\mathcal{T}_t}E_{\mathbb{Q}}[e^{-r(\tau-t)}(G_T - F_\tau^{(\beta)})^+|F_t^{(\beta)} = y]}{y} = 1,$$

where $(G_T - F_\tau^{(\beta)})^+ = \max(G_T - F_\tau^{(\beta)}, 0)$ and where $\mathcal{T}_t$ is the set of all stopping times larger than t and bounded by T.

Proof. Denote by $Put(t, S_t, T, G_T, \delta)$ the price at time t of an American put option with strike G_T and maturity T on a stock S_t paying dividends at a continuous rate δ. Using

$$\frac{S_T e^{-c(T-t)}}{S_t} < \frac{F_T^{(\beta)}}{F_t^{(\beta)}} < \frac{S_T}{S_t}, \qquad \text{a.s.},$$

it can be shown[12] that

$$Put(t, y, T, G_T, 0) \leqslant \sup_{\tau\in\mathcal{T}_t}E_{\mathbb{Q}}[e^{-r(\tau-t)}(G_T - F_\tau^{(\beta)})^+|F_t^{(\beta)} = y] \leqslant Put(t, y, T, G_T, c).$$

Since $\forall\delta \geqslant 0$, $\lim_{y\to\infty}Put(t, y, T, G_T, \delta) = 0$, the desired result follows from

$$\lim_{y\to\infty}\frac{E_{\mathbb{Q}}[e^{-r(T-t)}(G_T - F_T)^+|F_t^{(\beta)} = y]}{y} = 0.$$

$\square$

[12]To do so, use, for example, Equation (10) of Kim and Yu (1996).

Using Lemmas A.1 and A.2, we can now prove (5) by showing

$$\lim_{y\to\infty} \frac{V(t,y)}{y} = 1, \tag{15}$$

where $V(t,y)$ is the price at t of the VA contract with $\beta < \infty$, when $F_t^{(\beta)} = y$. First, we show that

$$\begin{aligned} E_{\mathbb{Q}}[e^{-r(T-t)}F_T^{(\beta)}{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du|F_t^{(\beta)} = y] \\ \leqslant V(t,F_t^{(\beta)}) \leqslant F_t^{(\beta)} + \sup_{\tau\in\mathcal{T}_t} E_{\mathbb{Q}}[e^{-r(\tau-t)}(G_T - F_\tau^{(\beta)})^+|\mathcal{F}_t]. \end{aligned} \tag{16}$$

To get the first inequality, observe that the full value of the contract is always less than or equal to its value without the surrender option.

To show the second inequality, recall that the payoff of the contract is either $(1-\kappa_u)F_u^{(\beta)}$ if the contract is surrendered at time $u < T$, or $F_u^{(\beta)} + (G_u - F_u^{(\beta)})^+ \leqslant F_u^{(\beta)} + (G_T - F_u^{(\beta)})^+$ a.s. at time $t < u \leqslant T$ if the policyholder dies before maturity or if the contract is kept until T (this follows from the definition of G_t, $0 \leqslant t \leqslant T$ in Section 2.2, which leads to $G_u \leqslant G_T$ for $u \leqslant T$). Thus, investing in the fund $F_t^{(\beta)}$ and buying an American put option on the fund with strike G_T will always yield a higher payout than the VA contract. It follows that

$$V(t,F_t^{(\beta)}) \leqslant F_t + \sup_{\tau\in\mathcal{T}_t} E_{\mathbb{Q}}[e^{-r(\tau-t)}(G_T - F_\tau^{(\beta)})^+|\mathcal{F}_t].$$

From (16),

$$\begin{aligned} &\frac{E_{\mathbb{Q}}[e^{-r(T-t)}F_T^{(\beta)}{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du|F_t^{(\beta)} = y]}{y} \\ &\qquad \leqslant \frac{V(t,y)}{y} \leqslant \frac{F_t^{(\beta)} + \sup_{\tau\in\mathcal{T}_t} E_{\mathbb{Q}}[e^{-r(\tau-t)}(G_T - F_\tau^{(\beta)})^+|F_t^{(\beta)} = y]}{y}. \end{aligned} \tag{17}$$

To complete the proof of (15), it suffices to take the limit of (17) as $y \to \infty$. The result follows from Lemma A.1 and Lemma A.2, since the first and the third terms of (17) both go to 1 in the limit. $\square$

B Additional details on the PDE pricing approach

We use a Crank-Nicolson finite difference scheme to solve the PDE in (3). The equation is discretized over a rectangular grid representing the discretized and upper truncated domain of $(t, F_t^{(\beta)})$. We consider time steps $dt = 0.001$ on the time domain $[0, T]$. Since the value of the VA contract can be computed exactly when $F_t^{(\beta)} = 0$ (see equation (4)), we use it to obtain the lower boundary in the grid in the space dimension. The upper truncation point of the grid must be large enough so that the asymptotic results derived in Section 2 can be used reliably to approximate the contract price at the highest fund values in the grid. In our numerical illustrations, we use a grid for $F_t^{(\beta)}$ which spans from 0 to 500 with steps of $dx = 0.01$. We tried higher levels for the upper bound, but found that this choice had marginal to no influence on numerical results. When an optimal surrender boundary exists for all $t \in [0, T]$, it is not necessary to consider values that are above this boundary, because the price of the contract is known exactly in this region (and equal to the value of the surrender benefit). In these cases, we use a lower maximal value to decrease computational time.

We validated the results with an explicit finite difference method, Monte Carlo simulations and the integral representation given in Bernard, Hardy, and MacKay (2013), and found no major convergence issues.

C Proof of Proposition 3.1

Proof. First, suppose that the surrender charge κ_u, for $t < u < T$, is sufficiently high to eliminate the surrender incentive for $t < u < T$. This situation is possible because we can consider the extreme case where $\kappa_u = 1$, for $t < u < T$. Then, the continuation value of the contract at time t must simply be the risk-neutral discounted expectation of death and maturity benefits, and be strictly greater than the surrender benefit:

$$\begin{aligned}V^*(t, F_t) &= E_{\mathbb{Q}}\left[e^{-r(T-t)} \max(G_T, F_T)_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)} \max(G_u, F_u)_{u-t}p_{x+t}\lambda_{x+u}du \,\middle|\, \mathcal{F}_t\right] \\ &> F_t(1-\kappa_t), \quad \forall F_t \geqslant 0.\end{aligned}$$

The previous inequality can be rewritten as $\kappa_t > 1 - \frac{V^*(t,F_t)}{F_t}$. Since $\forall F_t \geqslant 0$,

$$\begin{aligned}V^*(t, F_t) &= F_t\left(e^{-c(T-t)}{}_{T-t}p_{x+t} + \int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du\right) \\ &\quad + E_{\mathbb{Q}}\left[e^{-r(T-t)}(G_T - F_T)^+{}_{T-t}p_{x+t} + \int_t^T e^{-r(u-t)}(G_u - F_u)^+{}_{u-t}p_{x+t}\lambda_{x+u}du \,\middle|\, \mathcal{F}_t\right] \\ &> F_t\left(e^{-c(T-t)}{}_{T-t}p_{x+t} + \int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du\right),\end{aligned}$$

and $\lim_{F_t\to\infty} \frac{V^*(t,F_t)}{F_t} = e^{-c(T-t)}{}_{T-t}p_{x+t} + \int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du$, then the minimal surrender penalty that can be charged at time t while the inequality above is satisfied corresponds to

$$\begin{aligned}\kappa_t^\star &= \max\left(1-\inf_{F_t\geqslant 0}\left\{\frac{V^*(t,F_t)}{F_t}\right\},0\right)\\ &= \max\left(1-\lim_{F_t\to\infty}\frac{V^*(t,F_t)}{F_t},0\right)\\ &= 1-e^{-c(T-t)}{}_{T-t}p_{x+t}-\int_t^T e^{-c(u-t)}{}_{u-t}p_{x+t}\lambda_{x+u}du.\end{aligned}$$

□

D Proof of Proposition 4.1

Proof. We first consider the case without surrender charges, i.e., $\kappa_t = 0$, $\forall t$, that the fee is only paid below β, and that $F_t^{(\beta)} \geqslant \beta$ at time t. Consider the stopping time

$$\tau_\beta = \inf\left\{t < u < T : F_u^{(\beta)} < \beta\right\},$$

with the convention that $\tau_\beta = T$, if the barrier β is never reached. Then, we can write,

$$\begin{aligned}&V^*(t,F_t^{(\beta)}) = \sup_{\tau\in\mathcal{T}_t} E_{\mathbb{Q}}\left[e^{-r(\tau-t)}\psi(\tau,F_\tau^{(\beta)})_{\tau-t}p_{x+t}+\int_t^\tau e^{-r(u-t)}\max(G_u,F_u^{(\beta)})_{u-t}p_{x+t}\lambda_{x+u}du\;\middle|\;\mathcal{F}_t\right]\\ &\geqslant E_{\mathbb{Q}}\left[e^{-r(\tau_\beta-t)}\psi(\tau_\beta,F_{\tau_\beta}^{(\beta)})_{\tau_\beta-t}p_{x+t}+\int_t^{\tau_\beta} e^{-r(u-t)}\max(G_u,F_u^{(\beta)})_{u-t}p_{x+t}\lambda_{x+u}du\;\middle|\;\mathcal{F}_t\right]\\ &\geqslant E_{\mathbb{Q}}\left[e^{-r(\tau_\beta-t)}F_{\tau_\beta}^{(\beta)}{}_{\tau_\beta-t}p_{x+t}+\int_t^{\tau_\beta} e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du\;\middle|\;\mathcal{F}_t\right]\\ &= E_{\mathbb{Q}}\left[\left(e^{-r(\tau_\beta-t)}F_{\tau_\beta}^{(\beta)}{}_{\tau_\beta-t}p_{x+t}+\int_t^{\tau_\beta} e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du\right)\mathbb{1}_{\{\tau_\beta\in(t,T)\}}\;\middle|\mathcal{F}_t\right]\\ &\qquad + E_{\mathbb{Q}}\left[\left(e^{-r(T-t)}F_T^{(\beta)}{}_{T-t}p_{x+t}+\int_t^{T} e^{-r(u-t)}F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u}du\right)\mathbb{1}_{\{\tau_\beta=T\}}\;\middle|\mathcal{F}_t\right]\\ &\geqslant E_{\mathbb{Q}}\left[e^{-r(\tau_\beta-t)}F_{\tau_\beta}^{(\beta)}\mathbb{1}_{\{\tau_\beta\in(t,T)\}}\;\middle|\;\mathcal{F}_t\right]+E_{\mathbb{Q}}\left[e^{-r(T-t)}F_T^{(\beta)}\mathbb{1}_{\{\tau_\beta=T\}}\;\middle|\;\mathcal{F}_t\right]\\ &= \beta E_{\mathbb{Q}}\left[e^{-r(\tau_\beta-t)}\mathbb{1}_{\{\tau_\beta\in(t,T)\}}\mid\mathcal{F}_t\right]+E_{\mathbb{Q}}\left[e^{-r(T-t)}F_T^{(\beta)}\mathbb{1}_{\{\tau_\beta=T\}}\mid\mathcal{F}_t\right],\end{aligned}$$

where the first term is the payoff of a down rebate option which pays β if the fund $F_t^{(\beta)}$ reaches β before maturity T and zero otherwise, and the second term is the payoff of a down-and-out European call option with zero strike which pays $F_T^{(\beta)}$ at maturity T, provided that $F_u^{(\beta)} \geqslant \beta$, for $t \leqslant u \leqslant T$. Since the combined payoff of these two options can be replicated by holding the fund $F_t^{(\beta)}$ and selling it as soon as $F_t^{(\beta)} = \beta$, simple no-arbitrage arguments imply that the total price of these two options is exactly $F_t^{(\beta)}$, which gives

$$V^*(t, F_t^{(\beta)}) \geqslant F_t^{(\beta)}. \tag{18}$$

Note that the inequality in (18) becomes an equality if and only if τ_β is the optimal strategy that can be undertaken and $G_T < \beta$ (which implies that $\max(G_u, F_u^{(\beta)}) = F_u^{(\beta)}$, $\forall u \in [t, \tau_\beta]$). The equality, $V^*(t, F_t^{(\beta)}) = F_t^{(\beta)}$, therefore holds if and only if it is optimal to surrender the VA contract just below β, for all $t < u < T$. If this occurs, by definition of the optimal surrender region given in Section 2.2 ($V^*(t, F_t^{(\beta)}) = F_t^{(\beta)}$ induces surrender), the policyholder surrenders at time t, $\forall F_t^{(\beta)} \geqslant \beta$. However, strictly speaking, the policyholder is actually indifferent to lapse because the surrender benefit is exactly equal to the continuation value of the contract assuming a surrender strategy τ_β.

Now, consider the case where the surrender charge function κ_u, $t \leqslant u \leqslant T$, is a decreasing function of u, and is strictly positive at t:

$$\begin{aligned}
V^*(t, F_t^{(\beta)}) &\geqslant E_{\mathbb{Q}}\left[e^{-r(\tau_\beta - t)} \psi(\tau_\beta, F_{\tau_\beta}^{(\beta)})_{\tau_\beta - t}p_{x+t} + \int_t^{\tau_\beta} e^{-r(u-t)} \max(G_u, F_u^{(\beta)})_{u-t}p_{x+t}\lambda_{x+u} du \,\middle|\, \mathcal{F}_t \right] \\
&\geqslant E_{\mathbb{Q}}\left[\left(e^{-r(\tau_\beta - t)} F_{\tau_\beta}^{(\beta)} (1 - \kappa_{\tau_\beta})_{\tau_\beta - t}p_{x+t} + \int_t^{\tau_\beta} e^{-r(u-t)} F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u} du \right) \mathbb{1}_{\{\tau_\beta \in (t,T)\}} \,\middle|\, \mathcal{F}_t \right] \\
&\quad + E_{\mathbb{Q}}\left[\left(e^{-r(T-t)} F_T^{(\beta)}{}_{T-t}p_{x+t} + \int_t^{T} e^{-r(u-t)} F_u^{(\beta)}{}_{u-t}p_{x+t}\lambda_{x+u} du \right) \mathbb{1}_{\{\tau_\beta = T\}} \,\middle|\, \mathcal{F}_t \right] \\
&\geqslant E_{\mathbb{Q}}\left[e^{-r(\tau_\beta - t)} F_{\tau_\beta}^{(\beta)} (1 - \kappa_{\tau_\beta}) \mathbb{1}_{\{\tau_\beta \in (t,T)\}} \,\middle|\, \mathcal{F}_t \right] + E_{\mathbb{Q}}\left[e^{-r(T-t)} F_T^{(\beta)} \mathbb{1}_{\{\tau_\beta = T\}} \,\middle|\, \mathcal{F}_t \right] \\
&= \beta E_{\mathbb{Q}}\left[e^{-r(\tau_\beta - t)} (1 - \kappa_{\tau_\beta}) \mathbb{1}_{\{\tau_\beta \in (t,T)\}} \,\middle|\, \mathcal{F}_t \right] + E_{\mathbb{Q}}\left[e^{-r(T-t)} F_T^{(\beta)} \mathbb{1}_{\{\tau_\beta = T\}} \,\middle|\, \mathcal{F}_t \right] \\
&> (1 - \kappa_t) \left\{ \beta E_{\mathbb{Q}}\left[e^{-r(\tau_\beta - t)} \mathbb{1}_{\{\tau_\beta \in (t,T)\}} \,\middle|\, \mathcal{F}_t \right] + E_{\mathbb{Q}}\left[e^{-r(T-t)} F_T^{(\beta)} \mathbb{1}_{\{\tau_\beta = T\}} \,\middle|\, \mathcal{F}_t \right] \right\} \qquad (19) \\
&= (1 - \kappa_t) F_t^{(\beta)},
\end{aligned}$$

as the term inside the braces in equation (19) was shown to be exactly $F_t^{(\beta)}$. This result implies that in the presence of surrender charges, it is never optimal to surrender the VA contract when the fund is above or equal to the fee threshold barrier β. □