Hedging Costs for Variable Annuities

A PDE Regime-Switching Approach

by

Parsiad Azimzadeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A general methodology is described in which policyholder behaviour is decoupled from the pricing of a variable annuity based on the cost of hedging it, yielding two sequences of weakly coupled systems of partial differential equations (PDEs): the pricing and utility systems. The utility systems are used to generate policyholder withdrawal behaviour, which is in turn fed into the pricing systems as a means to determine the cost of hedging the contract. This approach allows us to incorporate the effects of utility-based pricing and factors such as taxation. As a case study, we consider the Guaranteed Lifelong Withdrawal and Death Benefits (GLWDB) contract. The pricing and utility systems for the GLWDB are derived under the assumption that the underlying asset follows a Markov regime-switching process. An implicit PDE method is used to solve both systems in tandem. We show that for a large class of utility functions, the two systems preserve homogeneity, allowing us to decrease the dimensionality of solutions. We also show that the associated control for the GLWDB is bang-bang, under which the work required to compute the optimal strategy is significantly reduced. We extend this result to provide the reader with sufficient conditions for a bang-bang control for a general variable annuity with a countable number of events (e.g. discontinuous withdrawals). Homogeneity and bang-bangness yield significant reductions in complexity and allow us to rapidly generate numerical solutions. Results are presented which demonstrate the sensitivity of the hedging expense to various parameters. The costly nature of the death benefit is documented. It is also shown that for a typical contract, the fee required to fund the cost of hedging calculated under the assumption that the policyholder withdraws at the contract rate is an appropriate approximation to the fee calculated assuming optimal consumption.

Acknowledgments

First and foremost, I would like to thank my supervisor, Peter Forsyth, for being an all-around excellent human being, introducing me to the field of computational finance, and enduring me for two years. Another thanks to Ken Vetzal for his work on variable annuities and his time. Thank you to all of the Scientific Computing Laboratory faculty and members who have influenced me and my work: Swathi Amarala, Ken Chan, George Labahn, Yuying Li, Kai Ma, Amir Memartoluie, Anthony Schmieder, Ad Tayal, Stephen Tse, and Justin Wan. Thanks to Matt Amy, Alexandre Laplante and Kyle Robinson for the many good times in Waterloo, to all of my friends from Vancouver, and to my family.

Dedication

To my parents.

Contents

Introduction

Variable annuities are tax-deferred, unit-linked insurance products. These products are a class of insurance vehicles that provide the buyer with particular guarantees without requiring them to sacrifice full control over the funds invested. These funds are often invested in a collective investment vehicle such as a mutual fund. The writer charges a premium that is deducted (over the entirety of the contract) as a proportional amount of the investment account. We term this proportion the *rider*. Note that unlike a vanilla option, this premium is not received up-front.

We propose a method for pricing such contracts when the value of the underlying investment follows a Markovian regime-switching process. Regime-switching was introduced by Hamilton (1989), while its application to long-term guarantees was popularized by Hardy (2001) who demonstrated its effectiveness by fitting to the S&P 500 and the Toronto Stock Exchange 300 indices. Regime switching has thus been suggested as a sensible model for pricing variable annuities (Siu 2005, Lin et al. 2009, Belanger et al. 2009, Yuen and Yang 2010, Ngai and Sherris 2011, Jin et al. 2011) due to their long-term nature. An alternative to this model is stochastic volatility (Hull and White 1987). However, it could be argued that due to the long-term nature of these guarantees, it is more useful to choose a model which allows for incorporation of a long term economic perspective. A regime-switching process has parameters which are economically meaningful, and it is straightforward to adjust these parameters to incorporate economic views. This is perhaps more difficult for a stochastic volatility model, which is typically calibrated to short term option prices. Furthermore, the adoption of stochastic volatility requires an additional dimension in the corresponding partial differential equation (PDE) while the regime-switching model adds complexity proportional to the number of regimes considered, and as a result is computationally less intensive. Moreover, it is straightforward (in the regime-switching framework) to allow for different levels of the risk-free interest rate across regimes. The alternative of incorporating an additional stochastic interest rate factor would add an extra dimension to the PDE, with the associated costs of complexity.

We demonstrate our methodology by considering a specific variable annuity: the Guaranteed Lifelong Withdrawal and Death Benefits (GLWDB) contract. The GLWDB is a response to a general reduction in availability of defined benefit pension plans, allowing the buyer to replicate the security of such a plan via a substitute. The GLWDB is bootstrapped via a lump sum payment to an insurer, $S(0)$, which is invested in risky assets. We term this the *investment account*. Associated with the GLWDB contract are the *guaranteed withdrawal benefit account* and the *guaranteed death benefit account*, hereafter referred to as the withdrawal and death benefits for brevity. We also refer to these as the *auxiliary accounts*. Both auxiliary accounts are initially set to $S(0)$. At a finite set of *withdrawal dates*, the policyholder is entitled to withdraw a predetermined fraction of the withdrawal benefit (or any lesser amount), even if the investment account diminishes to zero. This predetermined fraction is referred to as the *contract withdrawal rate*. If the policyholder wishes to withdraw in excess of the contract withdrawal rate, they can do so upon the payment of a penalty. Typical GLWDB contracts include penalty rates that are decreasing functions of time. Upon death, the policyholder's estate receives the maximum of the investment account and death benefit. These contracts are often bundled with *ratchets* (step-ups), a contract feature that periodically increases one or more of the auxiliary accounts to the investment account, provided that the investment account has grown larger than the respective auxiliary account. Moreover, *bonus* (roll-up) provisions are also often present, in which the withdrawal benefit is increased if the policyholder does not withdraw on a given withdrawal date.

This contract can be considered as part of a greater family of insurance vehicles offering guaranteed benefits that have emerged as a result of a recent trend away from defined benefits (Perkins 2011). Our approach can easily be extended to include features present in an arbitrary member of this family. There exists a maturing body of work on pricing these contracts. Bauer et al. (2008) introduces a general framework for pricing various products in this family. Monte Carlo and numerical integration are employed, and loss-maximizing (from the perspective of the insurer) withdrawal strategies are considered. Holz et al. (2007) compute the rider of Guaranteed Lifelong Withdrawal Benefit (GLWB) contracts via a Monte Carlo method. Milevsky and Salisbury (2006) employ a numerical PDE approach to the Guaranteed Minimum Withdrawal Benefits (GMWB) contract. Shah and Bertsimas (2008) introduce a GLWB model with stochastic volatility and consider static strategies. Kling et al. (2009) provide an extension of the variable annuity model under stochastic volatility. Piscopo and Haberman (2011) consider a model with stochastic mortality risk.

In the general area of financial derivatives, the traditional approach is to assume that the policyholder acts so as to maximize the value of owning the contract. The no-arbitrage price of the contract is then calculated as the cost to the writer of the contract of establishing a self-financing hedging strategy that is guaranteed to produce at least enough cash to pay off any future liabilities resulting from the policyholder's future decisions with respect to the contract (in the context of the assumed pricing model). Since derivative payoffs are a zero sum game, this is equivalent to establishing a price on the basis of assuming a worst case scenario to the contract writer. We will refer to the assumption of such behaviour by policyholders here as *loss-maximizing* strategies, as they represent worst case outcomes for the insurer. Such strategies produce an upper bound on the fair price of the contract, but it is far from clear that policyholders actually behave in this manner. Instead, for any of a number of reasons, a policyholder may deviate from the loss-maximizing strategy.

In order to account for this, we provide a new approach here in which we decouple policyholder withdrawal behaviour from the contract pricing equations, and generate said behaviour by considering a policyholder's utility. This general approach is applicable to any contract involving policyholder behaviour, and results in two sequences of weakly coupled systems of PDEs. In the context of GLWDBs, this allows for the easy modeling of complex phenomena such as risk aversion and taxation. Solving the PDEs backwards in time allows us to employ the Bellman principle to ensure that the policyholder is able to maximize their utility. Since our approach incorporates this added generality, we will generally avoid the use of the term "no-arbitrage" below, and instead refer to the cost of hedging. Of course, under the specific case of loss-maximizing behaviour by the policyholder, our cost of hedging coincides with the traditional no-arbitrage price.

In §1, we introduce a sequence of systems of regime-switching PDEs used to determine the hedging costs of the GLWDB contract. In §2, we introduce a sequence of systems of regime-switching PDEs used to model a policyholder's utility and describe how these systems are used alongside those introduced in §1 to determine the cost of hedging the guarantee assuming optimal consumption. In §3, we discuss theoretical results and our numerical methodology, with a particular focus on various methods used to reduce the computational cost of the pricing procedure. In §4, we present results under both the assumption that the policyholder behaves as to maximize the value of the guarantee (i.e. the loss-maximizing strategy) and the assumption that the policyholder behaves so as to maximize their utility. We summarize our contributions below.

- We model the long term behaviour of the underlying stock index (or mutual fund) by a Markovian regime-switching process.
- We introduce a general methodology that allows for the decoupling of policyholder behaviour from the cost of hedging the contract.
  - This approach yields two sequences of weakly coupled systems of PDEs: the pricing and utility systems.
  - This approach abandons the arguably flawed notion of a policyholder acting only so as to maximize the value of a guarantee.
- We present the pricing and utility systems for the GLWDB contract.
- We show sufficient conditions for the homogeneity of the systems. This result is computationally relevant, as it is used to reduce the dimensionality of the systems.
- In general, the set of possible actions a policyholder can perform at a withdrawal time is (uncountably) infinite. Hence, approximating the optimal action requires a linear search over a discretization of this space. We show sufficient conditions under which a general variable annuity admits a bang-bang strategy (i.e. the set of optimal actions is countable). In particular, for

the GLWDB, we show that the set of optimal choices is finite with cardinality three. Using this fact, we are able to reduce the work per linear search to a constant amount.

- We find that assuming optimal consumption yields a hedging cost fee that is very close to the fee calculated by assuming that the policyholder follows the static strategy of always withdrawing at the contract rate. This is a result of particular practical importance as it suggests that policyholders will generally withdraw at the contract rate. This substantiates pricing contracts under this otherwise seemingly naïve assumption.

- We demonstrate sensitivity to various parameters and we consider the adoption of exotic fee structures in which the proportional fee applies not just to the investment account but rather to the greater of this account and one or more of the auxiliary accounts.

- We find that the inclusion of a death benefit is often expensive. This may account for the failure to properly hedge this guarantee and the subsequent withdrawal of contracts including ratcheting death benefits from the Canadian market.

Chapter 1

Hedging costs

We begin by considering a basic model for pricing GLWDBs under which policyholder withdrawal behaviour is determined solely as to maximize the value of the guarantee (i.e. the loss-maximizing strategy). We extend previous work by Forsyth and Vetzal (2012) via the introduction of a death benefit. For simplicity, we first consider the single-regime case and subsequently extend this model to include regime-switching.

1.1 Pricing PDE

Let $\mathcal{M}(t)$ be defined as the instantaneous rate of mortality per unit interval. That is,

$$\int_t^{t+\Delta t} \mathcal{M}(s)\, ds$$

is the fraction of original owners of the contract who perish in the interval $(t, t+\Delta t)$. The fraction of owners still alive at time t is

$$\mathcal{R}(t) = 1 - \int_0^t \mathcal{M}(s)\, ds,$$

where $t = 0$ is the time at which the contract is purchased. Mortality tables are almost exclusively available in terms of integer ages. In light of this, let

$$\begin{aligned}
x_0 &= \text{insured's age at contract inception} \\
{}_y p_x &= \text{the probability that someone aged } x \text{ will survive to age } x+y \\
{}_y q_x &= 1 - {}_y p_x \\
x_0 + T &= \text{age beyond which survival is assumed impossible}
\end{aligned}$$

We take

$$\mathcal{M}(t) = {}_t p_{x_0} \cdot {}_1 q_{x_0+t},$$

which is piecewise constant over $[0, \infty)$.

Remark 1.1.1 (Diversifiable mortality risk). *We price GLWDB contracts under the assumption that there are a large number of insuree's aged x_0, allowing us to diversify mortality risk. This assumption is ubiquitous in the literature.*

Let $S(t)$ be the amount in the investment account of any policyholder of the GLWDB contract who is still alive at time t. Let $W(t)$ and $D(t)$ be the withdrawal and death benefits at time t. Assume that the underlying value of the investment account is described by

$$dS = (\mu - \alpha) S dt + \sigma S dZ$$

where Z is a Wiener process. The constant α determines the fee structure of the contract. α includes both the deduction required to hedge and manage the contract. This is expressed as

$$\alpha = \alpha_R + \alpha_M$$

where α_R, the *rider*, is the rate at which deductions from the investment account occur to create the premium for the option and α_M is the *management rate*, the rate used to determine the fee charged to manage the contract. If we suppose that α_M is fixed, the pricing problem becomes one of finding α_R such that the insurer's position is hedged. Such a value of α_R is termed the *hedging cost rider*. S tracks the index $\hat{S}$ which follows

$$d\hat{S} = \mu \hat{S} dt + \sigma \hat{S} dZ.$$

It is assumed that the insurer is unable to short S for fiduciary reasons.

Remark 1.1.2 (Fee structure). *Some contracts charge fees as a fraction of the guarantee account balance $W(t)$ or even $\max(S(t), W(t))$ instead. We consider the effect of more exotic fee structures in a subsequent section.*

We proceed by a hedging argument ubiquitous in the literature (Windcliff et al. 2001, Chen et al. 2008, Belanger et al. 2009). Let $U(S, W, D, t)$ be the no-arbitrage value of funding the withdrawal and death benefits at time t years after purchase for investment account value S, withdrawal benefit W, and death benefit D. The value of U is adjusted to account for the effects of mortality. We assume that this contract was purchased at time zero by a buyer aged x_0. Recall that time T corresponds to the time at which all buyers have passed away. The insurer has no obligations at time T and hence

$$U(S, W, D, T) = 0. \tag{1.1.1}$$

The writer creates a replicating portfolio Π by shorting one contract and taking a position of x units in the index $\hat{S}$. That is,

$$\Pi(S, W, D, t) = -U(S, W, D, t) + x\hat{S}(t).$$

The contractually specified times at which withdrawals and ratchets occur are referred to as *event times*, gathered in the set $\mathcal{T} = \{t_1, t_2, \ldots, t_{N-1}\}$ and ordered by

$$0 = t_1 < t_2 < \ldots < t_{N-1} < t_N = T.$$

Note that time zero (but not $t_N = T$) is also referred to as an event time even though it bears no withdrawals or ratchets.

Following standard portfolio dynamics arguments (see, e.g. Forsyth and Vetzal 2012) and noting that between event times dU is a function solely of S and t, we can use Itô's lemma (§A.2) to yield

$$d\Pi = -\left[\left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + (\mu - \alpha) S \frac{\partial U}{\partial S} + \frac{\partial U}{\partial t}\right) dt + \sigma S \frac{\partial U}{\partial S} dZ\right] + x\left[\mu \hat{S} dt + \sigma \hat{S} dZ\right] + \mathcal{R}(t) \alpha_R S dt - \mathcal{M}(t)\left[0 \vee (D - S)\right] dt,$$

where $a \vee b = \max(a, b)$. The term $\mathcal{R}(t) \alpha_R S dt$ represents the fees collected by the hedger, while $\mathcal{M}(t)\left[0 \vee (D - S)\right] dt$ represents the surplus generated by the death benefit as paid out to the estates of deceased policyholders. Taking

$$x = \frac{S}{\hat{S}} \frac{\partial U}{\partial S}$$

yields

$$d\Pi = \left(-\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + \alpha S \frac{\partial U}{\partial S} - \frac{\partial U}{\partial t} + \mathcal{R}(t) \alpha_R S - \mathcal{M}(t)\left[0 \vee (D - S)\right]\right) dt. \tag{1.1.2}$$

As this increment is deterministic, by the principle of no-arbitrage (§A.1), the corresponding portfolio process must grow at the risk-free rate. That is,

$$d\Pi = r\Pi dt = r\left(-U + \frac{S}{\hat{S}} \frac{\partial U}{\partial S} \hat{S}\right) dt. \tag{1.1.3}$$

Substituting (1.1.3) into (1.1.2),

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + (r - \alpha) S \frac{\partial U}{\partial S} + \frac{\partial U}{\partial t} - rU - \mathcal{R}(t) \alpha_R S + \mathcal{M}(t)\left[0 \vee (D - S)\right] = 0. \tag{1.1.4}$$

Let

$$V(S, W, D, t) = U(S, W, D, t) + \mathcal{R}(t) S \tag{1.1.5}$$

be the value of the entire contract at time t. Substituting into (1.1.4), we arrive at

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} (r - \alpha) S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - rV + \mathcal{R}(t) \alpha_M S + \mathcal{M}(t)(S \vee D) = 0. \tag{1.1.6}$$

Remark 1.1.3 (Notation)**.** *Let $\mathbb{R}_{\geqslant 0} = \{x \in \mathbb{R} \mid x \geqslant 0\}$. As to reduce clutter, we abuse notation slightly by writing $V(\mathbf{x}, t)$ with*

$$\mathbf{x} = \langle x_1, x_2, x_3 \rangle \in \mathbb{R}^3_{\geqslant 0}$$

to mean $V(x_1, x_2, x_3, t)$. We use the forms $V(\mathbf{x}, t)$ and $V(S, W, D, t)$ interchangeably in this work.

To summarize, let $\Omega = \mathbb{R}^3_{\geqslant 0}$ and $\Omega_{\mathcal{A}} = \Omega \times \mathcal{A}$ where $\mathcal{A}$ is an arbitrary set. Further let

$$\mathcal{L} = (r - \alpha) S \frac{\partial}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} - r$$

and

$$g(S, W, D, t) = -\mathcal{R}(t) \alpha_M S - \mathcal{M}(t)(S \vee D) \leqslant 0. \tag{1.1.7}$$

Let $(\cdot)$ denote an arbitrary point in Ω. We are interested in functions V solving the sequence of PDEs parameterized by n,

$$\begin{aligned} &\mathcal{L}V + \frac{\partial V}{\partial t} = g \text{ on } \Omega_{(t_n, t_{n+1})} \\ &V(\cdot, t_{n+1}) \text{ given.} \end{aligned} \tag{1.1.8}$$

Note that the boundary conditions at $S = 0$, $W = 0$ and $D = 0$ are obtained by substituting the corresponding value of S, W or D into (1.1.6), as captured by (1.1.8). By (1.1.1) and (1.1.5), when $n = N$, the Cauchy data to the above problem becomes $V(\cdot, t_N = T) = 0$. When $n < N$, the construction of the Cauchy data is outlined in §1.3.

1.2 Regime-switching

We extend the formulation to include a regime-switching framework in which shifts between states are controlled by a continuous-time Markov chain. Letting

$$\mathcal{S} = \{1, 2, \ldots, M\}$$

be the state-space consisting of M regimes, we assume that in regime $i \in \mathcal{S}$, the underlying investment account evolves according to

$$dS = (\mu_i - \alpha) S + \sigma_i S dZ + \sum_{j=1}^{M} S (J_{i \to j} - 1) dX_{i \to j} \tag{1.2.1}$$

where

$$dX_{i \to j} = \begin{cases} 1 & \text{with probability } \delta_{i,j} + q_{i \to j} dt \\ 0 & \text{with probability } 1 - (\delta_{i,j} + q_{i \to j} dt) \end{cases}$$

and

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Here, $q_{i \to j}$ is the objective ($\mathbb{P}$ measure) rate of transition from regime i to j whenever $i \neq j$ and

$$q_{i \to i} = -\sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}.$$

$J_{i\to j} \geqslant 0$ is the jump intensity from regime i to j. We take $J_{i\to i} = 1$ for all i so that jumps in the underlying are not experienced unless there is a change in regime. Let $V_i(S, W, D, t)$ (equivalently, $V_i(\mathbf{x}, t)$) be the no-arbitrage value of holding a GLWDB in regime i at time t years after purchase with investment account value S, withdrawal benefit W and death benefit D. Denote by $\mathbf{V}$ the vector consisting of functions V_1, V_2, …, V_M. Following a combination of the hedging arguments in §1.1 and §B, we arrive at the sequence of systems (parameterized by n),

$$\mathcal{L}_i V_i + \sum_{\substack{j=1 \\ j\neq i}}^{M} \left[q^{\mathbb{Q}}_{i\to j} V_j \left(J_{i\to j} S, W, D, t \right) \right] + \frac{\partial V_i}{\partial t} = g \text{ on } \Omega_{(t_n, t_{n+1})} \ \forall i \in \mathcal{S}$$

$$\mathbf{V}\left(\cdot, t_{n+1}\right) \text{ given} \tag{1.2.2}$$

where

$$\mathcal{L}_i = \frac{1}{2}\sigma_i^2 S^2 \frac{\partial^2}{\partial S^2} + \left(r_i - \alpha - \rho_i^{\mathbb{Q}} \right) S \frac{\partial}{\partial S} - \left(r_i - q^{\mathbb{Q}}_{i\to i} \right). \tag{1.2.3}$$

$q^{\mathbb{Q}}_{i\to j}$ is the risk-neutral rate of transition from regime i to j whenever $i \neq j$ and

$$q^{\mathbb{Q}}_{i\to i} = -\sum_{\substack{j=1 \\ j\neq i}}^{M} q^{\mathbb{Q}}_{i\to j}.$$

Furthermore, $\rho_i^{\mathbb{Q}}$ is defined as

$$\rho_i^{\mathbb{Q}} = \sum_{\substack{j=1 \\ j\neq i}}^{M} \left[q^{\mathbb{Q}}_{i\to j} \left(J_{i\to j} - 1 \right) \right] = \sum_{j=1}^{M} \left[q^{\mathbb{Q}}_{i\to j} J_{i\to j} \right].$$

For each n, (1.2.2) is referred to as a *pricing system*. This system is said to be *weakly coupled* to capture the fact it is coupled only in the terms which are not differentiated.

1.3 Events

At an event time, the policyholder is able to perform one of a number of *actions* based on the status of their contract and the observed regime. The set of all such actions is termed the *control set*, and denoted $\mathcal{C}$. In general, the control set need not be constant.

Across event times, $\mathbf{V}$ is not necessarily continuous as a function of time. We restrict $\mathbf{V}$ to the space of functions that are *càglàd* in time so that the limits

$$\mathbf{V}(\mathbf{x}, t) = \lim_{s\uparrow t} \mathbf{V}(\mathbf{x}, s)$$

and

$$\mathbf{V}\left(\mathbf{x}, t^{+}\right) = \lim_{s\downarrow t} \mathbf{V}(\mathbf{x}, s)$$

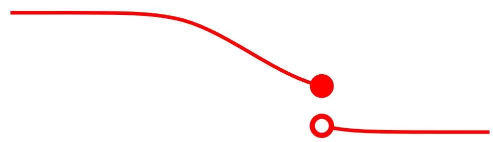


FIGURE 1.3.1: *A one dimensional càglàd function.*

exist, since the policyholder's actions are predictable (see Cont and Tankov (2004) for further discussion). Whenever t is an event time, $\mathbf{V}(\mathbf{x}, t)$ and $\mathbf{V}(\mathbf{x}, t^{+})$ can be regarded as the price of the contract "immediately before" and "immediately after" the event, respectively. An example of a càglàd function is shown in Figure 1.3.1.

1.3.1 Withdrawals

Let $\mathcal{T}_W \subset \mathcal{T}$ be the set of *withdrawal times*. Suppose that the policyholder performs action $\lambda \in \mathcal{C}$ at the point $\langle \mathbf{x}, t, i \rangle \in \Omega \times \mathcal{T}_W \times \mathcal{S}$. We take

- $\lambda = 0$ to indicate that the policyholder does not withdraw anything.
- $\lambda \in (0, 1]$ to indicate a nonzero withdrawal less than or equal to the *contract withdrawal amount*, the maximum amount one can withdraw without incurring a penalty.
- $\lambda \in (1, 2]$ to indicate withdrawal at more than the contract withdrawal amount.

In light of this, we take the control set for the GLWDB problem to be $\mathcal{C} = [0, 2]$. In particular, $\lambda = 2$ is referred to as a *full surrender*, as it corresponds to the scenario in which the policyholder withdraws the entirety of the investment account, while $\lambda \in (1, 2)$ is referred to as a *partial surrender*. We describe a withdrawal by considering the three cases enumerated above separately.

Bonus ($\lambda = 0$)

If the policyholder chooses not to withdraw, the withdrawal benefit is amplified by $1 + B(t)$, where $B(t)$ is the *bonus rate* available at t under regime i. By the principle of no-arbitrage,

$$V_i(S, W, D, t) = V_i\left(S,\ W(1 + B(t)),\ D,\ t^{+}\right). \tag{1.3.1}$$

Withdrawal not exceeding the contract rate ($\lambda \in (0, 1]$)

We assume that the contract withdrawal amount is of the form $G(t) W$, where $G(t)$, the *contract withdrawal rate* at time t, is specified by the contract. We express this type of withdrawal as

$$V_i(S, W, D, t) = V_i\left((S - \lambda G(t) W) \vee 0,\ W,\ (D - \lambda G(t) W) \vee 0,\ t^{+}\right) + \mathcal{R}(t) \lambda G(t) W. \tag{1.3.2}$$

Note that the cash flow $\lambda G(t) W$ is adjusted to account only for the fraction of policyholders still alive at time t, $\mathcal{R}(t)$.

Partial or full surrender ($\lambda \in (1, 2]$)

If the policyholder performs a surrender, the amount withdrawn is

$$w = G(t) W + (\lambda - 1) \kappa(t) S'$$

where

$$S' = (S - G(t) W) \vee 0$$

is the state of the investment account after a withdrawal at the contract withdrawal rate and $(1 - \kappa(t)) \in [0, 1]$ is the *penalty rate* incurred at t for withdrawing above the contract withdrawal rate. For a typical contract, the penalty rate is monotonically decreasing in time. We express this type of withdrawal as

$$V_i(S, W, D, t) = V_i\left((2 - \lambda) S',\ (2 - \lambda) W,\ (2 - \lambda)(D - G(t) W) \vee 0,\ t^+\right) + \mathcal{R}(t) w. \tag{1.3.3}$$

1.3.2 Ratchets

Let $\mathcal{T}_R \subset \mathcal{T}$ be the set of *ratchet times*. At a ratchet time $t \in \mathcal{T}_R$, the withdrawal benefit is increased to the value of the investment account, provided that the investment account has grown larger than the withdrawal benefit. We express this by

$$V_i(S, W, D, t) = V_i\left(S,\ S \vee W,\ D,\ t^+\right). \tag{1.3.4}$$

We also explore the possibility of contracts including a ratcheting death benefit. That is, we sometimes instead consider

$$V_i(S, W, D, t) = V_i\left(S,\ S \vee W,\ S \vee D,\ t^+\right).$$

Note that a ratchet is not controlled by the policyholder, and hence their action at the point $\langle \mathbf{x}, t, i \rangle \in \Omega \times \mathcal{T}_R \times \mathcal{S}$ can be, w.l.o.g., taken to be an arbitrary member of $\mathcal{C}$.

1.3.3 Notation

Although equations (1.3.1), (1.3.2), (1.3.3) and (1.3.4) are an intuitive description of the evolution of a GLWDB contract across event times, we seek a more compact representation in order to simplify subsequent analyses. Suppose the policyholder performs action $\lambda \in \mathcal{C}$ at the point $\langle \mathbf{x}, t, i \rangle \in \Omega \times \mathcal{T} \times \mathcal{S}$. We can parameterize an event occurring at this point by writing it in the form

$$V_i(\mathbf{x}, t) = V_i\left(\mathbf{f}(\mathbf{x}, t, \lambda), t^+\right) + \mathcal{R}(t) f(\mathbf{x}, t, \lambda) \tag{1.3.5}$$

where $\mathbf{f} \colon \Omega \times \mathcal{T} \times \mathcal{C} \to \Omega$ and $f \colon \Omega \times \mathcal{T} \times \mathcal{C} \to \mathbb{R}_{\geqslant 0}$. Here, f represents cash flow from the writer, while $\mathbf{f}$ represents the contract parameters (i.e. the state of the investment account, and withdrawal and death benefits) after a withdrawal. We can summarize (1.3.1), (1.3.2), (1.3.3) and (1.3.4) by letting

$$\mathbf{f}(\mathbf{x},t,\lambda)=\begin{cases}\langle x_1, x_2(1+B(t)), x_3\rangle & \text{if } t\in\mathcal{T}_W,\ \lambda=0\\ \langle (x_1-\lambda G(t)x_2)\vee 0, x_2, (x_3-\lambda G(t)x_2)\vee 0\rangle & \text{if } t\in\mathcal{T}_W,\ \lambda\in(0,1]\\ (2-\lambda)\mathbf{f}(\mathbf{x},t,1) & \text{if } t\in\mathcal{T}_W,\ \lambda\in(1,2]\\ \langle x_1, x_1\vee x_2, x_3\rangle & \text{if } t\in\mathcal{T}_R\\ \mathbf{x} & \text{if } t=t_1=0\end{cases} \tag{1.3.6}$$

and

$$f(\mathbf{x},t,\lambda)=\begin{cases}\lambda G(t)x_2 & \text{if } t\in\mathcal{T}_W,\ \lambda\in(0,1]\\ G(t)x_2+(\lambda-1)\kappa(t)((x_1-G(t)x_2)\vee 0) & \text{if } t\in\mathcal{T}_W,\ \lambda\in(1,2]\ .\\ 0 & \text{otherwise}\end{cases} \tag{1.3.7}$$

Although these functions may seem unwieldly, the form (1.3.5) will prove useful in characterizing properties of the relevant solution in §3. For completeness, (1.3.6) and (1.3.7) include the special case of event time $t_1 = 0$, at which no ratchets or withdrawals are prescribed to occur.

Remark 1.3.1 (Simultaneous events)**.** *We have, up until now, assumed that withdrawals and ratchets occur at separate times. In practice, this is not the case. Naturally, without a particular order, the pricing problem is not well-posed: the contract is ambiguous. If a withdrawal and a ratchet are prescribed to occur, we assume that the withdrawal occurs before the ratchet. As we are solving for the price of the contract backwards in time in order to apply Bellman's principle of optimality, these events are applied in reverse order. For example, when a withdrawal at the contract rate and a ratchet occur at $t \in \mathcal{T}$, we have*

$$V_i(\mathbf{x},t)=V_i\left((x_1-G(t)(x_1\vee x_2))\vee 0, x_1\vee x_2, (x_3-G(t)(x_1\vee x_2))\vee 0, t^+\right)\\ +\mathcal{R}(t)G(t)(x_1\vee x_2)\,.$$

W.l.o.g., we ignore this detail in subsequent analyses. The general theory developed in §3 applies regardless.

1.4 Loss-maximizing strategies

Definition 1.4.1 (Partial strategy)**.** *A partial strategy $\gamma_i^n\colon \Omega \to \mathcal{C}$, describes a policyholder's actions at event time t_n under regime i. That is, for each $\mathbf{x} \in \Omega$, there is a corresponding action $\lambda = \gamma_i^n(\mathbf{x}) \in \mathcal{C}$.*

Definition 1.4.2 (Strategy)**.** *A (full) strategy consists of a partial strategy for each event-regime pair (e.g. $\langle \gamma_1^1, \gamma_1^2, \ldots, \gamma_1^N, \gamma_2^1, \ldots, \gamma_M^N \rangle$).*

Remark 1.4.3 (Abstract strategy)**.** *We stress that we have not yet made any assumptions about policyholder behaviour and chosen instead to abstract it through the notion of a strategy. The decoupling of policyholder behaviour from the pricing equation is the guiding philosophy of this work, and allows us to model complex phenomena directly affecting the policyholder. The robustness of this approach is made concrete via the model developed in §2, which considers the effects of taxation and nonlinear utility functions on policyholder behaviour.*

Definition 1.4.4 (Loss-maximizing partial strategy)**.** *Let*

$$\Gamma_i^n(\mathbf{x}) = \operatorname*{arg\,max}_{\lambda \in [0,2]} \left[V_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right) + \mathcal{R}(t_n) f(\mathbf{x}, t_n, \lambda) \right].$$

A partial strategy γ_i^n is loss-maximizing whenever

$$\gamma_i^n(\mathbf{x}) \in \Gamma_i^n(\mathbf{x}) \ \ \forall \mathbf{x} \in \Omega. \tag{1.4.1}$$

A (full) strategy is said to be loss-maximizing when each of its partial strategies are so too.

If the writer is interested in computing the cost of a contract in the worst-case scenario, the underlying strategy should be taken to be loss-maximizing. Under a loss-maximizing strategy, the policyholder maximizes the cost of hedging the guarantee, thus maximizing the losses of the writer. Using the hedging cost rider computed under this assumption ensures the writer can, at least in theory, hedge a short position in the contract with no risk.

Remark 1.4.5 (An unfortunate choice of terms)**.** *A loss-maximizing strategy is often referred to as an optimal strategy in the literature. The adoption of the term optimal is an arguably unfortunate one, as an optimal strategy is not necessarily "optimal" for the policyholder. We stress that an optimal strategy (as typically referred to in the literature) is simply one that maximizes losses for the writer, and use instead the term "loss-maximizing" in order to avoid confusion.*

Chapter 2

Optimal consumption

Using a loss-maximizing strategy yields the largest hedging cost rider. Any other strategy will, by definition, yield a smaller rider. Using the rider generated by a loss-maximizing strategy ensures the writer can, at least in theory, hedge a short position in the contract with no risk. However, insurers are often interested in using a less conservative method for pricing contracts so as to decrease the hedging cost rider while minimizing their exposure. We now extend the framework introduced in §1 to strategies based on optimal consumption from the perspective of the policyholder. As usual, we first consider the single-regime case and subsequently provide the extension to include regime-switching.

2.1 Utility PDE

Let $\overline{V}(S, W, D, t)$ (equivalently, $\overline{V}(\mathbf{x}, t)$) be the mortality-adjusted utility of holding a GLWDB contract at time t years after purchase with investment account value S, withdrawal benefit W and death benefit D. Recall that the investment account evolves according to

$$dS = (\mu - \alpha) S dt + \sigma S dZ.$$

As usual, between event times, $d\overline{V}$ is a function solely of S and t so that by Itô's lemma,

$$d\overline{V} = \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \overline{V}}{\partial S^2} + (\mu - \alpha) S \frac{\partial \overline{V}}{\partial S} + \frac{\partial \overline{V}}{\partial t} + \mathcal{M}(t) u^B (S \vee D)\right) dt + \sigma S \frac{\partial \overline{V}}{\partial S} dZ \tag{2.1.1}$$

where $u^B(x)$ is the *bequest utility*, the utility received from bequeathing x. The expectation of $d\overline{V}$ is

$$\mathbb{E}\left[d\overline{V}\right] = \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \overline{V}}{\partial S^2} + (\mu - \alpha) S \frac{\partial \overline{V}}{\partial S} + \frac{\partial \overline{V}}{\partial t} + \mathcal{M}(t) u^B (S \vee D)\right) dt.$$

Note that the Wiener term in (2.1.1) vanishes due to the martingale property of the Itô integral. Introducing *rate of time preference* β and assuming that

$$\mathbb{E}\left[d\overline{V}\right] = \beta \overline{V} dt,$$

we express the evolution of a policyholder's utility by

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \overline{V}}{\partial S^2} + (\mu - \alpha) S \frac{\partial \overline{V}}{\partial S} + \frac{\partial \overline{V}}{\partial t} - \beta \overline{V} + \mathcal{M}(t) u^B (S \vee D) = 0. \tag{2.1.2}$$

Note that (2.1.2) depends on the real-world drift μ as opposed to the risk-free rate r. We represent the worthlessness of holding a GLWDB after all death benefits have been paid by

$$\overline{V}(S, W, D, T) = 0. \tag{2.1.3}$$

The drift-diffusion form (2.1.2) is often referred to as *additive utility*, a form of stochastic differential utility (Duffie and Epstein 1992).

Remark 2.1.1 (Kreps-Porteus utility)**.** *We note that an arguably more robust model is that of Kreps-Porteus utility (Kreps and Porteus 1978), another form of stochastic differential utility. Kreps-Porteus utility allows for the decoupling of policyholder substitution preferences and risk-aversion. We leave its incorporation into the GLWDB model to the interested reader.*

2.2 Regime-switching

Assuming that the underlying mutual fund evolves according to the regime-switching process introduced in (1.2.1), we use arguments similar to those in §2.1 to derive a system of expected utility PDEs under regime-switching. It should be noted that in this section, $q_{i \to j}$ denotes the objective ($\mathbb{P}$ measure) rate of transition from regime i to j whenever $i \neq j$ and

$$q_{i \to i} = -\sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}.$$

This should not be confused with the risk-neutral rates of transition denoted $q^{\mathbb{Q}}_{i \to j}$.

Let $\overline{V}_i(S, W, D, t)$ (equivalently, $\overline{V}_i(\mathbf{x}, t)$) be the mortality-adjusted utility of holding a GLWDB contract in regime i at time t years after purchase with investment account value S, withdrawal benefit W and death benefit D. As usual, between event times, the utility is independent of W and D so that

$$d\overline{V}_i = \left(\frac{1}{2}\sigma_i^2 S^2 \frac{\partial^2 \overline{V}_i}{\partial S^2} + (\mu_i - \alpha) S \frac{\partial \overline{V}_i}{\partial S} + \frac{\partial \overline{V}_i}{\partial t} + \mathcal{M}(t) u_i^B (S \vee D) \right) dt + \sigma_i S \frac{\partial \overline{V}_i}{\partial S} dZ + \sum_{j=1}^{M} \Delta \overline{V}_{i \to j} dX_{i \to j}$$

where

$$\Delta \overline{V}_{i \to j} = \overline{V}_j(J_{i \to j} S, t) - \overline{V}_i(S, t).$$

We have used the symbol u_i^B instead of the usual u^B to stress that the bequest utility can, in general, be regime-dependent. The expectation of $d\overline{V}_i$ is

$$\mathbb{E}\left[d\overline{V}_i\right] = \left(\frac{1}{2}\sigma_i^2 S^2 \frac{\partial^2 \overline{V}_i}{\partial S^2} + (\mu_i - \alpha) S \frac{\partial \overline{V}_i}{\partial S} + \frac{\partial \overline{V}_i}{\partial t} \right.$$
$$\left. + \mathcal{M}(t)\, u_i^B (S \vee D) + \sum_{\substack{j=1 \\ j \neq i}}^{M} \Delta \overline{V}_{i \to j} q_{i \to j} \right) dt.$$

Introducing rate of time preference β_i and assuming that

$$\mathbb{E}\left[d\overline{V}_i\right] = \beta_i \overline{V}_i dt,$$

yields

$$0 = \frac{1}{2}\sigma_i^2 S^2 \frac{\partial^2 \overline{V}_i}{\partial S^2} + (\mu_i - \alpha) S \frac{\partial \overline{V}_i}{\partial S} + \frac{\partial \overline{V}_i}{\partial t} - (\beta_i - q_{i \to i}) \overline{V}_i$$
$$+ \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[q_{i \to j} \overline{V}_j (J_{i \to j} S, W, D, t)\right] + \mathcal{M}(t)\, u_i^B (S \vee D).$$

In light of this, let

$$\overline{\mathcal{L}}_i = \frac{1}{2}\sigma_i^2 S^2 \frac{\partial^2}{\partial S^2} + (\mu_i - \alpha) S \frac{\partial}{\partial S} - (\beta_i - q_{i \to i})$$

and

$$\overline{g}_i (S, W, D, t) = -\mathcal{M}(t)\, u_i^B (S \vee D). \tag{2.2.1}$$

We are interested in solutions to the sequence of systems (parameterized by n)

$$\frac{\partial \overline{V}_i}{\partial t} + \overline{\mathcal{L}}_i \overline{V}_i + \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[q_{i \to j} \overline{V}_j (J_{i \to j} S, W, D, t)\right] = \overline{g}_i \text{ on } \Omega_{(t_n, t_{n+1})}\ \forall i \in \mathcal{S}$$
$$\overline{\mathbf{V}}(\cdot, t_{n+1}) \text{ given.} \tag{2.2.2}$$

For each n, (2.2.2) is referred to as a *utility system*.

2.3 Events

Suppose the policyholder performs action $\lambda \in \mathcal{C}$ at the point $\langle \mathbf{x}, t, i \rangle \in \Omega \times \mathcal{T} \times \mathcal{S}$. As before, we parameterize an event occurring at $t \in \mathcal{T}$ by writing it in the form

$$\overline{V}_i (\mathbf{x}, t) = \overline{V}_i \left(\mathbf{f}(\mathbf{x}, t, \lambda), t^+\right) + \mathcal{R}(t)\, u_i^C (f(\mathbf{x}, t, \lambda)) \tag{2.3.1}$$

where $\mathbf{f}$ and f are defined by (1.3.6) and (1.3.7). Although similar to (1.3.5), (2.3.1) differs in that cash flows are transformed by the *consumption utility*, u_i^C. Specifically, $u_i^C(x)$ is the utility gained from consuming x units of the numéraire under regime i.

2.4 Consumption-optimal strategies

Definition 2.4.1 (Consumption-optimal partial strategy)**.** *Let*

$$\overline{\Gamma}_i^n(\mathbf{x}) = \underset{\lambda\in[0,2]}{\arg\max}\left[\overline{V}_i\left(\mathbf{f}(\mathbf{x},t_n,\lambda),t_n^+\right) + \mathcal{R}(t_n)\,u_i^C(f(\mathbf{x},t_n,\lambda))\right]. \tag{2.4.1}$$

A partial strategy γ_i^n is consumption-optimal whenever

$$\gamma_i^n(\mathbf{x}) \in \overline{\Gamma}_i^n(\mathbf{x}) \;\; \forall \mathbf{x}\in\Omega. \tag{2.4.2}$$

A (full) strategy is said to be consumption-optimal when each of its partial strategies are so too.

Unlike a loss-maximizing strategy, the policyholder chooses λ in (2.4.1) to maximize their utility. It should be noted that we are not interested in the value of the numerical solution to the utility system but rather in the behaviour generated by it. Instead of adopting a loss-maximizing strategy as introduced in §1.4, we "feed" a strategy generated by the policyholder's utility into the pricing system. Given Cauchy data at time t_{n+1},

1. Solve $\mathbf{V}(\cdot,t_n^+)$ using (1.2.2) and Cauchy data $\mathbf{V}(\cdot,t_{n+1})$.
2. Solve $\overline{\mathbf{V}}(\cdot,t_n^+)$ using (2.2.2) and Cauchy data $\overline{\mathbf{V}}(\cdot,t_{n+1})$.
3. For each regime i,
   (a) Determine a consumption-optimal partial strategy γ_i^n using $\overline{V}_i(\cdot,t_n^+)$ and Definition 2.4.1. In doing so, determine $\overline{V}_i(\cdot,t_n)$ as in (2.3.1).
   (b) Use γ_i^n and $V_i(\cdot,t_n^+)$ to determine $V_i(\cdot,t_n)$ as in (1.3.5).

The propagation of information in this procedure is depicted in Figure 2.4.1.

Remark 2.4.2 (Ensuring uniqueness)**.** *Between event times, we find solutions to the systems (1.2.2) and (2.2.2) as described in steps 1 and 2. It is well-known that without additional assumptions on the rate of growth of solutions, classical solutions to parabolic systems are not necessarily unique (Friedman 1964). The term "solve" in these steps should hence be read "solve with sufficient regularity to ensure uniqueness."*

At an event time, step 3a requires determining a consumption-optimal partial strategy γ_i^n. The expression (2.4.2) suggests that this strategy need not be unique, prompting us to seek a way to break ties between consumption-optimal partial strategies. We take $c\colon 2^{\mathcal{C}} \to \mathcal{C}$ to be any ($\mathcal{F}(t_n)$-measurable; see §A.1) choice function on the power set of $\mathcal{C}$, $2^{\mathcal{C}}$. If condition (2.4.2) is substituted for

$$\gamma_i^n(\mathbf{x}) = c\left(\overline{\Gamma}_i^n(\mathbf{x})\right) \;\; \forall \mathbf{x}\in\Omega,$$

it is easy to show that uniqueness is maintained. Otherwise, a solution to the price of the contract need not be unique.

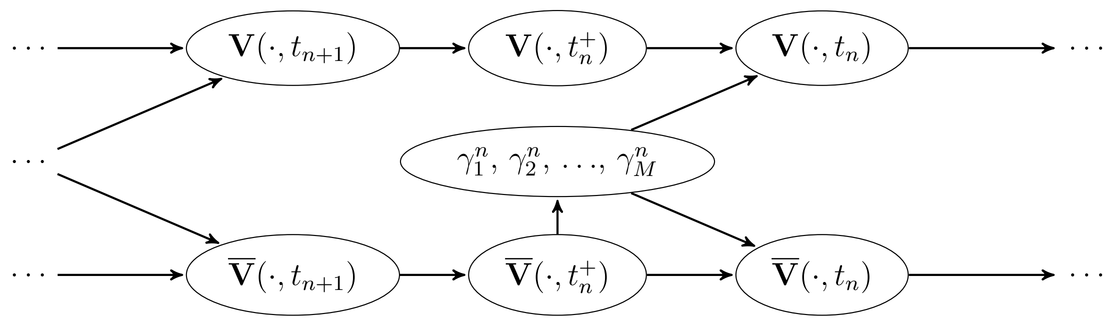


FIGURE 2.4.1: *A graph depicting the propagation of information in the pricing procedure.*

Intuitively, the introduction of a choice function corresponds to a policyholder preferring one action over another, even though the observed utility of both actions are equivalent. For example, a choice function c that selects the smallest element (e.g. $c(\{0,1,2\}) = 0$) corresponds to a policyholder that is "marginally surrender-averse".

2.5 Utility functions

The model introduced above is parameterized by our particular choices for the *utility functions* u_i^B and u_i^C. For any utility function $u : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$, we require that u be twice-differentiable on $(0, \infty)$ with $u'(x) > 0$ for all x. The classical notion of equivalence states that two functions u_1 and u_2 are equivalent if there exist a and $b > 0$ such that $u_1(x) = a + bu_2(x)$ for all x (i.e. equivalent under affine transformations). We write $u_1 \sim u_2$ to denote this equivalence.

Given a particular utility function u, the *Arrow-Pratt measure of absolute risk-aversion* (Arrow 1971, Pratt 1964) is defined to be

$$A(x; u) = -\frac{u''(x)}{u'(x)}.$$

The *Arrow-Pratt measure of relative risk-aversion* is defined to be

$$R(x; u) = xA(x; u).$$

Naturally, for any two utility functions u_1, u_2 with $u_1 \sim u_2$, $A(x; u_1) = A(x; u_2)$ for all x. Although the concavity of a utility function does not alone fully characterize its behaviour, the sign of $u''(x)$ does hold some meaning. Concavity (convexity) captures a policyholder's unwillingness (willingness) to accept small, actuarially neutral risks.

Several notable cases of utility functions are special cases of the general family of *hyperbolic absolute risk-aversion* (HARA) functions (Merton 1970). For ease of exposition, we present these notable cases separately before summarizing with HARA preferences.

2.5.1 Actuarially risk-neutral utility

We refer to a utility function u as (actuarially) *risk-neutral* (RN) when $u'(x)$ is constant for all x. In particular, when modeling risk-neutral utility, we often consider

$$u_{\mathrm{RN}}(x) = x.$$

2.5.2 Constant absolute risk-aversion

We refer to a utility function u as exhibiting *constant absolute risk-aversion* (CARA) whenever $A(x; u)$ is constant. In particular, when modeling CARA preferences, we often consider

$$u_{\mathrm{CARA}}(x; a) = 1 - e^{-ax}$$

for $a > 0$ with $A(x; u_{\mathrm{CARA}}) = a$.

2.5.3 Constant relative risk-aversion

We refer to a utility function u as exhibiting *constant relative risk-aversion* (CRRA) whenever $R(x; u)$ is constant. In particular, when modeling CRRA preferences, we often consider

$$(1-p)\frac{\left(\frac{ax}{1-p}\right)^p - 1}{p}, \tag{2.5.1}$$

for $a > 0$ and $p < 1$. Setting $a = 1 - p$ yields the more familiar form

$$\left[(1-p)\frac{x^p - 1}{p}\right] \sim \left[\frac{x^p - 1}{p}\right] \sim \left[\frac{x^p}{p}\right],$$

often termed the *power utility function* or *power law utility*. It should be noted that some authors also refer to the form (2.5.1) by this name. Note that

$$\lim_{p\to 0}(1-p)\frac{\left(\frac{ax}{1-p}\right)^p - 1}{p} = \ln(ax),$$

yielding *logarithmic utility*. We summarize this by defining

$$u_{\mathrm{CRRA}}(x; a, p) = \begin{cases} (1-p)\frac{\left(\frac{ax}{1-p}\right)^p - 1}{p} & \text{if } p \in (-\infty, 0) \cup (0, 1) \\ \ln(ax) & \text{if } p = 0 \end{cases}$$

with $R(x; u_{\mathrm{CRRA}}) = 1 - p$.

a	b	p	$\lim u_{\underline{\text{HARA}}}$	$\lim A_{\underline{\text{HARA}}}$	$\lim R_{\underline{\text{HARA}}}$	Qualitative Name
	1	$-\infty$	$1-e^{-ax}$	a	ax	Exponential
$1-p$	0	$(-\infty,0)$ $\cup(0,1)$	$(1-p)\,\frac{x^p-1}{p}$	$(1-p)\,/x$	$1-p$	Power Law
1	0	0	$\ln(x)$	$1/x$	1	Logarithmic
		1	ax	0	0	Risk-Neutral

TABLE 2.5.1: *Notable parameterizations of the HARA family with their corresponding measures of absolute and relative risk-aversion. We use the symbol "*$\lim$*" to denote that some of these quantities are acquired in the limit of particular values of* p *(we assume the limit w.r.t.* p *is taken after substitutions in* a *and* b *are made).*

2.5.4 Hyperbolic absolute risk-aversion

We refer to a utility function u as exhibiting *hyperbolic absolute risk-aversion* (HARA) whenever the measure of absolute risk-aversion has the form

$$A(x;u)=\left(\frac{x}{1-p}+\frac{b}{a}\right)^{-1}.$$

All functions exhibiting this risk-aversion can be expressed as

$$u_{\underline{\text{HARA}}}(x;a,b,p)=\frac{1-p}{p}\left(\left(\frac{ax}{1-p}+b\right)^{p}-1\right)$$

up to equivalence. This is a fairly flexible and general class of utility functions that can be parameterized so that marginal utility is finite at a consumption level of zero. This is potentially of interest in our context since it allows for the possibility that the policyholder will decide to not withdraw any amount at a withdrawal date. Otherwise, with infinite marginal utility at a consumption level of zero, the policyholder will always withdraw some positive amount.

We require $a>0$. $A(x;u)>0$ along with $a>0$ produces the requirement

$$\frac{ax}{1-p}+b>0.$$

Table 2.5.1 lists various notable parameterizations of the HARA family.

Let

$$u_{\text{HARA}}(x;a,b,p)=\frac{1-p}{p}\left(\frac{ax}{1-p}+b\right)^{p}$$

and note that $u_{\text{HARA}}\sim u_{\underline{\text{HARA}}}$ as long as $p\neq 0$ (equivalence for $p=1$ is understood in the limiting sense). We have chosen to first introduce the reader to the arguably more convenient form $u_{\underline{\text{HARA}}}$ as the limiting case for $p\to 0$ yields the familiar logarithmic utility (Table 2.5.1), whereas u_{HARA} does not have a two-sided limit as $p \to 0$. The form u_{HARA} is, however, more common in the literature. For the remainder of this work, we consider

$$u_i^C(x) = u_{\mathrm{HARA}}(x; a_i, b_i, p_i)$$

for all regimes $i \in \mathcal{S}$.

2.5.5 Proportional consumption on death

We refer to taking a policyholder's bequest utility as a proportion of the utility gained from consumption as *proportional consumption on death*. For the remainder of this work, we consider

$$u_i^B(x) = h_i u_i^C(x)$$

for all regimes $i \in \mathcal{S}$. Here, h_i is used to amplify (or attenuate) utility gained from bequeathing, and is hence termed the policyholder's *bequest motive*.

Chapter 3

Theory and numerical considerations

We outline our assumptions for this chapter below.

Definition 3.0.1 (Discontinuity of $\mathcal{M}$)**.** *Define $\mathcal{M}^\epsilon$ for $\epsilon > 0$ to be continuous with $\mathcal{M}^\epsilon \to \mathcal{M}$ as $\epsilon \to 0$. We define g^ϵ as the function resulting from the substitution of $\mathcal{M}^\epsilon$ for $\mathcal{M}$ appearing in g (1.1.7). That is,*

$$g^\epsilon(\mathbf{x}, t) = -\mathcal{R}(t)\,\alpha_M x_1 - \mathcal{M}^\epsilon(t)(x_1 \vee x_3).$$

To simplify analysis in ensuring the existence and uniqueness of classical solutions to the pricing problem, we substitute g^ϵ for g. We define $\overline{g}_i^\epsilon$ similarly, and substitute $\overline{g}_i^\epsilon$ for $\overline{g}_i$ (2.2.1). Specifically,

$$\overline{g}_i^\epsilon(\mathbf{x}, t) = -\mathcal{M}^\epsilon(t)\, u_i^B (x_1 \vee x_3).$$

Assumption 3.0.2 (Bounded and continuous coefficients)**.** *$\sigma_i^2 > 0$, r_i, α, $\rho_i^{\mathbb{Q}}$, $q_{i\to j}^{\mathbb{Q}}$, μ_i, β_i, $q_{i\to j}$ are bounded, continuous functions of t (independent of S, W and D).*

Assumption 3.0.3 (Unit jumps)**.** *We restrict our attention to the case of a regime-switching model with unit jumps (i.e. $J_{i\to j} = 1$).*

Remark 3.0.4. *The assumption of unit jumps is a fairly strong one; we believe that all of the results below hold under much weaker assumptions. The purpose of this assumption is primarily to allow the use of existing work in the development of the Green's function for use between event times (namely, by the parametrix method of Levi (1907)).*

Assumption 3.0.5 (Classical solutions)**.** *$\mathbf{V}$ and $\overline{\mathbf{V}}$ are twice differentiable in the investment account S, and once in time t, except possibly at any time $t \in \mathcal{T}$.*

Assumption 3.0.6 (Bounded solutions)**.** *For all $W, D \in \mathbb{R}_{\geqslant 0}$ there exist positive constants L and ℓ s.t.*

$$|\mathbf{V}(e^x, W, D, t)| \leqslant L e^{\ell x^2}$$

for all times $t \in (0, T]$ ($|\cdot|$ denotes any L^p norm). This bound is used to ensure uniqueness. We assume an identical bound on $\overline{\mathbf{V}}$.

We refer to the equations (1.3.5), (1.3.6), (1.3.7) and (2.3.1) extensively in this chapter; they are repeated below for convenience.

$$V_i(\mathbf{x}, t) = V_i\left(\mathbf{f}(\mathbf{x}, t, \lambda), t^+\right) + \mathcal{R}(t) f(\mathbf{x}, t, \lambda) \tag{1.3.5}$$

$$\mathbf{f}(\mathbf{x}, t, \lambda) = \begin{cases} \langle x_1, x_2(1 + B(t)), x_3 \rangle & \text{if } t \in \mathcal{T}_W,\ \lambda = 0 \\ \langle (x_1 - \lambda G(t) x_2) \vee 0, x_2, (x_3 - \lambda G(t) x_2) \vee 0 \rangle & \text{if } t \in \mathcal{T}_W,\ \lambda \in (0, 1] \\ (2 - \lambda) \mathbf{f}(\mathbf{x}, t, 1) & \text{if } t \in \mathcal{T}_W,\ \lambda \in (1, 2] \\ \langle x_1, x_1 \vee x_2, x_3 \rangle & \text{if } t \in \mathcal{T}_R \\ \mathbf{x} & \text{if } t = t_1 = 0 \end{cases} \tag{1.3.6}$$

$$f(\mathbf{x}, t, \lambda) = \begin{cases} \lambda G(t) x_2 & \text{if } t \in \mathcal{T}_W,\ \lambda \in (0, 1] \\ G(t) x_2 + (\lambda - 1) \kappa(t) ((x_1 - G(t) x_2) \vee 0) & \text{if } t \in \mathcal{T}_W,\ \lambda \in (1, 2] \\ 0 & \text{otherwise} \end{cases} \tag{1.3.7}$$

$$\overline{V}_i(\mathbf{x}, t) = \overline{V}_i\left(\mathbf{f}(\mathbf{x}, t, \lambda), t^+\right) + \mathcal{R}(t) u_i^C(f(\mathbf{x}, t, \lambda)) \tag{2.3.1}$$

3.1 Homogeneity

Definition 3.1.1 (Cone)**.** *A cone is a subset of a vector space closed under multiplication by positive scalars (e.g. Ω). Note that the positivity requirement suggests that this definition is sensible only when the corresponding scalar field is ordered.*

Definition 3.1.2 (Homogeneous function)**.** *A function $s : X \to Y$ between two cones is said to be homogeneous of order $k \in \mathbb{Z}$ if for all $\eta > 0$ and $\mathbf{x} \in X$,*

$$\eta^k s(\mathbf{x}) = s(\eta \mathbf{x}).$$

We say $\mathbf{V}$ is homogeneous if for each $i \in \mathcal{S}$, V_i is homogeneous.

Theorem 3.1.3 (Price homogeneity under loss-maximizing strategy)**.** *Suppose that a loss-maximizing (full) strategy is employed by the policyholder. Then, $\mathbf{V}(\mathbf{x}, t)$ is homogeneous of order 1 in $\mathbf{x}$.*

This fact is established via a series of lemmas. Namely, we show that if $\mathbf{V}(\mathbf{x}, t_{n+1})$ is homogeneous in $\mathbf{x}$, so too is $\mathbf{V}(\mathbf{x}, t_n^+)$ (Lemma 3.1.4). That is, the system composed of the operators $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_M$ preserves homogeneity. Then, we show that if $\mathbf{V}(\mathbf{x}, t_n^+)$ is homogeneous in $\mathbf{x}$, so too is $\mathbf{V}(\mathbf{x}, t_n)$ (Lemma 3.1.5). That is, homogeneity is preserved across event times under a loss-maximizing strategy. Noting that $\mathbf{V}(\mathbf{x}, t_N = T) = \mathbf{0}$ is trivially homogeneous, the desired result follows by induction.

Lemma 3.1.4 (Pricing system homogeneity between event times)**.** *Suppose that for some n with $1 \leqslant n < N$, $\mathbf{V}(\mathbf{x}, t_{n+1})$ is homogeneous of order 1 in $\mathbf{x}$. Then, for all $t \in (t_n, t_{n+1}]$, $\mathbf{V}(\mathbf{x}, t)$ is homogeneous of order 1 in $\mathbf{x}$.*

Proof. Let $\mathbf{u}(y, \tau) = \mathbf{V}(e^y, W, D, t_{n+1} - \tau)$ and $\Delta = t_{n+1} - t_n$. By mapping $S \to 0$ to $y \to -\infty$, we are able to consider operators that are uniformly elliptic on $\mathbb{R}$ instead of those defined by (1.2.3), which are degenerate on the boundary $S = 0 \in \mathbb{R}_{\geqslant 0}$. Specifically, we have that $\mathbf{u}$ satisfies

$$\mathcal{L}_i' u_i + \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[q_{i \to j}^{\mathbb{Q}} u_j \right] - \frac{\partial u_i}{\partial \tau} = g^{\epsilon}\left(e^y, W, D, t_{n+1} - \tau\right) \text{ on } \mathbb{R} \times (0, \Delta) \ \ \forall i \in \mathcal{S}, \tag{3.1.1}$$

where each $\mathcal{L}_i'$ is uniformly elliptic. Denote by $\mathbf{1}$ an $M \times 1$ vector of ones. Then, Assumption 3.0.6 allows us to write $\mathbf{u}$ as

$$\begin{aligned}\mathbf{u}(y, \tau) = &\int_{\mathbb{R}} F(y' - y, \tau, 0)\, \mathbf{u}(y', 0)\, dy' \\ &- \int_0^{\Delta} \int_{\mathbb{R}} F(y' - y, \tau, \tau') \left(g^{\epsilon}\left(e^{y'}, W, D, t_{n+1} - \tau'\right) \mathbf{1} \right) dy' d\tau',\end{aligned}$$

where F is the Green's function depending on y and y' through $y' - y$ alone (Friedman 1964). Hence, for $S > 0$,

$$\begin{aligned}\mathbf{V}(S, W, D, t) = &\int_{\mathbb{R}} F(y' - \log S, \tau, 0)\, \mathbf{V}\left(e^{y'}, W, D, t_{n+1}\right) dy' \\ &- \int_0^{\Delta} \int_{\mathbb{R}} F(y' - \log S, \tau, \tau') \left(g^{\epsilon}\left(e^{y'}, W, D, t_{n+1} - \tau'\right) \mathbf{1} \right) dy' d\tau'.\end{aligned}$$

The substitution $y' = \log(SS')$ yields

$$\begin{aligned}\mathbf{V}(S, W, D, t) = &\int_0^{\infty} F(\log S', \tau, 0)\, \mathbf{V}(SS', W, D, t_{n+1}) \frac{1}{S'} dS' \\ &- \int_0^{\Delta} \int_0^{\infty} F(\log S', \tau, \tau') \left(g^{\epsilon}(SS', W, D, t_{n+1} - \tau') \mathbf{1} \right) \frac{1}{S'} dS' d\tau'.\end{aligned} \tag{3.1.2}$$

Since $\mathbf{V}(\mathbf{x}, t_{n+1})$ and $g^{\epsilon}(\mathbf{x}, t_{n+1} - \tau')$ are homogeneous of degree 1 in $\mathbf{x}$, the homogeneity of $\mathbf{V}(\mathbf{x}, t)$ in $\mathbf{x}$ on $(\mathbb{R}_{\geqslant 0} \setminus \partial\mathbb{R}_{\geqslant 0}) \times \mathbb{R}_{\geqslant 0}^2 \times (t_n, t_{n+1}]$ follows. By the presumed continuity of $\mathbf{V}$, we can extend this to $\Omega_{(t_n, t_{n+1}]}$. □

Lemma 3.1.5 (Loss-maximizing partial strategy preserves homogeneity)**.** *Suppose that for some regime $i \in \mathcal{S}$ and for some n with $1 \leqslant n < N$, $V_i(\mathbf{x}, t_n^+)$ is homogeneous of order 1 in $\mathbf{x}$ and that the policyholder employs a loss-maximizing partial strategy γ_i^n. Then, $V_i(\mathbf{x}, t_n)$ is homogeneous of order 1 in $\mathbf{x}$.*

Proof. First, note that $\mathbf{f}$ and f ((1.3.6) and (1.3.7)) are homogeneous of order 1 in $\mathbf{x}$. Let $\eta > 0$. By Definition 1.4.4, and the homogeneity of $V_i\left(\mathbf{x}, t_n^+\right)$, $\mathbf{f}$ and f,

$$
\begin{aligned}
\gamma_i^n(\mathbf{x}) &\in \Gamma_i^n(\mathbf{x}) \\
&= \arg\max_{\lambda \in \mathcal{C}} \quad \left[V_i\left(\mathbf{f}\left(\ \mathbf{x}, t_n, \lambda\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\ \mathbf{x}, t_n, \lambda\right)\right] \\
&= \arg\max_{\lambda \in \mathcal{C}} \eta \left[V_i\left(\mathbf{f}\left(\ \mathbf{x}, t_n, \lambda\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\ \mathbf{x}, t_n, \lambda\right)\right] \\
&= \arg\max_{\lambda \in \mathcal{C}} \quad \left[V_i\left(\mathbf{f}\left(\eta \mathbf{x}, t_n, \lambda\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\eta \mathbf{x}, t_n, \lambda\right)\right] \\
&= \Gamma_i^n(\eta \mathbf{x}) \ni \gamma_i^n(\eta \mathbf{x}). \qquad (3.1.3)
\end{aligned}
$$

Now, using (1.3.5), (3.1.3), and the homogeneity of $V_i\left(\mathbf{x}, t_n^+\right)$, $\mathbf{f}$ and f,

$$
\begin{aligned}
\eta V_i\left(\mathbf{x}, t_n\right) &= \eta\left[V_i\left(\mathbf{f}\left(\ \mathbf{x}, t_n, \gamma_i^n\left(\ \mathbf{x}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\ \mathbf{x}, t_n, \gamma_i^n\left(\ \mathbf{x}\right)\right)\right] \\
&= \quad V_i\left(\mathbf{f}\left(\eta \mathbf{x}, t_n, \gamma_i^n\left(\ \mathbf{x}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\eta \mathbf{x}, t_n, \gamma_i^n\left(\ \mathbf{x}\right)\right) \\
&= \quad V_i\left(\mathbf{f}\left(\eta \mathbf{x}, t_n, \gamma_i^n\left(\eta \mathbf{x}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\eta \mathbf{x}, t_n, \gamma_i^n\left(\eta \mathbf{x}\right)\right) \\
&= \quad V_i\left(\eta \mathbf{x}, t_n\right).
\end{aligned}
$$

□

The homogeneity of the price under a loss-maximizing strategy allows us to reduce the dimensionality of the problem. By Theorem 3.1.3,

$$
\mathbf{V}(S, W, D, t) = \frac{1}{\eta} \mathbf{V}(\eta S, \eta W, \eta D, t).
$$

Suppose $W \neq 0$. Choosing $\eta = \frac{W^\star}{W}$ with $W^\star \neq 0$ yields

$$
\mathbf{V}(S, W, D, t) = \frac{W}{W^\star} \mathbf{V}\left(\frac{W^\star}{W} S, W^\star, \frac{W^\star}{W} D, t\right), \qquad (3.1.4)
$$

which reveals that we need only solve the problem for two values of the withdrawal benefit: $W^\star$ and zero. We refer to this reduction in dimensionality as a *similarity reduction*.

Assumption 3.1.6 (Choice function)**.** *Throughout the rest of* §*3.1, we assume that the choice function* $c \colon 2^{\mathcal{C}} \to \mathcal{C}$ *is used to break ties in the policyholder's withdrawal strategy (Remark 2.4.2).*

Theorem 3.1.7 (Utility homogeneity under consumption-optimal strategy)**.** *Suppose that a consumption-optimal (full) strategy is employed by the policyholder, and that for all regimes* $i \in \mathcal{S}$, u_i^B *and* u_i^C *are homogeneous of order* p*. Then,* $\mathbf{V}(\mathbf{x}, t)$ $\overline{\mathbf{V}}(\mathbf{x}, t)$ *are homogeneous of orders* 1 *and* p*, respectively, in* $\mathbf{x}$*.*

As with Theorem 3.1.3, this fact is established via a series of lemmas. Recall that if $\mathbf{V}(\mathbf{x}, t_{n+1})$ is homogeneous of order 1 in $\mathbf{x}$, so too is $\mathbf{V}\left(\mathbf{x}, t_n^+\right)$ (Lemma 3.1.4). Similarly, we establish that if $\overline{\mathbf{V}}(\mathbf{x}, t_{n+1})$ is homogeneous of order p in $\mathbf{x}$, so too is $\overline{\mathbf{V}}(\mathbf{x}, t_n^+)$ (Lemma 3.1.8). Then, we show that if $\mathbf{V}(\mathbf{x}, t_n^+)$ and $\overline{\mathbf{V}}(\mathbf{x}, t_n^+)$ are homogeneous of orders 1 and p, respectively, in $\mathbf{x}$, so too are $\mathbf{V}(\mathbf{x}, t_n)$ and $\overline{\mathbf{V}}(\mathbf{x}, t_n)$ (Lemma 3.1.9). Noting that $\mathbf{V}(\mathbf{x}, t_N = T) = \overline{\mathbf{V}}(\mathbf{x}, t_N = T) = \mathbf{0}$ are trivially homogeneous, the desired result follows by induction.

Lemma 3.1.8 (Utility system homogeneity between event times)**.** *Suppose that for some p and n with $1 \leqslant n < N$, $\overline{\mathbf{V}}(\mathbf{x}, t_{n+1})$ is homogeneous of order p in $\mathbf{x}$ and u_i^B is homogeneous of order p. Then, for all $t \in (t_n, t_{n+1}]$, $\overline{\mathbf{V}}(\mathbf{x}, t)$ is homogeneous of order p in $\mathbf{x}$.*

Proof sketch. The proof of this is almost identical to that of Lemma 3.1.4. Constructing the analogue to (3.1.2) for the utility system and noting that $\overline{\mathbf{V}}(\mathbf{x}, t_{n+1})$ and $\overline{g}_i^\epsilon(\mathbf{x}, t_{n+1} - \tau')$ are homogeneous of order p in $\mathbf{x}$, the desired result follows.

Lemma 3.1.9 (Consumption-optimal partial strategy preserves homogeneity)**.** *Suppose that for some regime $i \in \mathcal{S}$ and for some n with $1 \leqslant n < N$, $V_i(\mathbf{x}, t_n^+)$ and $\overline{V}_i(\mathbf{x}, t_n^+)$ are homogeneous of orders 1 and p, respectively, in $\mathbf{x}$. Further suppose that u_i^C is homogeneous of order p and that the policyholder employs a consumption-optimal partial strategy γ_i^n. Then, $V_i(\mathbf{x}, t_n)$ and $\overline{V}_i(\mathbf{x}, t_n)$ are homogeneous of orders 1 and p, respectively, in $\mathbf{x}$.*

Proof. Recall that $\mathbf{f}$ and f ((1.3.6) and (1.3.7)) are homogeneous of order 1 in $\mathbf{x}$. Let $\eta > 0$. Using choice function c (Remark 2.4.2), and the homogeneity of $\overline{V}_i(\mathbf{x}, t_n^+)$ and u_i^C,

$$\begin{aligned}
\gamma_i^n(\mathbf{x}) &= c\left(\overline{\Gamma}_i^n(\mathbf{x})\right) \\
&= c\left(\operatorname*{arg\,max}_{\lambda \in \mathcal{C}} \left[\overline{V}_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right) + \mathcal{R}(t_n)\, u_i^C\left(f(\mathbf{x}, t_n, \lambda)\right)\right]\right) \\
&= c\left(\operatorname*{arg\,max}_{\lambda \in \mathcal{C}} \eta^p \left[\overline{V}_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right) + \mathcal{R}(t_n)\, u_i^C\left(f(\mathbf{x}, t_n, \lambda)\right)\right]\right) \\
&= c\left(\operatorname*{arg\,max}_{\lambda \in \mathcal{C}} \left[\overline{V}_i\left(\mathbf{f}(\eta\mathbf{x}, t_n, \lambda), t_n^+\right) + \mathcal{R}(t_n)\, u_i^C\left(f(\eta\mathbf{x}, t_n, \lambda)\right)\right]\right) \\
&= c\left(\overline{\Gamma}_i^n(\eta\mathbf{x})\right) = \gamma_i^n(\eta\mathbf{x}).
\end{aligned}$$

Now, using (1.3.5) and the homogeneity of $V_i(\mathbf{x}, t_n^+)$,

$$\begin{aligned}
\eta V_i(\mathbf{x}, t_n) &= \eta\left[V_i\left(\mathbf{f}(\mathbf{x}, t_n, \gamma_i^n(\mathbf{x})), t_n^+\right) + \mathcal{R}(t_n)\, f(\mathbf{x}, t_n, \gamma_i^n(\mathbf{x}))\right] \\
&= V_i\left(\mathbf{f}(\eta\mathbf{x}, t_n, \gamma_i^n(\eta\mathbf{x})), t_n^+\right) + \mathcal{R}(t_n)\, f(\eta\mathbf{x}, t_n, \gamma_i^n(\eta\mathbf{x})) \\
&= V_i(\eta\mathbf{x}, t_n).
\end{aligned}$$

Likewise, using (2.3.1), and the homogeneity of $\overline{V}_i(\mathbf{x}, t_n^+)$ and u_i^C,

$$\begin{aligned}
\eta^p \overline{V}_i(\mathbf{x}, t_n) &= \eta^p\left[\overline{V}_i\left(\mathbf{f}(\mathbf{x}, t_n, \gamma_i^n(\mathbf{x})), t_n^+\right) + \mathcal{R}(t_n)\, u_i^C\left(f(\mathbf{x}, t_n, \gamma_i^n(\mathbf{x}))\right)\right] \\
&= \overline{V}_i\left(\mathbf{f}(\eta\mathbf{x}, t_n, \gamma_i^n(\eta\mathbf{x})), t_n^+\right) + \mathcal{R}(t_n)\, u_i^C\left(f(\eta\mathbf{x}, t_n, \gamma_i^n(\eta\mathbf{x}))\right) \\
&= \overline{V}_i(\eta\mathbf{x}, t_n).
\end{aligned}$$

$\square$

Remark 3.1.10. *Note that the above results on the homogeneity of the utility apply to any utility functions u_i^B and u_i^C satisfying the theorems' antecedents (provided they are bounded and continuous), not just those that are a part of the HARA family.*

Corollary 3.1.11 (Power law homogeneity)**.** *For all regimes $i \in \mathcal{S}$, take $b_i = 0$ and $p_i = p$ for some constant $p \neq 0$. Suppose that a consumption-optimal (full) strategy is employed by the policyholder. Then, $\mathbf{V}(\mathbf{x}, t)$ and $\overline{\mathbf{V}}(\mathbf{x}, t)$ are homogeneous of order 1 and p, respectively, in $\mathbf{x}$.*

Proof. This follows directly from Theorem 3.1.7 and the fact that $u_{\text{HARA}}(x; a, b, p)$ is homogeneous of order p in x and b. That is, for all $\eta > 0$,

$$u_{\text{HARA}}(\eta x; a, \eta b, p) = \eta^p u_{\text{HARA}}(x; a, b, p).$$

□

This encompasses a large family of actuarially relevant functions, namely the power law (a.k.a. isoelastic) utility functions. Under power law utility, we can once again reduce the dimensionality of the problem. In general, when the conditions of Theorem 3.1.7 hold,

$$\overline{\mathbf{V}}(S, W, D, t) = \frac{1}{\eta^p}\overline{\mathbf{V}}(\eta S, \eta W, \eta D, t).$$

Suppose $W \neq 0$. Choosing $\eta = \frac{W^\star}{W}$ with $W^\star \neq 0$ yields

$$\overline{\mathbf{V}}(S, W, D, t) = \left(\frac{W}{W^\star}\right)^p \overline{\mathbf{V}}\left(\frac{W^\star}{W}S, W^\star, \frac{W^\star}{W}D, t\right),$$

along with (3.1.4). This reveals that we need only solve the problem for two values of the withdrawal benefit: $W^\star$ and 0.

3.2 Control set discretization

At an event time $t_n \in \mathcal{T}$ under regime $i \in \mathcal{S}$, the action $\gamma_i^n(\mathbf{x})$ needs to be determined for each point $\mathbf{x}$ in our spatial discretization. Recall that the control set is $\mathcal{C} = [0, 2]$. We discretize this domain using the points

$$0 = \lambda_1 < \lambda_2 < \ldots < \lambda_P = 2.$$

In general, in order to ensure convergence, we compute a series of approximate solutions and let $\max_i [\lambda_{i+1} - \lambda_i] \to 0$ (see §3.4). We approximate a loss-maximizing partial strategy by taking

$$\gamma_i^n(\mathbf{x}) \in \underset{\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_P\}}{\arg\max} \left[V_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right) + \mathcal{R}(t_n) f(\mathbf{x}, t_n, \lambda)\right].$$

Similarly, we approximate a consumption-optimal partial strategy by taking

$$\gamma_i^n(\mathbf{x}) \in \underset{\lambda \in \{\lambda_1, \lambda_2, \ldots, \lambda_P\}}{\arg\max} \left[\overline{V}_i \left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+ \right) + \mathcal{R}(t_n) u_i^C \left(f(\mathbf{x}, t_n, \lambda) \right) \right].$$

A linear search over $\{\lambda_1, \lambda_2, \ldots, \lambda_P\}$ is used to obtain the maximum. Linear interpolation is used to poll values of $V_i(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+)$ (or $\overline{V}_i(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+)$) as $\mathbf{f}(\mathbf{x}, t, \lambda)$ is not constrained to lie on a grid point. We subsequently show that a solution to the pricing problem under a loss-maximizing strategy is convex and monotone. Linear interpolation preserves these qualities.

3.2.1 Bang-bangness

A partial strategy is said to be *bang-bang* when its range is a countable set. We demonstrate that there exists a loss-maximizing strategy composed of partial strategies γ_i^n satisfying

$$\gamma_i^n(\mathbf{x}) \in \{0, 1, 2\} \quad \forall \mathbf{x} \in \Omega. \tag{3.2.1}$$

This means that for a GLWDB, a policyholder behaving as to maximizing losses for the writer (up to equivalence of loss-maximizing strategies) will only ever perform a nonwithdrawal, a withdrawal at exactly the contract rate, or a full surrender. Since the plane $\{\mathbf{x} \in \Omega \mid x_3 = 0\}$ (corresponding to a death benefit of zero) is a subset of Ω, a GLWDB contract without death benefits (referred to as a GLWB) is bang-bang everywhere. The GLWB is observed to be bang-bang by Forsyth and Vetzal (2012). We generalize our findings in §3.2.2 to provide the reader with sufficient conditions for an arbitrary contract to be bang-bang.

W.r.t. the GLWDB, this allows us to "reduce" the control set $[0, 2]$ to the set $\{0, 1, 2\}$. This reduction is computationally relevant, as it allows us to take $\lambda_0 = 0$, $\lambda_1 = 1$ and $\lambda_2 = 2$ as a control set discretization, resulting in a constant amount of work per linear search. Related bang-bangness results are obtained by Dai et al. (2008), Huang and Kwok (2013) for Guaranteed Minimum Withdrawal Benefit (GMWB) contracts. The GMWB problem formulated in these works assumes continuous withdrawals, resulting in a Hamilton-Jacobi-Bellman equation. Hence, the methods used to establish the bang-bangness therein are inherently different from those found here.

Definition 3.2.1 (Convex function)**.** *A function $s : X \to \mathbb{R}$ defined on a convex set X in a vector space is said to be convex if for all $\theta \in (0, 1)$ and $\mathbf{x}, \mathbf{y} \in X$,*

$$s(\theta \mathbf{x} + (1 - \theta)\mathbf{y}) \leqslant \theta s(\mathbf{x}) + (1 - \theta) s(\mathbf{y}).$$

If $-s$ is convex, s is said to be concave. We say $\mathbf{V}$ is convex if for each $i \in \mathcal{S}$, V_i is convex.

Definition 3.2.2 (Monotone function)**.** *A function $s : X \to Y$ defined on partially ordered sets X and Y is monotone (monotone decreasing) if for all $x, y \in X$, $s(x) \leqslant s(y)$ ($s(x) \geqslant s(y)$) whenever $x \leqslant y$. We say $\mathbf{V}$ is monotone (monotone decreasing) if for each $i \in \mathcal{S}$, V_i is monotone (monotone decreasing).*

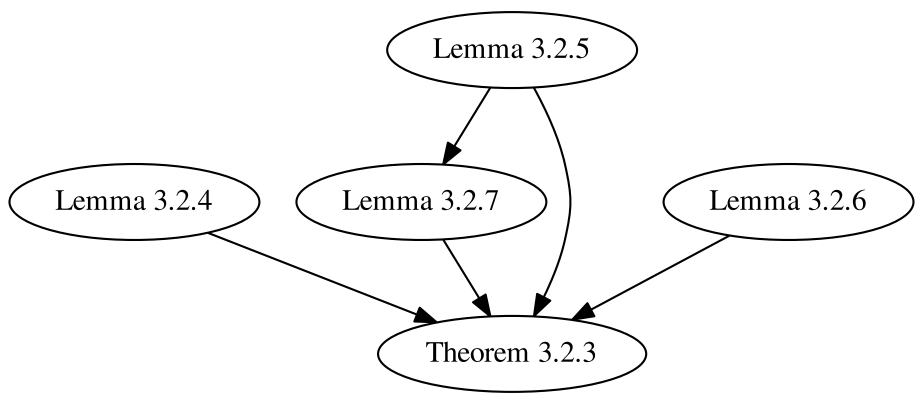


FIGURE 3.2.1: *A guide for the proof of Theorem 3.2.3. An edge from A to B indicates that the result B depends on the result A.*

Theorem 3.2.3 (Bang-bangness)**.** *There exists a loss-maximizing strategy composed of partial strategies γ_i^n satisfying (3.2.1).*

As usual, this result is established via a series of lemmas. We show that if $\mathbf{V}(\mathbf{x}, t_{n+1})$ is convex and monotone in $\mathbf{x}$, so too is $\mathbf{V}(\mathbf{x}, t_n^+)$ (Lemma 3.2.4). Then, we show that if $\mathbf{V}(\mathbf{x}, t_n^+)$ is convex and monotone in $\mathbf{x}$, for each regime i, there exists a loss-maximizing partial strategy γ_i^n satisfying (3.2.1) (Lemma 3.2.5) and that $\mathbf{V}(\mathbf{x}, t_n)$ is convex and monotone in $\mathbf{x}$ (Lemmas 3.2.6 and 3.2.7). Noting that $\mathbf{V}(\mathbf{x}, t_N = T) = \mathbf{0}$ is trivially convex and monotone in $\mathbf{x}$, the desired result follows by induction. Figure 3.2.3 serves as a pictorial guide for this process.

Lemma 3.2.4 (Pricing system convexity and monotonicity between event times)**.** *Suppose that for some n with $1 \leqslant n < N$, $\mathbf{V}(\mathbf{x}, t_{n+1})$ is convex and monotone in $\mathbf{x}$. Then, for all $t \in (t_n, t_{n+1}]$, $\mathbf{V}(\mathbf{x}, t)$ is convex and monotone in $\mathbf{x}$.*

Proof. As before, a change of variables affords us (3.1.2). Since $F \geqslant 0$ almost everywhere (Garroni and Menaldi 1992), and $\mathbf{V}(\mathbf{x}, t_{n+1})$ and $-g^\epsilon(\mathbf{x}, t_{n+1} - \tau')$ are convex and monotone in $\mathbf{x}$, the convexity and monotonicity of $\mathbf{V}(\mathbf{x}, t)$ in $\mathbf{x}$ on $(\mathbb{R}_{\geqslant 0} \setminus \partial\mathbb{R}_{\geqslant 0}) \times \mathbb{R}^2_{\geqslant 0} \times (t_n, t_{n+1}]$ follows. By the presumed continuity of $\mathbf{V}$, we can extend this to $\Omega_{(t_n, t_{n+1}]}$. □

Lemma 3.2.5 (Bang-bangness for a convex and monotone contract)**.** *Suppose that for some regime $i \in \mathcal{S}$ and for some n with $1 \leqslant n < N$, $V_i(\mathbf{x}, t_n^+)$ is convex and monotone in $\mathbf{x}$. Then, there exists a loss-maximizing partial strategy γ_i^n satisfying (3.2.1).*

Proof. W.l.o.g., we need only consider withdrawal times $\mathcal{T}_W$ since ratchets and the event occurring at time zero are not controlled by the policyholder. In light of this, consider the point $\langle \mathbf{x}, t_n, i \rangle \in \Omega \times \mathcal{T}_W \times \mathcal{S}$. Let

$$v_\lambda = V_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right)$$

be the contract value "immediately after" the withdrawal time t_n assuming the policyholder performs action λ.

Suppose that $\{0,1\}\cap\Gamma_i^n(\mathbf{x})=\emptyset$ and $(0,1)\cap\Gamma_i^n(\mathbf{x})\neq\emptyset$ (i.e. no loss-maximizing partial strategy γ_i^n exists with $\gamma_i^n(\mathbf{x})$ indicating nonwithdrawal or withdrawal at exactly the contract rate, but a loss-maximizing partial strategy γ_i^n does exist with $\gamma_i^n(\mathbf{x})$ indicating nonzero withdrawal at strictly below the contract rate). Then, by (1.3.5), (1.3.7) and Definition 1.4.4, there exists $\lambda\in(0,1)$ s.t.

$$v_\lambda+\mathcal{R}(t_n)\,f(\mathbf{x},t_n,\lambda)>v_0+\mathcal{R}(t_n)\,f(\mathbf{x},t_n,0) \tag{3.2.2}$$

and

$$v_\lambda+\mathcal{R}(t_n)\,f(\mathbf{x},t_n,\lambda)>v_1+\mathcal{R}(t_n)\,f(\mathbf{x},t_n,1). \tag{3.2.3}$$

Furthermore, (1.3.6) suggests that

$$\begin{aligned}
&\lambda\mathbf{f}(\mathbf{x},t_n,1)+(1-\lambda)\,\mathbf{f}(\mathbf{x},t_n,0)\\
&=\lambda\left\langle(x_1-G(t_n)\,x_2)\vee 0,x_2,(x_3-G(t_n)\,x_2)\vee 0\right\rangle+(1-\lambda)\left\langle x_1,x_2\left(1+B(t_n)\right),x_3\right\rangle\\
&\geqslant\lambda\left\langle(x_1-G(t_n)\,x_2)\vee 0,x_2,(x_3-G(t_n)\,x_2)\vee 0\right\rangle+(1-\lambda)\left\langle x_1,x_2,x_3\right\rangle\\
&=\left\langle\lambda(x_1-G(t_n)\,x_2)\vee 0+(1-\lambda)\,x_1,x_2,\lambda(x_3-G(t_n)\,x_2)\vee 0+(1-\lambda)\,x_3\right\rangle.
\end{aligned} \tag{3.2.4}$$

Let

$$x_1^+=\lambda(x_1-G(t)\,x_2)\vee 0+(1-\lambda)\,x_1,$$

corresponding to the first element in (3.2.4). If $x_1-G(t_n)\,x_2\geqslant 0$,

$$x_1^+=\lambda(x_1-G(t_n)\,x_2)+(1-\lambda)\,x_1=x_1-\lambda G(t_n)\,x_2\geqslant(x_1-\lambda G(t_n)\,x_2)\vee 0.$$

If, however, $x_1-G(t_n)\,x_2<0$, then $(1-\lambda)\,x_1>x_1-\lambda G(t_n)\,x_2$. This along with $(1-\lambda)\,x_1\geqslant 0$ yields

$$x_1^+=(1-\lambda)\,x_1\geqslant(x_1-\lambda G(t_n)\,x_2)\vee 0. \tag{3.2.5}$$

An identical argument can be carried out to show that

$$x_3^+=(1-\lambda)\,x_3+\lambda(x_3-G(t_n)\,x_2)\vee 0\geqslant(x_3-\lambda G(t_n)\,x_2)\vee 0. \tag{3.2.6}$$

Recalling that

$$\mathbf{f}(\mathbf{x},t_n,\lambda)=\left\langle(x_1-\lambda G(t_n)\,x_2)\vee 0,x_2,(x_3-\lambda G(t_n)\,x_2)\vee 0\right\rangle,$$

we conclude that by (3.2.4), (3.2.5) and (3.2.6),

$$\mathbf{f}(x,t_n,\lambda)\leqslant\lambda\mathbf{f}(\mathbf{x},t_n,1)+(1-\lambda)\,\mathbf{f}(\mathbf{x},t_n,0). \tag{3.2.7}$$

From (1.3.7), it is trivial to deduce

$$f(\mathbf{x},t_n,\lambda)\leqslant\lambda f(\mathbf{x},t_n,1)+(1-\lambda)\,f(\mathbf{x},t_n,0), \tag{3.2.8}$$

as the above holds with equality. (3.2.2) implies that

$$\mathcal{R}(t_n)\,f(\mathbf{x},t_n,\lambda)>v_0-v_\lambda+\mathcal{R}(t_n)\,f(\mathbf{x},t_n,0). \tag{3.2.9}$$

Multiplying (3.2.8) by $\mathcal{R}(t_n) \geqslant 0$ yields

$$\mathcal{R}(t_n)(\lambda f(\mathbf{x}, t_n, 1) + (1-\lambda) f(\mathbf{x}, t_n, 0)) \geqslant \mathcal{R}(t_n) f(\mathbf{x}, t_n, \lambda). \tag{3.2.10}$$

Combining (3.2.9) and (3.2.10) yields

$$\mathcal{R}(t_n)(f(\mathbf{x}, t_n, 1) - f(\mathbf{x}, t_n, 0)) > \frac{v_0 - v_\lambda}{\lambda}. \tag{3.2.11}$$

Similarly, (3.2.3) along with (3.2.10) imply that

$$\frac{v_\lambda - v_1}{(1-\lambda)} > \mathcal{R}(t_n)(f(\mathbf{x}, t_n, 1) - f(\mathbf{x}, t_n, 0)). \tag{3.2.12}$$

Combining (3.2.11) and (3.2.12) yields

$$\frac{v_\lambda - v_1}{(1-\lambda)} > \frac{v_0 - v_\lambda}{\lambda}.$$

Rearranging terms,

$$v_\lambda > \lambda v_1 + (1-\lambda) v_0.$$

By (3.2.7) along with the presumed convexity and monotonicity of $V_i(\mathbf{x}, t_n^+)$,

$$\begin{aligned}
v_\lambda &> \lambda v_1 + (1-\lambda) v_0. \\
&= \lambda V_i\left(\mathbf{f}(\mathbf{x}, t_n, 1), t_n^+\right) + (1-\lambda) V_i\left(\mathbf{f}(\mathbf{x}, t_n, 0), t_n^+\right) \\
&\geqslant V_i\left(\lambda \mathbf{f}(\mathbf{x}, t_n, 1) + (1-\lambda)\mathbf{f}(\mathbf{x}, t_n, 0), t_n^+\right) \\
&\geqslant V_i\left(\mathbf{f}(\mathbf{x}, t_n, \lambda), t_n^+\right) = v_\lambda,
\end{aligned}$$

a contradiction.

Now, suppose that $\{1,2\} \cap \Gamma_i^n(\mathbf{x}) = \emptyset$ and $(1,2) \cap \Gamma_i^n(\mathbf{x}) \neq \emptyset$ (i.e. no loss-maximizing partial strategy γ_i^n exists with $\gamma_i^n(\mathbf{x})$ indicating withdrawal at exactly the contract rate or full surrender, but a loss-maximizing partial strategy γ_i^n does exist with $\gamma_i^n(\mathbf{x})$ indicating partial surrender). Then, by (1.3.5), (1.3.7) and Definition 1.4.4, there exists $\lambda \in (1,2)$,

$$v_\lambda + \mathcal{R}(t_n) f(\mathbf{x}, t_n, \lambda) > v_1 + \mathcal{R}(t_n) f(\mathbf{x}, t_n, 1) \tag{3.2.13}$$

and

$$v_\lambda + \mathcal{R}(t_n) f(\mathbf{x}, t_n, \lambda) > v_2 + \mathcal{R}(t_n) f(\mathbf{x}, t_n, 2). \tag{3.2.14}$$

Let $\theta = \lambda - 1 \in (0,1)$. It is trivial to show that

$$\mathbf{f}(\mathbf{x}, t_n, \lambda) \leqslant \theta \mathbf{f}(\mathbf{x}, t_n, 2) + (1-\theta)\mathbf{f}(\mathbf{x}, t_n, 1), \tag{3.2.15}$$

and

$$f(\mathbf{x}, t_n, \lambda) \leqslant \theta f(\mathbf{x}, t_n, 2) + (1-\theta) f(\mathbf{x}, t_n, 1) \tag{3.2.16}$$

as both statements hold true with equality. As before, we use (3.2.13), (3.2.14) and (3.2.16) to arrive at

$$v_\lambda > \theta v_2 + (1-\theta) v_1.$$

Using (3.2.15) along with the presumed convexity and monotonicity of $V_i\left(\mathbf{x}, t_n^+\right)$,

$$\begin{aligned}
v_\lambda &> \theta v_2 + (1-\theta)\, v_1 \\
&= \theta V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, 2\right), t_n^+\right) + (1-\theta)\, V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, 1\right), t_n^+\right) \\
&\geqslant V_i\left(\theta \mathbf{f}\left(\mathbf{x}, t_n, 2\right) + (1-\theta)\, \mathbf{f}\left(\mathbf{x}, t_n, 1\right), t_n^+\right) \\
&\geqslant V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \lambda\right), t_n^+\right) = v_\lambda,
\end{aligned}$$

a contradiction. □

Lemma 3.2.6 (Loss-maximizing partial strategy preserves convexity)**.** *Suppose that for some regime $i \in \mathcal{S}$ and for some n with $1 \leqslant n < N$, $V_i\left(\mathbf{x}, t_n^+\right)$ is convex and monotone in $\mathbf{x}$ and that the policyholder employs a loss-maximizing partial strategy γ_i^n. Then $V_i\left(\mathbf{x}, t_n\right)$ is convex in $\mathbf{x}$.*

Proof. First, note that for fixed t and λ, $\mathbf{f}$ and f can be written as compositions of functions convex in $\mathbf{x}$, and hence we conclude that $\mathbf{f}$ and f are convex in $\mathbf{x}$. Let $\mathbf{z} = \theta \mathbf{x} + (1-\theta)\, \mathbf{y}$. By the convexity of $\mathbf{f}$,

$$\mathbf{f}\left(\mathbf{z}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right) \leqslant \theta \mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right) + (1-\theta)\, \mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right). \tag{3.2.17}$$

By (3.2.17), the convexity of f and the optimality of a loss-maximizing partial strategy (Definition 1.4.4),

$$\begin{aligned}
V_i\left(\mathbf{z}, t_n\right) &= V_i\left(\mathbf{f}\left(\mathbf{z}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{z}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right) \\
&\leqslant V_i\left(\theta \mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right) + (1-\theta)\, \mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right), t_n^+\right) \\
&\qquad\qquad\qquad\qquad + \mathcal{R}\left(t_n\right) f\left(\theta \mathbf{x} + (1-\theta)\, \mathbf{y}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right) \\
&\leqslant \theta \left[V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right)\right] \\
&\qquad + (1-\theta) \left[V_i\left(\mathbf{f}\left(\mathbf{y}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{y}, t_n, \gamma_i^n\left(\mathbf{z}\right)\right)\right] \\
&\leqslant \theta \left[V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{x}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{x}\right)\right)\right] \\
&\qquad + (1-\theta) \left[V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{y}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{y}, t_n, \gamma_i^n\left(\mathbf{y}\right)\right)\right] \\
&= \theta V_i\left(\mathbf{x}, t_n\right) + (1-\theta)\, V_i\left(\mathbf{y}, t_n\right).
\end{aligned}$$

□

Lemma 3.2.7 (Loss-maximizing partial strategy preserves monotonicity)**.** *Suppose that for some regime $i \in \mathcal{S}$ and for some n with $1 \leqslant n < N$, $V_i\left(\mathbf{x}, t_n^+\right)$ is convex and monotone in $\mathbf{x}$ and that the policyholder employs a loss-maximizing partial strategy γ_i^n. Then, $V_i\left(\mathbf{x}, t_n\right)$ is monotone in $\mathbf{x}$.*

Proof. Let $\mathbf{x} \geqslant \mathbf{y}$. By (1.3.5) and the optimality of a loss-maximizing partial strategy (Definition 1.4.4),

$$\begin{aligned}
V_i\left(\mathbf{x}, t_n\right) &= V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{x}\right)\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{x}, t_n, \gamma_i^n\left(\mathbf{x}\right)\right) \\
&\geqslant V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \lambda\right), t_n^+\right) + \mathcal{R}\left(t_n\right) f\left(\mathbf{x}, t_n, \lambda\right) \;\; \forall \lambda \in \mathcal{C}.
\end{aligned} \tag{3.2.18}$$

By Lemma 3.2.5, it is sufficient to consider the cases $\gamma_i^n(\mathbf{y}) = 0$, $\gamma_i^n(\mathbf{y}) = 1$ and $\gamma_i^n(\mathbf{y}) = 2$. For each case, by (3.2.18), we need only find λ s.t.

$$\begin{aligned} V_i(\mathbf{y}, t_n) &= V_i\left(\mathbf{f}\left(\mathbf{y}, t_n, \gamma_i^n(\mathbf{y})\right), t_n^+\right) + \mathcal{R}(t_n) f\left(\mathbf{y}, t_n, \gamma_i^n(\mathbf{y})\right) \\ &\leqslant V_i\left(\mathbf{f}\left(\mathbf{x}, t_n, \lambda\right), t_n^+\right) + \mathcal{R}(t_n) f\left(\mathbf{x}, t_n, \lambda\right). \end{aligned}$$

When $t_n \in \mathcal{T}_R$ (i.e. t_n corresponds to a ratchet) or $t_n = t_1 = 0$, the result is trivial since the policyholder has no control over the contract. Hence, we consider the case of $t_n \in \mathcal{T}_W$ (i.e. t_n is a withdrawal time).

1. Suppose $\gamma_i^n(\mathbf{y}) = 0$. Take $\lambda = 0$ to get $f(\mathbf{x}, t_n, 0) = f(\mathbf{y}, t_n, 0) = 0$ and $\mathbf{f}(\mathbf{x}, t_n, 0) \geqslant \mathbf{f}(\mathbf{y}, t_n, 0)$.
2. Suppose $\gamma_i^n(\mathbf{y}) = 1$. Take $\lambda = y_2/x_2$ to get $f(\mathbf{x}, t_n, \lambda) = f(\mathbf{y}, t_n, \gamma_i^n(\mathbf{y}))$ and $\mathbf{f}(\mathbf{x}, t_n, \lambda) \geqslant \mathbf{f}(\mathbf{y}, t_n, \gamma_i^n(\mathbf{y}))$.
3. Suppose $\gamma_i^n(\mathbf{y}) = 2$. If $y_1 \leqslant G(t_n) y_2$, we have $f(\mathbf{y}, t_n, 1) = f(\mathbf{y}, t_n, 2) = G(t_n) y_2$ and $\mathbf{f}(\mathbf{y}, t_n, 1) \geqslant \mathbf{f}(\mathbf{y}, t_n, 2)$ and hence we can once again take $\lambda = y_2/x_2$ to arrive at a contradiction. Therefore, we can safely assume $y_1 > G(t_n) y_2$ so that

$$f(\mathbf{y}, t_n, 2) = (1 - \kappa(t_n)) G(t_n) y_2 + \kappa(t_n) y_1 \leqslant y_1.$$

   This inequality is not strict to account for the case of $\kappa(t_n) = 1$.

   (a) Suppose $x_1 > G(t_n) x_2$. Take $\lambda = 2$ to get $f(\mathbf{x}, t_n, 2) = (1 - \kappa(t_n)) G(t_n) x_2 + \kappa(t_n) x_1 \geqslant f(\mathbf{y}, t_n, 2)$ and $\mathbf{f}(\mathbf{x}, t_n, 2) = \mathbf{f}(\mathbf{y}, t_n, 2) = \mathbf{0}$.

   (b) Suppose $x_1 \leqslant G(t_n) x_2$. Take $\lambda = 1$ to get $f(\mathbf{x}, t_n, 1) = G(t_n) x_2 \geqslant x_1 \geqslant y_1 \geqslant f(\mathbf{y}, t_n, 2)$ and $\mathbf{f}(\mathbf{x}, t_n, 1) \geqslant \mathbf{f}(\mathbf{y}, t_n, 2)$.

□

The bang-bang property relies heavily on the convexity of the contract. Figure 3.2.2 illustrates the convexity of a single-regime contract under a loss-maximizing partial strategy. Unlike loss-maximizing partial strategies, for arbitrary utility functions, consumption-optimal partial strategies will not, in general, be bang-bang.

We believe that the above result can be generalized to the case of a local volatility model. This is conjectured below.

Conjecture 3.2.8 (Bang-bangness under local volatility)**.** *Relax Assumption 3.0.2 to allow σ_i to depend on S, with σ_i Hölder continuous in S. There exists a loss-maximizing strategy composed of partial strategies γ_i^n satisfying (3.2.1).*

For relevant work regarding the convexity preservation properties of options written on assets following geometric Brownian motion, we refer the reader to a work by Bergman et al. (1996) (see also Janson and Tysk (2004) for a more general result applicable to parabolic operators).

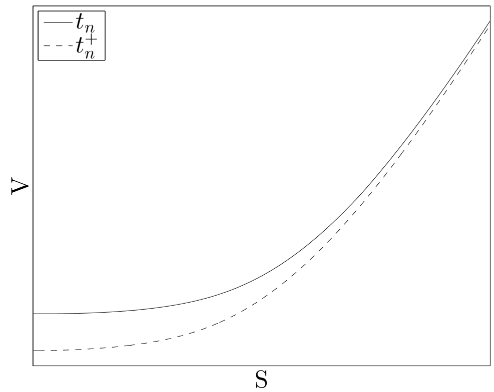


FIGURE 3.2.2: *A typical contract for fixed withdrawal and death benefits across an event time under a loss-maximizing partial strategy.*

3.2.2 Generalization

We now extend the above result to a general variable annuity, a function of one or more possibly correlated assets evolving according to geometric Brownian motion (e.g. the investment account in a GLWDB) and a "state variable" that does not contribute to the price of the variable annuity between events (e.g. the state of the withdrawal and death benefits in a GLWDB). We assume each asset is in $\mathbb{R}_{\geqslant 0}$, while the state variable is in a partially ordered convex set. Assuming a countable number of events between which we can write the solution in a form analogous to (3.1.2), monotonicity and convexity (concavity) are preserved between events. It then remains to check conditions at event times in order to ensure that

(a) Monotonicity is preserved.

(b) Convexity (concavity) is preserved.

(c) The bang-bang property holds.

We outline a set of conditions which ensure that (a-c) hold.

Let $\mathcal{S} = \{1, 2, \ldots, M\}$ and $\Omega = \Omega_1 \times \Omega_2$ where $\Omega_1 = \mathbb{R}^d_{\geqslant 0}$ and Ω_2 is a partially ordered convex set. Denote a point $\mathbf{x} \in \Omega$ by writing $\langle x_1, x_2, \ldots, x_d, x_{d+1} \rangle$ where $x_{d+1} \in \Omega_2$. We consider the problem posed on $\Omega \times (0, T)$ with a single event occurring at time zero. The extension to a countable number of events follows by induction. Let

$$\mathcal{L}_i = \sum_{j,k=1}^{d} a_i^{j,k} x_j x_k \frac{\partial^2}{\partial x_j \partial x_k} + \sum_{j=1}^{d} b_i^j x_j \frac{\partial}{\partial x_j}.$$

Consider $\mathbf{V} \colon \Omega \times (0, T) \to \mathbb{R}^M$ satisfying

$$\mathcal{L}_i V_i + \sum_{m=1}^{M} [c_{i,m} V_m] + \frac{\partial V_i}{\partial t} = g_i \text{ on } \Omega \times (0, T) \ \forall i \in \mathcal{S} \tag{3.2.19}$$

in the classical sense. That is, for each i, the derivatives $\partial V_i / \partial x_j \partial x_k$ for $j, k \in \{1, 2, \ldots, d\}$ exist everywhere and $\partial V_i / \partial t$ exists everywhere except possibly at time zero. Specifically, $\mathbf{V}$ is *càglàd* in time. The coefficients $a_i^{j,k}$, b_i^j and $c_{i,m}$ are bounded, continuous functions of t (independent of $\mathbf{x}$), and $g_i = g_i(\mathbf{x}, t)$ is bounded and continuous in $\mathbf{x}$ and t. For all i and t, we assume the positive semidefiniteness of the matrix with $a_i^{j,k}(t)$ in the j^{th} row, k^{th} column, so that under the log transformation

$$\mathbf{u}(y_1, y_2, \ldots, y_d, y_{d+1}, \tau) = \mathbf{V}\left(e^{y_1}, e^{y_2}, \ldots, e^{y_d}, y_{d+1}, \tau - T\right),$$

(3.2.19) becomes a relation involving a uniformly elliptic operator. As usual, we assume a bound strong enough to ensure uniqueness: for all $x_{d+1} \in \Omega_2$, there exist positive constants L and ℓ s.t.

$$\left|\mathbf{V}\left(e^{x_1}, e^{x_2}, \ldots, e^{x_d}, x_{d+1}, t\right)\right| \leqslant L e^{\ell\left(x_1^2 + x_2^2 + \ldots + x_d^2\right)} \; \forall t \in (0, T].$$

For all $i \in \mathcal{S}$, define $\lambda_i^1, \lambda_i^2, \ldots \in \mathbb{R}$ via the ordering $\lambda_i^1 \leqslant \lambda_i^2 \leqslant \ldots$, and the sets $\mathcal{C}_i$ and $\mathcal{C}_i'$ by

$$\mathcal{C}_i' = \left\{\lambda_i^1, \lambda_i^2, \ldots\right\} \subset \left[\lambda_i^1, \lambda_i^2\right] \cup \left[\lambda_i^2, \lambda_i^3\right] \cup \ldots = \mathcal{C}_i.$$

Lastly, let $\mathbf{f}_i \colon \Omega \times \mathcal{C}_i \to \Omega$ and $f_i \colon \Omega \times \mathcal{C}_i \to \mathbb{R}$.

Let $\mathbb{Z}_{>0} = \{1, 2, \ldots,\}$. We refer to the following list of propositions in the statement of the general theory:

(A1$_i$) g_i is concave in $\mathbf{x}$.

(A2$_i$) g_i is convex in $\mathbf{x}$.

(B1$_i$) g_i is monotone decreasing in $\mathbf{x}$.

(B2$_i$) g_i is monotone increasing in $\mathbf{x}$.

(C1$_i$) $\mathbf{f}_i$ is convex in $\mathbf{x}$.

(C2$_i$) $\mathbf{f}_i$ is concave in $\mathbf{x}$.

(D1$_i$) f_i is convex in $\mathbf{x}$.

(D2$_i$) f_i is concave in $\mathbf{x}$.

(E1$_i$) For all $\mathbf{x} \in \Omega$, $k \in \mathbb{Z}_{>0}$ and $\lambda \in \left(\lambda_i^k, \lambda_i^{k+1}\right)$,

$$\mathbf{f}_i(\mathbf{x}, \lambda) \leqslant \frac{\lambda - \lambda_i^k}{\lambda_i^{k+1} - \lambda_i^k} \mathbf{f}_i\left(\mathbf{x}, \lambda_i^{k+1}\right) + \frac{\lambda_i^{k+1} - \lambda}{\lambda_i^{k+1} - \lambda_i^k} \mathbf{f}_i\left(\mathbf{x}, \lambda_i^k\right)$$

(E2$_i$) For all $\mathbf{x} \in \Omega$, $k \in \mathbb{Z}_{>0}$ and $\lambda \in \left(\lambda_i^k, \lambda_i^{k+1}\right)$,

$$\mathbf{f}_i(\mathbf{x}, \lambda) \geqslant \frac{\lambda - \lambda_i^k}{\lambda_i^{k+1} - \lambda_i^k} \mathbf{f}_i\left(\mathbf{x}, \lambda_i^{k+1}\right) + \frac{\lambda_i^{k+1} - \lambda}{\lambda_i^{k+1} - \lambda_i^k} \mathbf{f}_i\left(\mathbf{x}, \lambda_i^k\right)$$

(F1$_i$) For all $\mathbf{x} \in \Omega$, $k \in \mathbb{Z}_{>0}$ and $\lambda \in \left(\lambda_i^k, \lambda_i^{k+1}\right)$,

$$f_i\left(\mathbf{x}, \lambda\right) \leqslant \frac{\lambda - \lambda_i^k}{\lambda_i^{k+1} - \lambda_i^k} f_i\left(\mathbf{x}, \lambda_i^{k+1}\right). + \frac{\lambda_i^{k+1} - \lambda}{\lambda_i^{k+1} - \lambda_i^k} f_i\left(\mathbf{x}, \lambda_i^k\right)$$

(F2$_i$) For all $\mathbf{x} \in \Omega$, $k \in \mathbb{Z}_{>0}$ and $\lambda \in \left(\lambda_i^k, \lambda_i^{k+1}\right)$,

$$f_i\left(\mathbf{x}, \lambda\right) \geqslant \frac{\lambda - \lambda_i^k}{\lambda_i^{k+1} - \lambda_i^k} f_i\left(\mathbf{x}, \lambda_i^{k+1}\right). + \frac{\lambda_i^{k+1} - \lambda}{\lambda_i^{k+1} - \lambda_i^k} f_i\left(\mathbf{x}, \lambda_i^k\right)$$

(G1$_i$) For all $\mathbf{x} \geqslant \mathbf{y}$ and $\lambda_{\mathbf{y}} \in \mathcal{C}_i'$, there exists $\lambda_{\mathbf{x}} \in \mathcal{C}_i$ s.t. $\mathbf{f}_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \geqslant \mathbf{f}_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$ and $f_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \geqslant f_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$.

(G2$_i$) For all $\mathbf{x} \geqslant \mathbf{y}$ and $\lambda_{\mathbf{y}} \in \mathcal{C}_i'$, there exists $\lambda_{\mathbf{x}} \in \mathcal{C}_i$ s.t. $\mathbf{f}_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \leqslant \mathbf{f}_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$ and $f_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \geqslant f_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$.

(G3$_i$) For all $\mathbf{x} \geqslant \mathbf{y}$ and $\lambda_{\mathbf{y}} \in \mathcal{C}_i'$, there exists $\lambda_{\mathbf{x}} \in \mathcal{C}_i$ s.t. $\mathbf{f}_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \geqslant \mathbf{f}_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$ and $f_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \leqslant f_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$.

(G4$_i$) For all $\mathbf{x} \geqslant \mathbf{y}$ and $\lambda_{\mathbf{y}} \in \mathcal{C}_i'$, there exists $\lambda_{\mathbf{x}} \in \mathcal{C}_i$ s.t. $\mathbf{f}_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \leqslant \mathbf{f}_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$ and $f_i\left(\mathbf{x}, \lambda_{\mathbf{x}}\right) \leqslant f_i\left(\mathbf{y}, \lambda_{\mathbf{y}}\right)$.

(H1$_i$) $V_i\left(\mathbf{x}, T\right)$ is convex in $\mathbf{x}$.

(H2$_i$) $V_i\left(\mathbf{x}, T\right)$ is concave in $\mathbf{x}$.

(I1$_i$) $V_i\left(\mathbf{x}, T\right)$ is monotone in $\mathbf{x}$.

(I2$_i$) $V_i\left(\mathbf{x}, T\right)$ is monotone decreasing in $\mathbf{x}$.

(J1$_i$) For all $\mathbf{x} \in \Omega$, $V_i\left(\mathbf{x}, 0\right) = \sup_{\lambda \in \mathcal{C}_i}\left[V_i\left(\mathbf{f}_i\left(\mathbf{x}, \lambda\right), 0^+\right) + f_i\left(\mathbf{x}, \lambda\right)\right]$.

(J2$_i$) For all $\mathbf{x} \in \Omega$, $V_i\left(\mathbf{x}, 0\right) = \inf_{\lambda \in \mathcal{C}_i}\left[V_i\left(\mathbf{f}_i\left(\mathbf{x}, \lambda\right), 0^+\right) + f_i\left(\mathbf{x}, \lambda\right)\right]$.

(K1$_i$) $V_i\left(\mathbf{x}, 0\right)$ is convex in $\mathbf{x}$.

(K2$_i$) $V_i\left(\mathbf{x}, 0\right)$ is concave in $\mathbf{x}$.

(L1$_i$) $V_i\left(\mathbf{x}, 0\right)$ is monotone in $\mathbf{x}$.

(L2$_i$) $V_i\left(\mathbf{x}, 0\right)$ is monotone decreasing in $\mathbf{x}$.

(M1$_i$) For all $\mathbf{x} \in \Omega$, $V_i\left(\mathbf{x}, 0\right) = \sup_{\lambda \in \mathcal{C}_i'}\left[V_i\left(\mathbf{f}_i\left(\mathbf{x}, \lambda\right), 0^+\right) + f_i\left(\mathbf{x}, \lambda\right)\right]$.

(M2$_i$) For all $\mathbf{x} \in \Omega$, $V_i\left(\mathbf{x}, 0\right) = \inf_{\lambda \in \mathcal{C}_i'}\left[V_i\left(\mathbf{f}_i\left(\mathbf{x}, \lambda\right), 0^+\right) + f_i\left(\mathbf{x}, \lambda\right)\right]$.

Note that (F1$_i$) simply states that for each $\mathbf{x}$ and k, the line segment connecting $\left\langle\lambda_i^k, f\left(\mathbf{x}, \lambda_i^k\right)\right\rangle$ to $\left\langle\lambda_i^{k+1}, f\left(\mathbf{x}, \lambda_i^{k+1}\right)\right\rangle$ is contained in the epigraph of $f\left(\mathbf{x}, \lambda\right)$ as a function of λ. Figure 3.2.3 illustrates this. Clearly, if f is piecewise convex in λ on the intervals of the form $\left(\lambda_i^k, \lambda_i^{k+1}\right)$, then this condition is trivially satisfied. Similar conclusions can be drawn about (E1$_i$), (E2$_i$), and (F2$_i$). Summarizing,

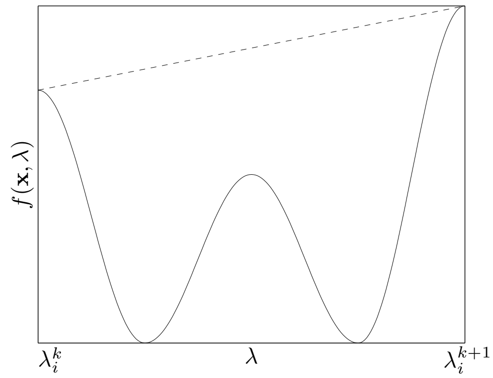


FIGURE 3.2.3: *The line segment connecting* $\left\langle \lambda_i^k, f\left(\mathbf{x}, \lambda_i^k\right)\right\rangle$ *to* $\left\langle \lambda_i^{k+1}, f\left(\mathbf{x}, \lambda_i^{k+1}\right)\right\rangle$ *is contained in the epigraph of the function.*

- $\mathbf{f}$ is convex in λ on $\left(\lambda_i^k, \lambda_i^{k+1}\right)$ for all $k \in \mathbb{Z}_{>0} \Rightarrow$ (E1$_i$)
- $\mathbf{f}$ is concave in λ on $\left(\lambda_i^k, \lambda_i^{k+1}\right)$ for all $k \in \mathbb{Z}_{>0} \Rightarrow$ (E2$_i$)
- f is convex in λ on $\left(\lambda_i^k, \lambda_i^{k+1}\right)$ for all $k \in \mathbb{Z}_{>0} \Rightarrow$ (F1$_i$)
- f is concave in λ on $\left(\lambda_i^k, \lambda_i^{k+1}\right)$ for all $k \in \mathbb{Z}_{>0} \Rightarrow$ (F2$_i$)

Theorem 3.2.9 (Bang-bang variable annuities)**.**

(i) If for all $i \in \mathcal{S}$, (A1$_i$), (B1$_i$), (C1$_i$), (D1$_i$), (E1$_i$), (F1$_i$), (G1$_i$), (H1$_i$), (I1$_i$) and (J1$_i$) are satisfied, then, for all $i \in \mathcal{S}$, (K1$_i$), (L1$_i$) and (M1$_i$) follow.

(ii) If for all $i \in \mathcal{S}$, (A1$_i$), (B2$_i$), (C2$_i$), (D1$_i$), (E2$_i$), (F1$_i$), (G2$_i$), (H1$_i$) (I2$_i$) and (J1$_i$) are satisfied, then, for all $i \in \mathcal{S}$, (K1$_i$), (L2$_i$) and (M1$_i$) follow.

(iii) If for all $i \in \mathcal{S}$, (A2$_i$), (B2$_i$), (C1$_i$), (D2$_i$), (E1$_i$), (F2$_i$), (G3$_i$), (H2$_i$), (I2$_i$) and (J2$_i$) are satisfied, then, for all $i \in \mathcal{S}$, (K2$_i$), (L2$_i$) and (M2$_i$) follow.

(iv) If for all $i \in \mathcal{S}$, (A2$_i$), (B1$_i$), (C2$_i$), (D2$_i$), (E2$_i$), (F2$_i$), (G4$_i$), (H2$_i$), (I1$_i$) and (J2$_i$) are satisfied, then, for all $i \in \mathcal{S}$, (K2$_i$), (L1$_i$) and (M2$_i$) follow.

Remark 3.2.10 (Controls as functions of time)**.** *As mentioned above, a variable annuity with a countable number of events is handled by applying Theorem 3.2.9 inductively. In this extension, the control sets $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_M$ along with the actions $\lambda_1^1, \lambda_1^2, \ldots, \lambda_2^1, \ldots, \lambda_M^1, \ldots$ can be taken to be nonconstant in time so long as they satisfy the antecedents of Theorem 3.2.9 at each event time.*

Remark 3.2.11 (Connection to GLWDB)**.** *The bang-bang property of the GLWDB follows from (i) in Theorem 3.2.9.*

3.3 Localized problem and boundary conditions

We approximate the original problem, posed on

$$\langle S, W, D, t\rangle \in \mathbb{R}^3_{\geqslant 0} \times [0, T],$$

on the truncated domain

$$\langle S, W, D, t\rangle \in [0, S_{\text{Max}}] \times \mathcal{W} \times [0, D_{\text{Max}}] \times [0, T], \tag{3.3.1}$$

where $\mathcal{W} = [0, \infty)$ when a similarity reduction is applied and $\mathcal{W} = [0, W_{\text{Max}}]$ otherwise. We clamp regime-switching jumps that drive the underlying above S_{Max}. Our particular handling of jumps is detailed in §D. No boundary conditions are needed at $S = 0$, $W = 0$, $D = 0$, $W = W_{\text{Max}}$ and $D = D_{\text{Max}}$. It is sufficient to substitute one of the aforementioned boundary values of S, W or D into (1.2.2) and (2.2.2) to retrieve the relevant behaviour. At $S = S_{\text{Max}}$, we impose instead the linearity conditions (Windcliff et al. 2004)

$$\mathbf{V}(S_{\text{Max}}, W, D, t) = \mathbf{C}(t) S_{\text{Max}} \text{ and } \overline{\mathbf{V}}(S_{\text{Max}}, W, D, t) = \overline{\mathbf{C}}(t) S_{\text{Max}} \tag{3.3.2}$$

in an attempt to estimate the true asymptotic behaviour of the contract. The details of this can be found in §C. Errors introduced by the above approximations are small in the region of interest, as verified by numerical experiments. At $t = T$, (1.1.1) and (2.1.3) suggest

$$\mathbf{V}(S, W, D, T) = \overline{\mathbf{V}}(S, W, D, T) = \mathbf{0}.$$

We solve numerically using finite differences. The truncated domain (3.3.1) is discretized using variable-size timestepping (see Johnson (2009) for an expository treatment) and a rectilinear grid in space. If a similarity reduction is used, the withdrawal benefit dimension is discretized by the points 0 and $W^\star = S(0)$.

We use Crank-Nicolson time-stepping with Rannacher smoothing (Rannacher 1984). We discretize the diffusive term using a second-order centered difference, while the convective term is discretized using a centered difference only when the corresponding backward Euler scheme is monotone. Otherwise, an upwind discretization is employed. The details of our discretization are discussed in §D. The resulting system is solved using fixed-point iteration. The details of this approach are available in d'Halluin et al. (2005), Kennedy (2007).

3.4 Determining the hedging cost rider

At contract inception, the withdrawal and death benefits are set to the initial value of the investment account, $S(0)$. That is, $W(0) = S(0)$ and $D(0) = S(0)$. If we overload our previous definition of $\mathbf{V}$ as parameterized by the rider, α_R, the problem becomes one of determining $\alpha_R \in \mathbb{R}$ s.t.

$$V_I(S(0), W(0), D(0), 0; \alpha_R) - \underbrace{\mathcal{R}(0)}_{1} S(0) = 0 \tag{3.4.1}$$

where I is the regime observed at time zero. This is a requirement stating that α_R must be selected so as to compensate the writer for the hedging costs. We term such a value of α_R the *hedging cost rider*. Equation (3.4.1) is solved numerically using Newton's method.

A *refinement level* $E \in \mathbb{N}$ refers to a particular discretization of the domain. We refer to the original grid and control discretization as refinement level 0, and refinement level E is produced by taking the previous level, $E - 1$, and inserting a grid node between each pair of adjacent spatial nodes on the S and D axes. If a bang-bang withdrawal strategy does not exist, the control axis is refined similarly. If a similarity reduction is not applicable (e.g. in the case of exotic utility functions) W is also refined. At refinement level 0, an arbitrary initial guess is used to bootstrap Newton's method. At refinement level $E > 0$, we use the approximation to the hedging cost rider computed at refinement level $E - 1$ as the initial guess.

Chapter 4

Results

We begin by performing experiments under the assumption (i) that the policyholder behaves so as to maximize the writer's losses and (ii) that the policyholder always withdraws at the contract rate. We consider a handful of numerical tests based on perturbations to the base case data in Table 4.2.1. We subsequently move to considering consumption-optimal strategies, in which we use the base case data in Tables 4.2.1 and 4.3.1. Throughout this section, various rates are presented in basis points (bps).

4.1 Validation

We validate the numerical method against a Monte Carlo formulation of the pricing problem. Specifically, we determine the hedging cost rider as in §3.4 under the assumption that the policyholder always withdraws at the contract rate. We refer to this as the *contract rate withdrawal strategy*. We use this particular value of the rider in a Monte Carlo simulation to verify that the writer has indeed hedged their position properly. The contract rate withdrawal strategy is an example of a *static strategy*: given a particular withdrawal time, policyholder behaviour depends only on observable quantities at that time. Static (also referred to as greedy) strategies are trivially simulated by a Monte Carlo simulation (§E).

Validation is performed using a set of contract parameters described in Holz et al. (2007), with the addition of ratcheting and nonratcheting death benefits. These parameters are replicated in Table 4.1.1. To compare with previous work, we assume that death benefit payments are processed annually. A formulation with discrete death benefits is given by Forsyth and Vetzal (2012). Results from computing the hedging cost rider are shown in Table 4.1.2. The Monte Carlo method detailed in §E is used to compute the value of the contract at time zero using computed values of the hedging cost rider. Results from the Monte Carlo simulations are shown in Table 4.1.3. Under the hedging cost rider, the value of the contract should be equal to the amount of wealth initially invested, $S(0) = 100$. Table 4.1.3 substantiates the validity of our method by showing (nonmonotone) convergence to $S(0)$.

Parameter		Value
Volatility	σ	0.15
Risk-free rate	r	0.04
Initial investment	$S(0)$	100
Management rate	α_M	0
Contract rate	G	0.05
Bonus rate	B	0
Initial age	x_0	65
Expiry	T	60
Mortality data		Pasdika et al. (2005)
Mortality payments		Annual
Ratchets		Annual
Withdrawals		Annual
Strategy		Contract rate withdrawal

Time t	Penalty $1 - \kappa(t)$
1	0.05
2	0.04
3	0.03
4	0.02
5	0.01
$\geqslant 6$	0

TABLE 4.1.1: *Single-regime parameters in Holz et al. (2007).*

	Hedging cost rider α_M (bps)	
Refinement	Nonratcheting death benefits	Ratcheting death benefits
0	84.3238355810	160.854056408
1	84.4592646554	161.845696720
2	84.5321112997	162.178675089
3	84.5509146820	162.263487810

TABLE 4.1.2: *Hedging cost rider values acquired using the data in Table 4.1.1 with nonratcheting and ratcheting death benefits (under the contract rate withdrawal strategy).*

	Ratcheting death benefits		Nonratcheting death benefits	
Simulations	Mean	Standard error	Mean	Standard error
10^3	98.9105	$2.39230 \cdot 10^{+0}$	98.8808	$2.59831 \cdot 10^{+0}$
10^4	99.7417	$7.53455 \cdot 10^{-1}$	99.7508	$8.18451 \cdot 10^{-1}$
10^5	100.005	$2.44526 \cdot 10^{-1}$	100.048	$2.65138 \cdot 10^{-1}$
10^6	100.024	$7.73968 \cdot 10^{-2}$	100.034	$8.37591 \cdot 10^{-2}$
10^7	100.001	$2.43845 \cdot 10^{-2}$	99.995	$2.63799 \cdot 10^{-2}$

TABLE 4.1.3: *Monte Carlo validations (under the contract rate withdrawal strategy) for the nonratcheting death benefit with $\alpha_R = 84.5509146820$ bps and ratcheting death benefit with $\alpha_R = 162.263487810$ bps (Table 4.1.2). Standard error is computed at the 99% confidence level.*

Refinement	Hedging cost rider (bps)
0	64.6459730380
1	64.8346526598
2	64.8988278056
3	64.9153982141
4	64.9196165169
5	64.9206819769
6	64.9209497393

TABLE 4.1.4: *Hedging cost rider values acquired using the data in Table 4.1.1 with no death benefits (under the contract rate withdrawal strategy).*

This validation experiment reveals some interesting preliminary trends. Specifically, the data suggests that under the contract rate withdrawal strategy, a ratcheting death benefit is significantly more valuable than a nonratcheting one. Furthermore, Table 4.1.4, showing the values computed for the contract specified by Table 4.1.1 without a death benefit (validated against Forsyth and Vetzal (2012)), suggests that even the nonratcheting death benefit is valuable. Lastly, rates of convergence are calculated in Table 4.1.5, and suggest that the method achieves quadratic convergence.

4.2 Loss-maximizing and contract rate withdrawal

All tests in this section are performed on perturbations to the base case data in Table 4.2.1. Table 4.2.2 documents wide variation in the hedging cost rider across different volatility and interest rate parameters for the two regimes considered, and for the cases with a ratcheting death benefit, with a nonratcheting death benefit, and without a death benefit. Of course, in any otherwise identical scenario, the loss-maximizing withdrawal assumption results in a higher rider since this represents the worst case scenario for the insurer. As we might expect, higher volatility is associated with an increase in the cost of hedging and thus a higher rider. The rider is also quite sensitive to the levels of the risk-free interest rate across the two regimes. The presence of a death benefit results in a notably increased rider, particularly if this feature is ratcheting.

Withdrawal analysis

We now turn to a brief exploration of loss-maximizing withdrawal strategies by the policyholder. Figures 4.2.1 and 4.2.2 show these strategies under each regime

Refinement	Absolute error	Change in error	Ratio of errors
0	$2.1819 \cdot 10^{-2}$		
1	$6.8474 \cdot 10^{-3}$	$V^{(1)} - V^{(0)} = 1.4972 \cdot 10^{-2}$	
2	$1.7553 \cdot 10^{-3}$	$V^{(2)} - V^{(1)} = 5.0921 \cdot 10^{-3}$	$\frac{V^{(1)}-V^{(0)}}{V^{(2)}-V^{(1)}} = 2.94$
3	$4.4048 \cdot 10^{-4}$	$V^{(3)} - V^{(2)} = 1.3148 \cdot 10^{-3}$	$\frac{V^{(2)}-V^{(1)}}{V^{(3)}-V^{(2)}} = 3.87$
4	$1.0578 \cdot 10^{-4}$	$V^{(4)} - V^{(3)} = 3.3470 \cdot 10^{-4}$	$\frac{V^{(3)}-V^{(2)}}{V^{(4)}-V^{(3)}} = 3.93$
5	$2.1245 \cdot 10^{-5}$	$V^{(5)} - V^{(6)} = 8.4535 \cdot 10^{-5}$	$\frac{V^{(4)}-V^{(3)}}{V^{(5)}-V^{(4)}} = 3.96$

TABLE 4.1.5: *Convergence table. The value of the contract at time zero is computed using the data in Table 4.1.1 with no death benefit and* $\alpha_R = 64.9209497393$ *(Table 4.1.4). Let* $V^{(E)}$ *be the numerical solution at the* E^{th} *level of refinement. Since* $V^{(6)} \approx 0$ *(due to our choice of hedging cost rider), we take* $V^{(E)}$ *to be the error at each previous level of refinement.*

(Table 4.2.1) at $t = 1, 2, \ldots, 6$ assuming that the corresponding hedging cost rider is charged for hedging the contract and that $D = 100$. In either regime, if W is much bigger than S, the strategy always involves withdrawing at the contract rate, but the strategy in other regions can be quite complex. We note that in the less volatile regime (Figure 4.2.1), the withdrawal strategy does not involve surrender for $t \leqslant 3$, prior to the vanishing of surrender charges at $t > 3$ (Table 4.2.1). However, in the more volatile regime (Figure 4.2.2), the policyholder is more willing to surrender the contract, despite the large penalties at times $t = 1$ and $t = 2$. Regardless, both regimes experience a sudden change in behaviour when surrender charges vanish. Also note that in this regime, the policyholder's willingness to surrender (for large values of S) vanishes at $t = 3$ in anticipation of the triennial ratchet. Both regimes experience a large region of nonwithdrawal at $t = 3$. The complexity of these loss-maximizing strategies provides some further motivation for our consumption-based approach, since it may seem implausible that individual policyholders would actually implement such strategies.

Management rate

Figure 4.2.3 shows the relationship between the hedging cost rider and the management rate. As is to be expected, the rider grows superlinearly as a function of the management rate, since the management rate acts as a drag on the investment account. This confirms the observation in Forsyth and Vetzal (2012) that use of mutual funds with high management fees as the underlying investment for variable annuities results in higher costs for the insurer compared to a policy written on funds with low management fees (e.g. exchange-traded index funds). We also see that for both the loss-maximizing and contract rate withdrawal strategy, the death benefit adds significant value to the contract, consistent with the results reported in

Parameter				Value
Volatility	σ_1	σ_2	0.0832	0.2141
Risk-free rate	r_1	r_2	0.0521	0.0521
Rate of transition	$q^{\mathbb{Q}}_{1\to 2}$	$q^{\mathbb{Q}}_{2\to 1}$	0.0525	0.1364
Jump magnitude	$J_{1\to 2}$	$J_{2\to 1}$	1	1
Initial regime	I			1
Initial investment	$S\left(0\right)$			100
Management rate	α_M			100 bps
Contract rate	G			0.05
Bonus rate	B			0.05
Initial age	x_0			65
Expiry time	T			60
Mortality data				Pasdika et al. (2005)
Mortality payments				Continuous
Ratchets				Triennial
Withdrawals				Annual

Time t	Penalty $1-\kappa\left(t\right)$
1	0.03
2	0.02
3	0.01
$\geqslant 4$	0

TABLE 4.2.1: *Pricing system base case data with regime-dependent parameters obtained from O'Sullivan and Moloney (2010) by calibration to FTSE 100 options in January 2007.*

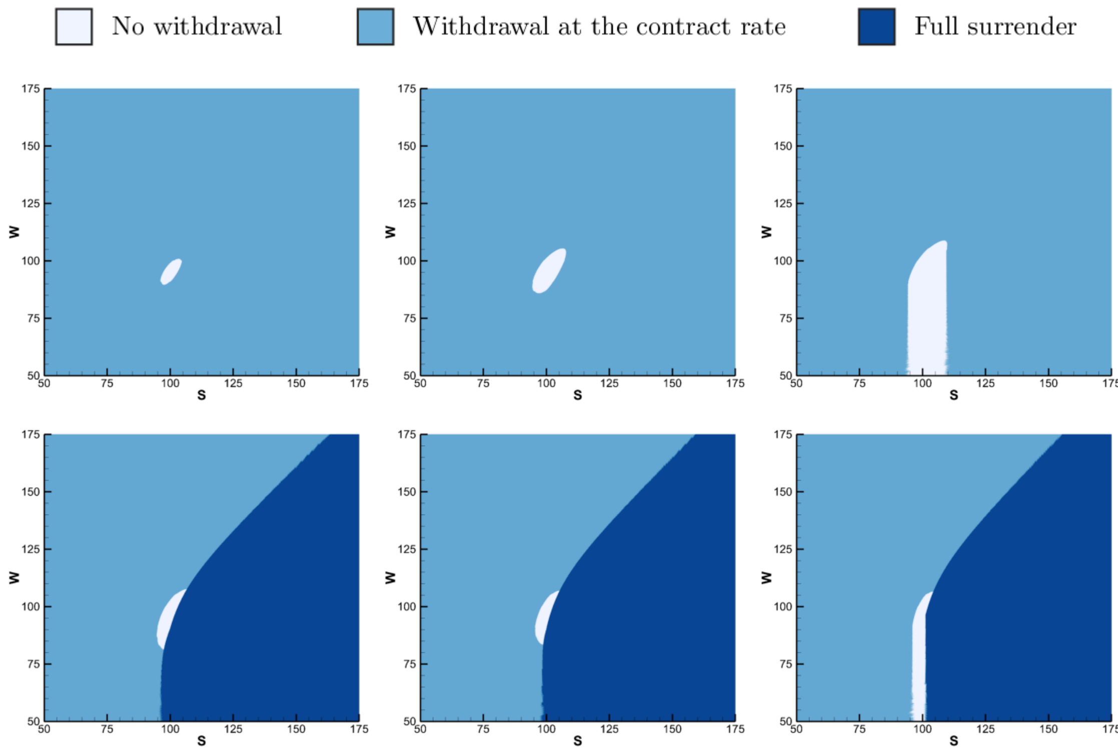


FIGURE 4.2.1: *Observed loss-maximizing strategies at $D = 100$ under regime 1. The hedging cost rider $\alpha_R \approx 37$ bps is used (Table 4.2.2). The subfigures, from top-left to bottom-right, correspond to $t = 1, 2, \ldots, 6$.*

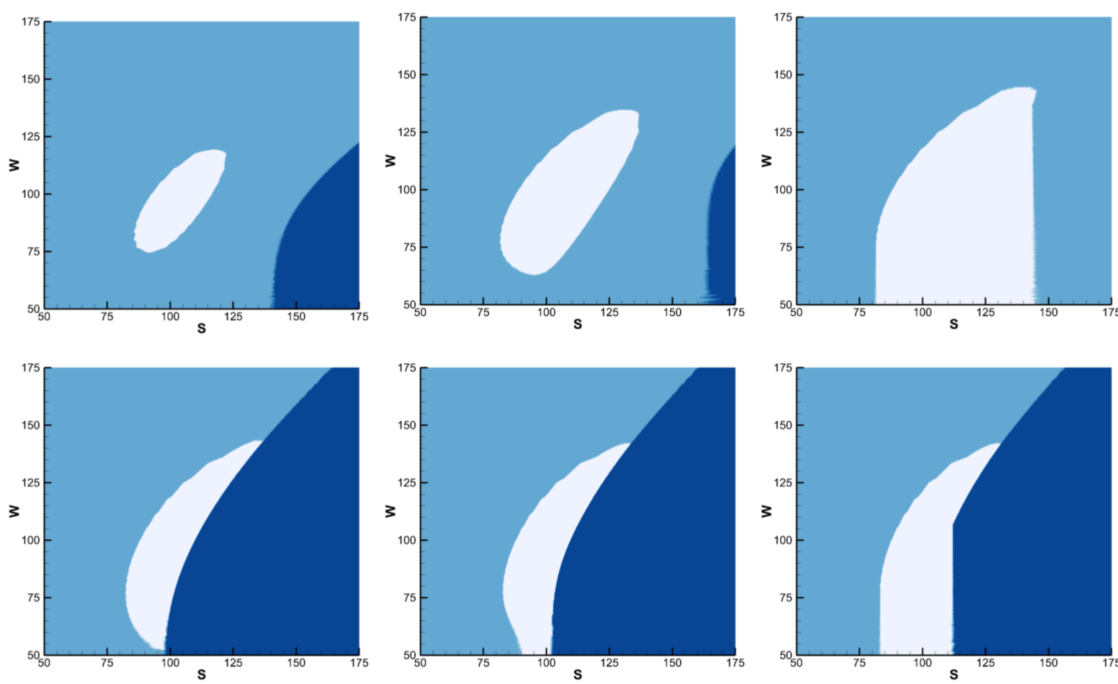


FIGURE 4.2.2: *Observed loss-maximizing strategies at $D = 100$ under regime 2. The hedging cost rider $\alpha_R \approx 139$ bps is used (Table 4.2.2). The subfigures, from top-left to bottom-right, correspond to $t = 1, 2, \ldots, 6$.*

Parameters	Hedging cost rider α_R (bps)					
	Ratcheting Death Benefit		Nonratcheting Death Benefit		No Death Benefit	
Base case (Table 4.2.1)	54	48	37	24	27	19
Initial regime = 2	158	113	139	75	86	52
$(r_1, r_2) = (0.04, 0.06)$	79	72	62	43	44	33
$(r_1, r_2) = (0.03, 0.07)$	124	114	106	76	73	57
$(r_1, r_2) = (0.02, 0.08)$	239	212	224	156	129	104
$(\sigma_1, \sigma_2) = (0.10, 0.20)$	62	56	45	29	31	22
$(\sigma_1, \sigma_2) = (0.15, 0.25)$	133	123	107	69	70	51

TABLE 4.2.2: *The value of the hedging cost rider for perturbations to the data in Table 4.2.1. For each perturbation, riders are calculated under the loss-maximizing (left) and contract rate withdrawal (right) strategies. Values are reported to the nearest basis point.*

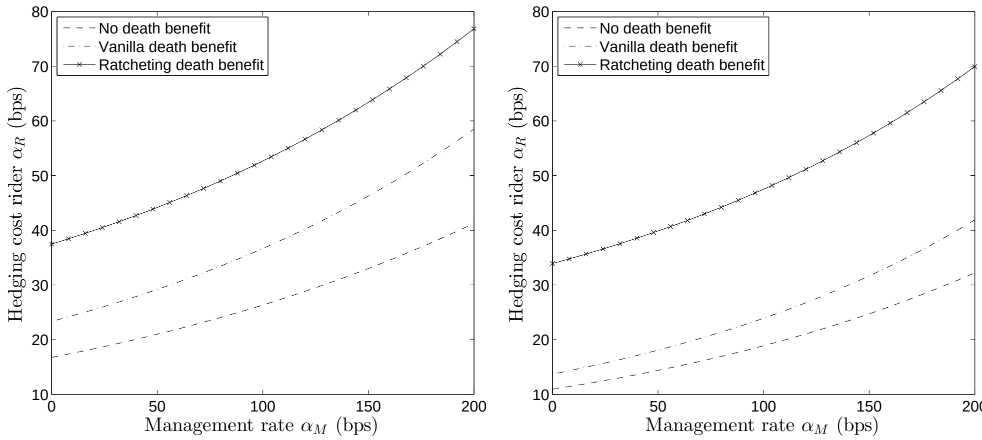
Table 4.2.2. Again, the disparity between the ratcheting and nonratcheting death benefit is even more pronounced.

Alternate fee structure

Some insurers have adopted alternate fee structures that are functions of the auxiliary accounts. In general, the risky account evolves according to

$$dS = (\mu S - \alpha F(S, W, D))\, dt + \sigma S dZ.$$

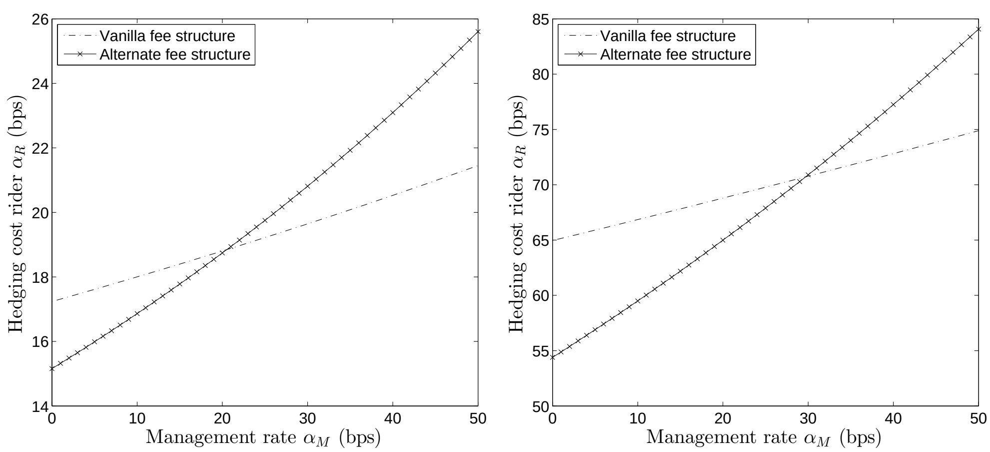
A comparison of the usual fee structure $F = S$ with $F = S \vee W$ on a contract without death benefits for various values of the management rate α_M under the loss-maximizing strategy is shown in Figure 4.2.4. We see that for sufficiently small management rates, the alternate fee structure reduces the hedging cost rider. However, as the management fee increases, the rider calculated under the alternate fee structure surpasses its vanilla counterpart. When the management rate is relatively low, it has a comparatively small impact in terms of decreasing the value of the investment account and so a limited influence on the value of the guarantee. Moreover, since the total rate (i.e. management rate plus rider) applies to the greater of the investment account and the guarantee benefit, the size of the rider in such cases is comparatively small. However, as the management rate increases, the value of the guarantee rises and eventually a higher rider is needed to fund the cost of hedging.



(a) *Loss-maximizing strategy*

(b) *Contract-rate strategy*

FIGURE 4.2.3: *Sensitivity of hedging cost rider to the management rate.*



(a) *Initial regime $I = 1$; no death benefit*

(b) *Initial regime $I = 2$; no death benefit*

FIGURE 4.2.4: *Sensitivity of hedging cost rider to the management rate for different fee structures.*

Parameter			Value	
Drift rate	μ_1	μ_2	0.1	0.1
Time preference	β_1	β_2	0.032	0.032
HARA scaling	a_1	a_2	1	1
HARA offset	b_1	b_2	0	0
Risk-aversion	p_1	p_2	0.5	0.5
Bequest motive	h_1	h_2	1	1
Rate of transition	$q_{1\to 2}$	$q_{2\to 1}$	0.0525	0.1364

TABLE 4.3.1: *Consumption system base case data with rate of time preference obtained from Nishiyama and Smetters (2005).*

4.3 Consumption-optimal withdrawal

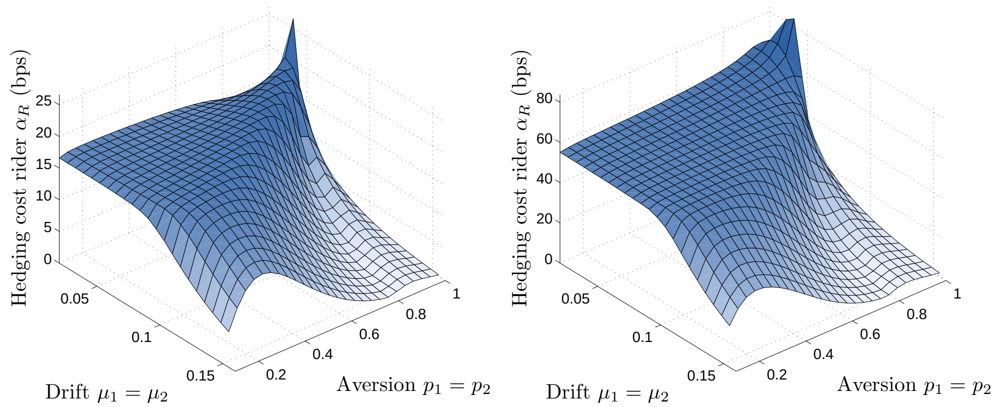
Risk-aversion

Suppose the management rate, α_M, is zero. If for all regimes $i \in \mathcal{S}$ we take the parameterization shown in Table 4.3.2, the consumption-optimal strategy reduces to the loss-maximizing strategy (this can be verified by direct substitution). Reflecting this, we refer to this parameterization as the *degeneracy parameterization*. Since the degeneracy parameterization corresponds to the loss-maximizing strategy, it is guaranteed to yield the highest possible hedging cost rider. We stress that this holds only when the management rate is zero, as reiterated in Table 4.3.2. The utility parameters under this parameterization $u_i^B(x) = h_i u_i^C(x; a_i = 1, b_i = 0, p_i = 1)$ correspond to the case of risk-neutral utility: $u_i^B(x) = u_i^C(x) = x$.

Although the above only holds under the degeneracy parameterization, we expect to see large hedging cost riders under parameterizations that are close to the degeneracy parameterization. Figure 4.3.1 shows the effect of simultaneously varying the regime-dependent drifts μ_1 and μ_2 and risk-aversion parameters p_1 and p_2 on the hedging cost rider for the base case data in Tables 4.2.1 and 4.3.1 for a contract without death benefits. When $\mu_1 = \mu_2 = 0.0521$ and $p_1 = p_2 = 1$, a global maximum appears on each surface. As expected, the parameterization $\mu_1 = \mu_2 = 0.0521$ and $p_1 = p_2 = 1$ is close to the degeneracy parameterization (Tables 4.2.1 and 4.3.1 specify $\alpha = 100$ bps ≈ 0 and $\beta_i = 0.032 \approx 0.04 = r_i$), and hence these maxima (27 bps and 84 bps, rounded to the nearest basis point) are very close to the hedging cost riders for each regime calculated under the loss-maximizing strategy (27 and 86 bps, rounded to the nearest basis point; see Table 4.2.2). Realistically, these maxima are not of great interest to the insurer as they occur where the drift of the investment account is equal to the risk-free rate of return. Both surfaces exhibit a large "plateau" region (i.e. where the gradient is approximately zero) for which the consumption-optimal hedging cost rider is close to that calculated under the

Parameter	α_M	μ_i	β_i	a_i	b_i	p_i	h_i
Value	0	$r_i - \rho_i^{\mathbb{Q}}$	r_i	1	0	1	1

TABLE 4.3.2: *Degeneracy parameterization.*



(a) *Initial regime* $I = 1$ (b) *Initial regime* $I = 2$

FIGURE 4.3.1: *Effects of varying drift and risk-aversion on the hedging cost rider.*

contract rate withdrawal strategy. This suggests that for a large family of parameters, the policyholder withdraws at nearly the contract rate. This can be verified by comparing the hedging cost rider here for the two regimes with those shown in Table 4.2.2.

Taxation

It has been suggested by Moenig and Bauer (2011) that a policyholder's strategy depends on the taxation of their withdrawals. We assume that withdrawals are taxed on the American *last-in first-out* (LIFO) basis and that earnings in the underlying investment account grow on a tax-deferred basis.

This requires the addition of another process $Q(t)$, which is referred to as the *tax base* at time t. The tax base denotes what amount of the underlying investment account is nontaxable. Initially, $Q(0) = S(0)$. Q is piecewise constant between withdrawals. When a withdrawal of size w is made at time t,

$$Q(t) = Q\left(t^-\right) - \underbrace{\left(w - \left[S\left(t^-\right) - Q\left(t^-\right)\right] \vee 0\right) \vee 0}_{\text{Nontaxable portion of the withdrawal}}.$$

The introduction of the tax base variable introduces an additional dimension for which the PDEs must be solved. We assume that the policyholder optimizes their after-tax consumption.

Table 4.3.3 shows the effect of varying the tax rate on the hedging cost rider for the base case data in Tables 4.2.1 and 4.3.1 for a contract without death benefits.

	0%	10%	20%	30%	40%	50%
Initial regime $I = 1$	18.0	18.9	19.2	18.7	17.7	16.3
Initial regime $I = 2$	54.7	55.8	56.3	56.7	57.0	57.2

TABLE 4.3.3: *Sensitivity of the hedging cost rider to the tax rate. Values are reported to the nearest tenth of a basis point.*

Although it is tempting to hypothesize that higher levels of taxation will reduce the hedging cost rider, Table 4.3.3 suggests that this is not the case. For example, consider a contract in which the loss-maximizing strategy involves large regions of surrender. Taxation will generally dissuade the policyholder from surrendering, and hence it is not obvious that the hedging cost rider is monotonically decreasing in the tax rate. In general, we find that for typical levels of risk-aversion, taxation has a small effect on the rider. As the policyholder's risk-aversion approaches risk-neutrality, the effects of taxation increase. Even for extreme tax rates of 50%, the hedging cost rider changes by at most several basis points.

Conclusion

We have introduced a general methodology that allows for the decoupling of policyholder behaviour from the pricing (i.e. determining the cost of hedging) of a variable annuity. Assuming that the underlying investment follows a regime-switching process, this yields two sequences of weakly coupled systems of PDEs: the pricing and utility systems. When considering strategies contingent on the policyholder's level of consumption, the utility systems are used to generate policyholder withdrawal behaviour, which is in turn fed into the pricing systems as a means to determine the no-arbitrage value of the contract. Our methodology is general enough to allow us to consider any withdrawal strategy contingent on either the cost of hedging the contract or the policyholder's level of consumption.

We have adopted the GLWDB as a case study. A similarity reduction transforms our systems of three-dimensional PDEs to systems of two-dimensional PDEs, allowing us to generate numerical solutions with speed. In the absence of a death benefit, these systems further simplify into systems involving one-dimensional PDEs, which (for a reasonable number of regimes) can be solved on the order of hundreds of milliseconds on a modern desktop. Under the loss-maximizing strategy assumption, the control was shown to be bang-bang, allowing us to further optimize the pricing procedure. A generalization of this result is provided via sufficient conditions for a bang-bang control in a variable annuity.

Since GLWDB contracts are held over long periods of time, regime-switching serves as a natural model for the process followed by the underlying asset. This process can incorporate stochastic interest rates and volatility in a simple and intuitive manner. It is also possible to have policyholder preferences which differ between regimes. Results obtained under various regime-switching processes indicate that the hedging cost rider is extremely sensitive to the regime-dependent parameters.

We show that the inclusion of a death-benefit yields large riders for typical contract values under both the loss-maximizing strategy and the static strategy of always withdrawing at the contract rate. We observe an even more pronounced disparity between the no-arbitrage rider generated by a contract with nonratcheting death benefits compared to a contract with ratcheting death benefits. These findings are consistent with the phasing out of products including ratcheting death benefits from the Canadian market.

We find that for a large family of utility functions, the consumption-optimal strategy yields a rider that is very close to the rider calculated by assuming that the policyholder withdraws at the contract rate. This can be understood as substantiating the otherwise seemingly naïve assumption that the policyholder "generally" withdraws at the contract rate. Adopting the contract rate withdrawal strategy renders the pricing problem computationally simple, as this strategy is deterministic and can easily be implemented in either the PDE or an equivalent Monte Carlo formulation. We also find that for typical levels of risk-aversion, the introduction of taxation does not affect the hedging cost rider greatly.

Appendices

Appendix A

Preliminaries

A.1 Principle of no-arbitrage

As is usual, we assume that the market is free of arbitrage opportunities. For each $t \in [0,T]$, let $\langle \Theta, \mathcal{F}(t), \mathbb{P} \rangle$ be a probability space with sample space Θ, σ-algebra $\mathcal{F}(t)$ ($\mathcal{F}$ is a filtration) and "relevant" (i.e. objective or equivalent to) measure $\mathbb{P}$. Suppose Π and $\hat{\Pi}$ are two portfolios (adapted to $\mathcal{F}$) with

$$\mathbb{P}\left[\Pi(t) > \hat{\Pi}(t)\right] > 0$$

and

$$\mathbb{P}\left[\Pi(t) \geqslant \hat{\Pi}(t)\right] = 1.$$

Then it must hold that

$$\Pi(s) > \hat{\Pi}(s) \quad \forall s \leqslant t.$$

This is termed the *principle of no-arbitrage.*

A.2 Itô's lemma

Let X be a process evolving according to

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dZ(t)$$

where Z is a Wiener process. For any sufficiently smooth $f : \mathbb{R}^2 \to \mathbb{R}$,

$$df(t, X(t)) = \frac{\partial f}{\partial t}(t, X(t)) + \frac{\partial f}{\partial x}(t, X(t))\,dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X(t))\,d[X,X](t),$$

where $[X,Y]$ denotes the quadratic variation of X and Y.

Appendix B

Regime-switching model

B.1 Continuous-time Markov chains

A stochastic process X is said to be a *continuous-time Markov chain* if it takes on values from a particular *state space*,

$$\mathcal{S} = \{1, 2, \ldots, M\}$$

and satisfies the *Markov property*,

$$\mathbb{P}\left[X\left(t + \Delta t\right) = x \mid \mathcal{F}\left(t\right)\right] = \mathbb{P}\left[X\left(t + \Delta t\right) = x \mid X\left(t\right)\right]$$

where X is adapted to the filtration $\mathcal{F}$, $x \in \mathcal{S}$ and Δt is nonnegative.

Let $p_{i \to j}\left(t\right)$, where i, $j \in \mathcal{S}$, be the probability that $X\left(t\right) = j$ given $X\left(0\right) = i$. That is,

$$p_{i \to j}\left(t\right) = \mathbb{P}\left[X\left(t\right) = j \mid X\left(0\right) = i\right].$$

$p_{i \to j}$ is termed the *transition probability* of X from i to j. We are particularly interested in the subset of Markov chains that are *time-homogeneous*. X is time-homogeneous if

$$\mathbb{P}\left[X\left(t + \Delta t\right) = j \mid X\left(t\right) = i\right] = \mathbb{P}\left[X\left(s + \Delta t\right) = j \mid X\left(s\right) = i\right]$$

for all t, s, Δt, j, and i. If the Markov chain is time-homogeneous, by choosing $s = 0$, we get that

$$\mathbb{P}\left[X\left(t + \Delta t\right) = j \mid X\left(t\right) = i\right] = \mathbb{P}\left[X\left(\Delta t\right) = j \mid X\left(0\right) = i\right] = p_{i \to j}\left(\Delta t\right).$$

We henceforth assume the time-homogeneity of the process X.

In particular, we consider the ordinary differential equation

$$p_i'\left(t\right) = p_i\left(t\right) Q$$

where

$$p_i\left(t\right) = \left\langle p_{i \to 1}\left(t\right), p_{i \to 2}\left(t\right), \ldots, p_{i \to M}\left(t\right)\right\rangle.$$

We would further impose the condition

$$p_i(0) = \mathbf{e}_i^\dagger$$

where $\mathbf{e}_i$ is the i^{th} Euclidean basis column vector and $\dagger$ denotes transpose to express that a transition cannot occur unless some nonzero time has passed. The above has solutions of the form

$$\begin{aligned} p_i(t) &= \mathbf{e}_i^\dagger \exp(Qt) \\ &= \left[I + \sum_{n=1}^{\infty} \frac{(Qt)^n}{n!}\right]_i \\ &= \left[I + Qt + \mathcal{O}\left(t^2\right)\right]_i . \end{aligned} \tag{B.1.1}$$

In order to familiarize ourselves with the intuition behind (B.1.1), we now consider the analogous discrete process. Let $p_{i\to j}^n$ be the value of the corresponding discrete process at time $n\Delta t$. Define

$$p_i^n = \langle p_{i\to 1}^n, p_{i\to 2}^n, \ldots, p_{i\to M}^n \rangle .$$

Naturally, p_i^n should satisfy the properties of a *probability vector*. That is

$$\sum_{j=1}^{M} p_{i\to j}^n = 1, \quad p_{i\to j}^n \geqslant 0. \tag{B.1.2}$$

Suppose the discrete process evolves according to

$$\begin{aligned} \frac{p_i^{n+1} - p_i^n}{\Delta t} &= p_i^n Q \\ p_i^{n+1} &= p_i^n (I + Q\Delta t) . \end{aligned}$$

The operator mapping p_i^n to p_i^{n+1} should preserve the properties of probability vectors (B.1.2). Summing over the components of p_i^{n+1} yields

$$\begin{aligned} 1 &= \sum_{j=1}^{M} p_{i\to j}^{n+1} \\ &= \sum_{j=1}^{M} \left[p_i^n (I + Q\Delta t)\right]_j \\ &= \sum_{j=1}^{M} \left[p_i^n\right]_j + \Delta t \sum_{j=1}^{M} \left[p_i^n Q\right]_j \\ &= 1 + \Delta t \sum_{j=1}^{M} \left[p_i^n Q\right]_j \\ 0 &= \Delta t \left(p_{i\to 1}^n \sum_{k=1}^{M} q_{1\to k} + p_{i\to 2}^n \sum_{k=1}^{M} q_{2\to k} + \ldots + p_{i\to M}^n \sum_{k=1}^{M} q_{M\to k} \right) \end{aligned}$$

where we have written $[Q]_{i,j}$ as $q_{i\to j}$. This reveals that $\sum_{k=1}^{M} q_{j\to k} = 0$ for all j, as the above relation must hold for all possible p_i^n. Setting

$$q_{j\to j} = -\sum_{\substack{k=1\\k\neq j}}^{M} q_{j\to k}$$

yields

$$\begin{aligned} p_{i\to j}^{n+1} &= p_{i\to j}^{n} + \left(q_{1\to j}\cdot p_{i\to 1}^{n} + q_{2\to j}\cdot p_{i\to 2}^{n} + \ldots + q_{M\to j}\cdot p_{i\to M}^{n}\right)\Delta t \\ &= p_{i\to j}^{n} + \left(\sum_{\substack{k=1\\k\neq j}}^{M} q_{k\to j}p_{i\to k} - q_{j\to k}p_{i\to j}\right)\Delta t. \end{aligned}$$

In this form, it is clear that $q_{k\to j}$ parameterizes the rate of transition from k to j.

We can exploit the fact that the Markov chains we are interested in are time-homogeneous. By (B.1.1),

$$p_{i\to j}(\Delta t) = \delta_{i,j} + q_{i\to j}\Delta t + \mathcal{O}\left((\Delta t)^2\right) \tag{B.1.3}$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Hence, the differential describing the evolution of the continuous-time Markov chain can be described by

$$dX_{i\to j}(t) = \begin{cases} 1 & \text{with probability } \delta_{i,j} + q_{i\to j}dt \\ 0 & \text{with probability } 1 - (\delta_{i,j} + q_{i\to j}dt) \end{cases}$$

(dropping higher-order terms from (B.1.3)). It is understood that each row of $dX(t)$ has exactly one nonzero entry (i.e. at most one transition can occur between t and $t + dt$).

B.2 Regime-switching PDEs

Consider the M-regime processes S evolving according to

$$dS(t) = a_i(S(t), t)\,dt + b_i(S(t), t)\,dZ(t) + \sum_{j=1}^{M} S(t)(J_{i\to j} - 1)\,dX_{i\to j}(t)$$

in which dS describes the increment of S assuming that the regime at time t is i. We restrict $J_{i\to i} = 1$ for all i so that jumps in the underlying are not experienced unless there is a change in regime.

In the relevant literature, it is often mentioned that the introduction of the regime-switching underlying S yields an incomplete market (Zhou and Yin 2003, Elliott et al. 2005), if the hedging portfolio contains only the underlying asset and the risk-free account. We complete the market by adding M hedging instruments to the portfolio, making it possible to hedge the contract perfectly. Note that the assumption of the availability of M instruments is not one that is farfetched; we need only find M instruments written on the regime-switching underlying S. Often, it is possible to take S itself as one of these instruments (this scenario is detailed in §B.3).

We follow the formulation of a regime-switching framework in Kennedy (2007). Consider a portfolio Π short an option V and with positions in instruments $F^{(1)}$, $F^{(2)}$, ..., $F^{(M)}$. We assume that the trading instruments depend only on $S(t)$ and t. Let B represent the money market process with risk-free rate r. Denote by V_i and $F_i^{(k)}$ the values of the option and k^{th} instrument in regime i. Assuming that regime i is observed at time t,

$$\Pi\left(S\left(t\right),t\right) = -V_i\left(S\left(t\right),t\right) + \sum_{k=1}^{M}\left[\omega^{(k)}F_i^{(k)}\left(S\left(t\right),t\right)\right] + B\left(t\right). \tag{B.2.1}$$

The increment of the above portfolio can be written as

$$d\Pi\left(S\left(t\right),t\right) = -dV_i\left(S\left(t\right),t\right) + \sum_{k=1}^{M}\left[\omega^{(k)}dF_i^{(k)}\left(S\left(t\right),t\right)\right] + dB\left(t\right). \tag{B.2.2}$$

where

$$\begin{aligned}
dV_i &= \hat{\mu}_i dt + \hat{\sigma}_i dZ + \sum_{j=1}^{M}\Delta V_{i\to j}dX_{i\to j}\\
\hat{\mu}_i &= \frac{1}{2}b_i^2\frac{\partial^2 V_i}{\partial S^2} + a_i\frac{\partial V_i}{\partial S} + \frac{\partial V_i}{\partial t}\\
\hat{\sigma}_i &= b_i\frac{\partial V_i}{\partial S}\\
\Delta V_{i\to j} &= V_j\left(J_{i\to j}S,t\right) - V_i\left(S,t\right)
\end{aligned}$$

and

$$\begin{aligned}
dF_i^{(k)} &= \bar{\mu}_i^{(k)}dt + \bar{\sigma}_i^{(k)}dZ + \sum_{j=1}^{M}\Delta F_{i\to j}^{(k)}dX_{i\to j}\\
\bar{\mu}_i^{(k)} &= \frac{1}{2}b_i^2\frac{\partial^2 F_i^{(k)}}{\partial S^2} + a_i\frac{\partial F_i^{(k)}}{\partial S} + \frac{\partial F_i^{(k)}}{\partial t}\\
\bar{\sigma}_i^{(k)} &= b_i\frac{\partial F_i^{(k)}}{\partial S}\\
\Delta F_{i\to j}^{(k)} &= F_j^{(k)}\left(J_{i\to j}S,t\right) - F_i^{(k)}\left(S,t\right).
\end{aligned}$$

Substituting these expressions into (B.2.2) yields

$$d\Pi(t) = \left[\sum_{k=1}^{M}\left[\omega^{(k)}\bar{\mu}_i^{(k)}\right] + rB - \hat{\mu}_i\right]dt + \left[\sum_{k=1}^{M}\left[\omega^{(k)}\bar{\sigma}_i^{(k)}\right] - \hat{\sigma}_i\right]dZ$$
$$+ \sum_{j=1}^{M}\left[\sum_{k=1}^{M}\left[\omega^{(k)}\Delta F_{i\to j}^{(k)}\right] - \Delta V_{i\to j}\right]dX_{i\to j}. \tag{B.2.3}$$

To make the portfolio deterministic, we eliminate Brownian risk by

$$\sum_{k=1}^{M}\omega^{(k)}\bar{\sigma}_i^{(k)} = \hat{\sigma}_i \tag{B.2.4}$$

and jump risk by

$$\sum_{k=1}^{M}\omega^{(k)}\Delta F_{i\to j}^{(k)} = \Delta V_{i\to j} \quad \forall j \in \mathcal{S}. \tag{B.2.5}$$

Note that the jump risk equation corresponding to $j = i$ relates a zero change in the hedging instruments to zero change in the option, so that to eliminate jump risk, we need only satisfy $M - 1$ equations.

Given that the portfolio is deterministic, the principle of no-arbitrage requires $r\Pi dt = d\Pi$. Using the expressions (B.2.1) and (B.2.3), we write this as

$$\sum_{k=1}^{M}\omega^{(k)}\left(\bar{\mu}_i^{(k)} - rF_i^{(k)}\right) = \hat{\mu}_i - rV_i. \tag{B.2.6}$$

Equations (B.2.4), (B.2.5) and (B.2.6) make for a total of $M + 1$ equations in M unknowns. This system has a solution if and only if one of the equations is a linear combination of the others. We denote by ξ_i, $q_{i\to 1}^{\mathbb{Q}}$, $q_{i\to 2}^{\mathbb{Q}}$, ..., $q_{i\to M}^{\mathbb{Q}}$ the weights under which the linear dependence requirement

$$\xi_i\left(\sum_{k=1}^{M}\left[\omega^{(k)}\bar{\sigma}_i^{(k)}\right] - \hat{\sigma}_i\right) = \sum_{\substack{j=1\\j\neq i}}^{M} q_{i\to j}^{\mathbb{Q}}\left(\sum_{k=1}^{M}\left[\omega^{(k)}\Delta F_{i\to j}^{(k)}\right] - dV_{i\to j}\right)$$
$$+ \sum_{k=1}^{M}\left[\omega^{(k)}\left(\bar{\mu}^{(k)} - rF_i^{(k)}\right)\right] - (\hat{\mu}_i - rV_i)$$

holds true. Rearranging this expression,

$$0 = \sum_{k=1}^{M}\left[\omega^{(k)}\left(\xi_i\bar{\sigma}_i^{(k)} - \sum_{\substack{j=1\\j\neq i}}^{M}\left[q_{i,j}^{\mathbb{Q}}\Delta F_{i\to j}^{(k)}\right] - \left(\bar{\mu}_i - rF_i^{(k)}\right)\right)\right]$$
$$- \xi_i\hat{\sigma}_i + \sum_{\substack{j=1\\j\neq i}}^{M}\left[q_{i\to j}^{\mathbb{Q}}\Delta V_{i\to j}\right] + \hat{\mu}_i - rV_i.$$

Since this must hold for any position $\omega^{(1)}$, $\omega^{(2)}$, ..., $\omega^{(M)}$, we write the above as

$$\xi_i \bar{\sigma}_i^{(k)} - \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} \Delta F_{i \to j}^{(k)} = \left(\bar{\mu}_i^{(k)} - r F_i^{(k)} \right) \quad \forall k \in \mathcal{S} \tag{B.2.7}$$

and

$$\xi_i \hat{\sigma}_i - \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} \Delta V_{i \to j} = \hat{\mu}_i - r V_i. \tag{B.2.8}$$

This procedure effectively decouples the hedging instruments from the option V. Resolving the symbols $\hat{\mu}_i$ and $\hat{\sigma}_i$ in (B.2.8) yields

$$\frac{1}{2} b_i^2 \frac{\partial^2 V_i}{\partial S^2} + (a_i - \xi_i b_i) \frac{\partial V_i}{\partial S} - r V_i + \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[q_{i \to j}^{\mathbb{Q}} \Delta V_{i \to j} \right] + \frac{\partial V_i}{\partial t} = 0, \tag{B.2.9}$$

which describes a system of M PDEs: one for each regime. The more familiar form above reveals $a_i - \xi_i b_i$ as the risk-neutral drift and the $q_{i \to j}^{\mathbb{Q}}$ terms as the risk-neutral transition intensities.

We express this more compactly by defining

$$q_{i \to i}^{\mathbb{Q}} = - \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}}$$

and noting that

$$\sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} \Delta V_{i \to j} = \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} V_j \left(J_{i \to j} S, t \right) - V_i \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} = \sum_{\substack{j=1 \\ j \neq i}}^{M} q_{i \to j}^{\mathbb{Q}} V_j \left(J_{i \to j} S, t \right) + q_{i \to i}^{\mathbb{Q}} V_i$$

so that (B.2.9) becomes

$$\frac{1}{2} b_i^2 \frac{\partial^2 V_i}{\partial S^2} + (a_i - \xi_i b_i) \frac{\partial V_i}{\partial S} - \left(r - q_{i \to i}^{\mathbb{Q}} \right) V_i + \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[q_{i \to j}^{\mathbb{Q}} V_j \left(J_{i \to j} S, t \right) \right] + \frac{\partial V_i}{\partial t} = 0. \tag{B.2.10}$$

B.3 Eliminating the market price of risk

It is often possible to eliminate the market price of risk $\xi_i b_i$ from (B.2.10) (Kennedy 2007). For example, let

$$a_i \left(S(t), t \right) = (\mu_i - \alpha) S(t)$$

and

$$b_i\left(S\left(t\right),t\right)=\sigma_i S\left(t\right).$$

Under these parameters, (B.2.10) becomes

$$\frac{1}{2}\sigma_i^2 S^2\frac{\partial^2 V_i}{\partial S^2}+\left(\mu_i-\alpha-\xi_i\sigma_i\right)S\frac{\partial V_i}{\partial S}-\left(r-q^{\mathbb{Q}}_{i\to i}\right)V_i+\sum_{\substack{j=1\\ j\neq i}}^{M}\left[q^{\mathbb{Q}}_{i\to j}V_j\left(J_{i\to j}S,t\right)\right]+\frac{\partial V_i}{\partial t}=0. \tag{B.3.1}$$

Suppose further that S itself is not tradeable but tracks the tradeable index $\hat{S}$ with

$$d\hat{S}\left(t\right)=\mu_i\hat{S}\left(t\right)dt+\sigma_i\hat{S}\left(t\right)dZ\left(t\right).$$

Take the 1st instrument, $F^{(1)}$, to be $\hat{S}$ so that

$$\begin{aligned}\bar{\mu}_i^{(1)}&=\mu_i\hat{S}\\ \bar{\sigma}_i^{(1)}&=\sigma_i\hat{S}\\ \Delta F^{(1)}_{i\to j}&=\hat{S}\left(J_{i\to j}-1\right).\end{aligned}$$

Substituting this into (B.2.7) for $k=1$ yields

$$\xi_i\sigma_i\hat{S}-\sum_{\substack{j=1\\ j\neq i}}^{M}q^{\mathbb{Q}}_{i\to j}\hat{S}\left(J_{i\to j}-1\right)=\xi_i\sigma_i\hat{S}-\rho_i^{\mathbb{Q}}\hat{S}=\mu_i\hat{S}-r\hat{S}.$$

More compactly, we write this as

$$\xi_i\sigma_i\hat{S}=\left(\rho_i^{\mathbb{Q}}+\mu_i-r\right)\hat{S} \tag{B.3.2}$$

where

$$\rho_i^{\mathbb{Q}}=\sum_{\substack{j=1\\ j\neq i}}^{M}q^{\mathbb{Q}}_{i\to j}\left(J_{i\to j}-1\right)=\sum_{j=1}^{M}q^{\mathbb{Q}}_{i\to j}J_{i\to j}. \tag{B.3.3}$$

Whenever $\hat{S}$ is equal to 0, S is necessarily 0 so that the term associated with the market price of risk in (B.3.1) also vanishes. We are thus only interested in the case in which $\hat{S}\neq 0$, under which (B.3.2) states that

$$\xi_i\sigma_i=\rho_i^{\mathbb{Q}}+\mu_i-r.$$

Substituting the above into (B.3.1),

$$\frac{1}{2}\sigma_i^2 S^2\frac{\partial^2 V_i}{\partial S^2}+\left(r-\alpha-\rho_i^{\mathbb{Q}}\right)S\frac{\partial V_i}{\partial S}-\left(r-q^{\mathbb{Q}}_{i\to i}\right)V_i+\sum_{\substack{j=1\\ j\neq i}}^{M}\left[q^{\mathbb{Q}}_{i\to j}V_j\left(J_{i\to j}S,t\right)\right]+\frac{\partial V_i}{\partial t}=0.$$

Appendix C

Linearity conditions

C.1 Pricing system

This section is concerned with deriving expressions for the boundary conditions for the pricing system (1.2.2) at the boundary $S = S_{\text{Max}}$. Recall the linearity condition (3.3.2),

$$V_i\left(S_{\text{Max}}, W, D, t\right) = C_i\left(t\right) S_{\text{Max}} \ \forall i \in \mathcal{S}.$$

Substituting this into the pricing system (1.2.2),

$$\begin{aligned} 0 = \Bigg(& C_i'\left(t\right) + \left(r_i - \alpha - \rho_i^{\mathbb{Q}}\right) C_i\left(t\right) - \left(r_i - q_{i\to i}^{\mathbb{Q}}\right) C_i\left(t\right) \\ & + \sum_{\substack{j=1 \\ j\neq i}}^{M} \left[q_{i\to j}^{\mathbb{Q}} C_j\left(t\right) J_{i\to j}\right] + \mathcal{R}\left(t\right)\alpha_M + \mathcal{M}\left(t\right)\left(1 \vee \frac{D}{S_{\text{Max}}}\right)\Bigg) S_{\text{Max}}. \end{aligned}$$

Simplifying and dividing by $S_{\text{Max}} > 0$ yields

$$0 = C_i'\left(t\right) - \left(\alpha + \rho_i^{\mathbb{Q}}\right) C_i\left(t\right) + \sum_{j=1}^{M} \left[q_{i\to j}^{\mathbb{Q}} C_j\left(t\right) J_{i\to j}\right] + \mathcal{R}\left(t\right)\alpha_M + \mathcal{M}\left(t\right)\left(1 \vee \frac{D}{S_{\text{Max}}}\right). \tag{C.1.1}$$

Letting

$$\Psi_i = -\left(\alpha + \rho_i^{\mathbb{Q}}\right)$$

and

$$\Phi_i\left(t\right) = \sum_{j=1}^{M} \left[q_{i\to j}^{\mathbb{Q}} C_j\left(t\right) J_{i\to j}\right] + \mathcal{R}\left(t\right)\alpha_M + \mathcal{M}\left(t\right)\left(1 \vee \frac{D}{S_{\text{Max}}}\right)$$

we compact the expression for $C_i'\left(t\right)$ into

$$C_i'\left(t\right) + \Psi_i C_i\left(t\right) = -\Phi_i\left(t\right).$$

Multiplying by the integrating factor $\exp\left(\Psi_i t\right)$ and integrating from t to $t + \Delta t$,

$$e^{\Psi_i (t+\Delta t)}\left(C_i\left(t+\Delta t\right) - e^{-\Psi_i \Delta t} C_i\left(t\right)\right) = -\int_t^{t+\Delta t} e^{\Psi_i s}\Phi_i\left(s\right) ds.$$

Approximating the integral by

$$\begin{aligned}\int_t^{t+\Delta t} e^{\Psi_i s}\Phi_i(s)\,ds &\approx \Phi_i(t)\int_t^{t+\Delta t} e^{\Psi_i s} ds\\ &= \Phi_i(t)\frac{e^{\Psi_i(t+\Delta t)}\left(1-e^{-\Psi_i\Delta t}\right)}{\Psi_i}\end{aligned}$$

we get

$$C_i(t) \approx e^{\Psi_i\Delta t}C_i(t+\Delta t) + \Phi_i(t)\frac{e^{\Psi_i\Delta t}-1}{\Psi_i}. \tag{C.1.2}$$

C.1.1 Regime-switching behaviour in the limit

As $S \to \infty$, we expect $C_i(t) \leftrightarrow C_j(t)$ for any regimes i and j (the price of the contract will be independent of the current regime and dominated solely by the value of the underlying). In light of this, we assume $C_i(t) = C_j(t)$ for all regimes i and j whenever $S = S_{\text{Max}}$. The above, along with (B.3.3) suggests that

$$\sum_{j=1}^{M}\left[C_j(t)\,q_{i\to j}^{\mathbb{Q}}J_{i\to j}\right] = C_i(t)\sum_{j=1}^{M}\left[q_{i\to j}^{\mathbb{Q}}J_{i\to j}\right] = C_i(t)\,\rho_i^{\mathbb{Q}}.$$

Substituting this expression into (C.1.1) and simplifying,

$$C_i'(t) - \alpha C_i(t) = -\left(\mathcal{R}(t)\,\alpha_M + \mathcal{M}(t)\left(1 \vee \frac{D}{S_{\text{Max}}}\right)\right),$$

which yields the analogous equation to (C.1.2),

$$C_i(t) \approx e^{-\alpha\Delta t}C_i(t+\Delta t) + \left(\mathcal{R}(t)\,\alpha_M + \mathcal{M}(t)\left(1 \vee \frac{D}{S_{\text{Max}}}\right)\right)\frac{1-e^{-\alpha\Delta t}}{\alpha}.$$

We have verified through numerical tests that this approximation produces insignificant error when S_{Max} is sufficiently large.

C.2 Utility system

We now turn our attention to the boundary conditions for the utility system (2.2.2) at the boundary $S = S_{\text{Max}}$. Recall the linearity condition (3.3.2),

$$\overline{V}_i(S_{\text{Max}}, W, D, t) = \overline{C}_i(t)\,S_{\text{Max}}\ \forall i \in \mathcal{S}$$

Letting

$$\varrho_i = \alpha + \beta - \mu - \sum_{j=1}^{M}\left[q_{i\to j}J_{i\to j}\right]$$

and following the arguments above,

$$\overline{C}_i(t)\,S_{\text{Max}} \approx e^{-\varrho_i\Delta t}\overline{C}_i(t+\Delta t)\,S_{\text{Max}} + \mathcal{M}(t)\,u_B(S_{\text{Max}} \vee D)\frac{1-e^{-\varrho_i\Delta t}}{\varrho_i}.$$

Appendix D

Discretization

D.1 Pricing system

This section is concerned with developing a numerical scheme for solving the pricing system (1.2.2) on the truncated domain (3.3.1). We discretize space using a rectilinear grid with

$$0 = S_1 < S_2 < \ldots < S_{\hat{s}} = S_{\text{Max}},$$

$$0 = W_1 < W_2 < \ldots < W_{\hat{\ell}} = W_{\text{Max}},$$

and

$$0 = D_1 < D_2 < \ldots < D_{\hat{d}} = D_{\text{Max}}.$$

Recall that we are solving the system from t_n to t_{n-1} (i.e. moving backwards in time) and hence we use $t_n - \tau^k$ to denote the k^{th} timestep with

$$0 = \tau^1 < \tau^2 < \ldots < \tau^K = t_n - t_{n-1}.$$

Denote by $V_{i,s,w,d}^{n,k}$ the numerical solution of (1.2.2) at a particular point in the discretization. (3.3.1). Specifically,

$$V_{i,s,w,d}^{n,k} \approx V_i\left(S_s, W_w, D_d, t_n - \tau^k\right).$$

Numerical approximations to the first and second spatial derivatives are written in a similar fashion. Specifically,

$$\left[\frac{\partial V}{\partial S}\right]_{i,s,w,d}^{n,k} \approx \frac{\partial V_i}{\partial S}\left(S_s, W_w, D_d, t_n - \tau^k\right)$$

and

$$\left[\frac{\partial^2 V}{\partial S^2}\right]_{i,s,w,d}^{n,k} \approx \frac{\partial^2 V_i}{\partial S^2}\left(S_s, W_w, D_d, t_n - \tau^k\right).$$

We solve the PDE (1.2.2) using Crank-Nicolson. To simplify notation, we assume that the coefficients of the PDE do not depend on space or time. The removal of this assumption does not alter the derivation below. The Crank-Nicolson scheme as applied to (1.2.2) can be written as

$$\frac{V^{n,k+1}_{i,s,w,d} - V^{n,k}_{i,s,w,d}}{\Delta\tau^k} = \frac{1}{2}\left(\mathcal{G}_{i,s,w,d}\left[V^{n,k+1}\right] + \mathcal{G}_{i,s,w,d}\left[V^{n,k}\right]\right)$$

where $\Delta\tau^k = \tau^{k+1} - \tau^k$ and $\mathcal{G}_{i,s,w,d}$ is the discrete spatial operator for regime i, which we develop in this section. For notational succinctness, let

$$\Delta S_s = S_{s+1} - S_{s-1} \quad \text{and} \quad \Delta S_{s+\frac{1}{2}} = S_{s+1} - S_s.$$

the diffusive term in (1.2.2) be approximated using a centred difference, so that

$$\left[\frac{\partial^2 V}{\partial S^2}\right]^{n,k}_{i,s,w,d} = \frac{V^{n,k}_{i,s-1,w,d}}{\Delta S_{s-\frac{1}{2}}\Delta S_s} - \frac{V^{n,k}_{i,s,w,d}}{\Delta S_{s-\frac{1}{2}}\Delta S_{s+\frac{1}{2}}} + \frac{V^{n,k}_{i,s,+1,w,d}}{\Delta S_{s+\frac{1}{2}}\Delta S_s}.$$

Our discretization of the convection term, on the other hand, is contingent on whether the resulting backward Euler scheme is monotone to achieve higher-order convergence while preserving monotonicity (Rannacher 1984). In the notation above, the corresponding backward Euler scheme is

$$\frac{V^{n,k+1}_{i,s,w,d} - V^{n,k}_{i,s,w,d}}{\Delta\tau^k} = \mathcal{G}_{i,s,w,d}\left[V^{n,k+1}\right] \tag{D.1.1}$$

Denote the forward difference of the convection term by

$$\left[\frac{\overrightarrow{\partial V}}{\partial S}\right]^{n,k}_{i,s,w,d} = \frac{V^{n,k}_{i,s+1,w,d} - V^{n,k}_{i,s,w,d}}{\Delta S_{s+\frac{1}{2}}},$$

the backward difference by

$$\left[\frac{\overleftarrow{\partial V}}{\partial S}\right]^{n,k}_{i,s,w,d} = \frac{V^{n,k}_{i,s,w,d} - V^{n,k}_{i,s-1,w,d}}{\Delta S_{s+\frac{1}{2}}},$$

and the central difference by

$$\left[\frac{\overleftrightarrow{\partial V}}{\partial S}\right]^{n,k}_{i,s,w,d} = \frac{V^{n,k}_{i,s+1,w,d} - V^{n,k}_{i,s-1,w,d}}{\Delta S_{s+\frac{1}{2}}}.$$

We approximate the jump term by linear interpolation. Let

$$\theta_{i\to j,s} = \min\left(J_{i\to j} S_s, S_{\hat{s}}\right)$$

where $\hat{s}$ indexes the rightmost grid point. (i.e. $S_1 < S_2 < \ldots < S_{\hat{s}} = S_{\text{Max}}$) Let $u_{i\to j,s}$ be such that

$$S_u \leqslant \theta_{i\to j,s} \leqslant S_{u+1}$$

where we have suppressed u's dependence on i, j and s for brevity. Further define

$$\ell_{i\to j,s}=\frac{S_{u+1}-\theta_{i\to j,s}}{S_{u+1}-S_u}$$

so that a linear approximation to the contract value after a jump occurs is

$$V_j\left(J_{i\to j}S_s,W_w,D_d,t_n-\tau^k\right)\approx\begin{cases}\ell_{i\to j,s}V^{n,k}_{j,u,w,d}+\left(1-\ell_{i\to j,s}\right)V^k_{j,u+1,w,d} & \text{if } J_{i\to j}S_s<S_{\hat{s}}\\ V^k_{j,\hat{s},w,d} & \text{otherwise}\end{cases}.$$

We write $\mathcal{G}$ as the sum of two operators, $\mathcal{J}$ and $\mathcal{K}$ with

$$\mathcal{J}_{i,s,w,d}\left[V^{n,k}\right]=\sigma_i^2S_s^2\left(\frac{V^{n,k}_{i,s-1,w,d}}{\Delta S_{s-\frac{1}{2}}\Delta S_s}-\frac{V^{n,k}_{i,s,w,d}}{\Delta S_{s-\frac{1}{2}}\Delta S_{s+\frac{1}{2}}}+\frac{V^{n,k}_{i,s,+1,w,d}}{\Delta S_{s+\frac{1}{2}}\Delta S_s}\right)$$
$$+\left(r_i-\alpha-\rho_i^{\mathbb{Q}}\right)S_s\left[\frac{\partial V}{\partial S}\right]^{n,k}_{i,s,w,d}$$

and

$$\mathcal{K}_{i,s,w,d}\left[V^{n,k}\right]=-\left(r_i-q^{\mathbb{Q}}_{i\to i}\right)V^{n,k}_{i,s,w,d}$$
$$+\sum_{\substack{j=0\\ j\neq i}}^{m-1}\left[q^{\mathbb{Q}}_{i\to j}\left(\ell_{i\to j,s}V^{n,k}_{j,u,w,d}+\left(1-\ell_{i\to j,s}\right)V^{n,k}_{j,u+1,w,d}\right)\right]$$
$$+\mathcal{R}\left(t_n-\tau^k\right)\alpha_MS_s+\mathcal{M}\left(t_n-\tau^k\right)\left(S_s\vee D_d\right)$$

Assuming $r_i-q^{\mathbb{Q}}_{i\to i}\geqslant 0$, the terms appearing in $\mathcal{K}$ do not prohibit the resulting backward Euler scheme (D.1.1) from being monotone. Therefore, in choosing how to discretize the convective term, it is sufficient to consider $\mathcal{J}$ alone. Rearranging the terms in $\mathcal{G}$ yields

$$\frac{V^{n,k+1}_{i,s,w,d}-V^{n,k}_{i,s,w,d}}{\Delta\tau^k}=\varphi_{i,s}\left(V^{n,k+1}_{i,s-1,w,d}-V^{n,k+1}_{i,s,w,d}\right)$$
$$+\psi_{i,s}\left(V^{n,k+1}_{i,s+1,w,d}-V^{n,k+1}_{i,s,w,d}\right)+\mathcal{K}_{i,s,w,d}\left[V^{n,k+1}\right].$$

In this form, it is clear that the problem of choosing a discretization of the convective term so that the resulting scheme is monotone is equivalent to that of choosing a scheme in which $\varphi_{i,s}$ and $\psi_{i,s}$ are nonnegative. The results of substituting the various types of discretizations (i.e. forward, backward, central) of the diffusion term into (1.2.2) are listed in Table D.1.1.

D.2 Utility system

This section is concerned with developing a numerical scheme for solving the utility system (2.2.2). Denote by $\overline{V}^{n,k}_{i,s,w,d}$ the numerical solution of (2.2.2) at a particular point in the truncated domain (3.3.1) Specifically,

$$\overline{V}^{n,k}_{i,s,w,d}\approx\overline{V}_i\left(S_s,W_w,D_d,t_n-\tau^k\right).$$

	$\varphi_{i,s}$	$\psi_{i,s}$
Forward	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}} + \frac{\left(r_i - \alpha - \rho_i^{\mathbb{Q}}\right) S_s}{\Delta S_{s+\frac{1}{2}}}$
Backward	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s} - \frac{\left(r_i - \alpha - \rho_i^{\mathbb{Q}}\right) S_s}{\Delta S_{s-\frac{1}{2}}}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}}$
Central	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s} - \frac{\left(r_i - \alpha - \rho_i^{\mathbb{Q}}\right) S_s}{\Delta S_s}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}} + \frac{\left(r_i - \alpha - \rho_i^{\mathbb{Q}}\right) S_s}{\Delta S_s}$

TABLE D.1.1: *Values for $\varphi_{i,s}$ and $\psi_{i,s}$ under the various discretizations.*

	$\overline{\varphi}_{i,s}$	$\overline{\psi}_{i,s}$
Forward	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}} + \frac{(\mu_i - \alpha) S_s}{\Delta S_{s+\frac{1}{2}}}$
Backward	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s} - \frac{(\mu_i - \alpha) S_s}{\Delta S_{s-\frac{1}{2}}}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}}$
Central	$\frac{\sigma_i^2 S_i^2}{\Delta S_{s-\frac{1}{2}} \Delta S_s} - \frac{(\mu_i - \alpha) S_s}{\Delta S_s}$	$\frac{\sigma_i^2 S_i^2}{\Delta S_s \Delta S_{s+\frac{1}{2}}} + \frac{(\mu_i - \alpha) S_s}{\Delta S_s}$

TABLE D.2.1: *Values for $\overline{\phi}_{i,s}$ and $\overline{\psi}_{i,s}$ under the various discretizations.*

Following arguments similar to those above, the corresponding Crank-Nicolson scheme for the system (2.2.2) is

$$\frac{\overline{V}^{n,k+1}_{i,s,w,d} - \overline{V}^{n,k}_{i,s,w,d}}{\Delta \tau^k} = \frac{1}{2}\left(\overline{\mathcal{G}}_{i,s,w,d}\left[\overline{V}^{n,k+1}\right] + \overline{\mathcal{G}}_{i,s,w,d}\left[\overline{V}^{n,k}\right]\right)$$

with $\overline{\mathcal{G}}$ being the sum of two operators, $\overline{\mathcal{J}}$ and $\overline{\mathcal{K}}$, in which $\overline{\mathcal{K}}$ contains all regime-switching and time-inhomogeneous terms. The corresponding backward Euler scheme is

$$\frac{\overline{V}^{n,k+1}_{i,s,w,d} - \overline{V}^{n,k}_{i,s,w,d}}{\Delta \tau^k} = \overline{\mathcal{G}}_{i,s,w,d}\left[\overline{V}^{n,k+1}\right].$$

As before, we need only consider $\overline{\mathcal{J}}$ to ensure that the resulting backward Euler scheme is monotone. The results of substituting the various types of discretizations (i.e. forward, backward, central) of the diffusion term into (2.2.2) are listed in Table D.2.1.

Appendix E

Monte Carlo for the contract rate withdrawal strategy

For brevity, we present a single-regime algorithm under the assumptions that events occur yearly and that r, σ and α are constant with $\alpha_M = 0$. Death benefits are assumed to be paid out annually (Forsyth and Vetzal 2012). Set

$$\mathbf{V}^0 = \mathbf{0}, \quad \mathbf{S}^0 = S(0)\,\mathbf{1}, \quad \mathbf{W}^0 = S(0)\,\mathbf{1}, \quad \text{and} \quad \mathbf{D}^0 = S(0)\,\mathbf{1}.$$

where each of the above vectors is in $\mathbb{R}^{m\times 1}$ and m is the number of samples. The amount withdrawn at time $n+1$ is

$$\mathbf{w}^{n+1} = G\mathbf{W}^n.$$

By Itô's lemma (§A.2), we are able to evolve the investment account according to

$$\mathbf{S}^{n+1} = \left(\operatorname{diag}\left(\exp\left(r - \alpha - \frac{1}{2}\sigma^2\right)\mathbf{1} + \sigma\mathbf{z}^{n+1}\right)\mathbf{S}^n - \mathbf{w}^{n+1}\right) \vee \mathbf{0}$$

where $\mathbf{z}^{n+1}$ is a vector of randomly generated samples from a standard normal distribution and $\operatorname{diag}(\mathbf{y})$ forms the matrix A with $A_{i.i} = \mathbf{y}_i$ and $A_{i,j} = 0$ whenever $i \neq j$. $\vee$ is understood to be an element-wise operator in this context. If the withdrawal benefit has a ratchet at time $n+1$, we set

$$\mathbf{W}^{n+1} = \mathbf{S}^{n+1} \vee \mathbf{W}^n$$

and similarly for $\mathbf{D}^{n+1}$ if a ratcheting death benefit is specified. Lastly, set

$$\mathbf{V}^{n+1} = \mathbf{V}^n + e^{-r}\left[\left(\mathcal{R}(n+1) - \mathcal{R}(n)\right)\left(\mathbf{S}^{n+1} \vee \mathbf{D}^n\right) + \mathcal{R}(n+1)\,\mathbf{w}^{n+1}\right]$$

An approximation to the risk-neutral price is given by the sample mean $\left(\mathbf{1}^\dagger\mathbf{V}^T\right)/m$, where $\dagger$ denotes transpose. We use an implementation of the Mersenne Twister (Matsumoto and Nishimura 1998) and the Box-Muller transform (Box and Muller 1958) to generate samples from $\mathcal{N}(0,1)$.

Appendix F

DAV 2004 R

The mortality table in Pasdika et al. (2005) is reproduced here for the reader's convenience in terms of ${}_1q_x$ (the probability that a person aged x will die within the next year).

x	${}_1q_x$	x	${}_1q_x$	x	${}_1q_x$	x	${}_1q_x$	x	${}_1q_x$
65	0.008886	66	0.009938	67	0.011253	68	0.012687	69	0.014231
70	0.015887	71	0.017663	72	0.019598	73	0.021698	74	0.023990
75	0.026610	76	0.029533	77	0.032873	78	0.036696	79	0.041106
80	0.046239	81	0.052094	82	0.058742	83	0.066209	84	0.074583
85	0.083899	86	0.094103	87	0.105171	88	0.116929	89	0.129206
90	0.141850	91	0.154860	92	0.168157	93	0.181737	94	0.195567
95	0.209614	96	0.223854	97	0.238280	98	0.252858	99	0.267526
100	0.278816	101	0.293701	102	0.308850	103	0.324261	104	0.339936
105	0.355873	106	0.372069	107	0.388523	108	0.405229	109	0.422180
110	0.439368	111	0.456782	112	0.474411	113	0.492237	114	0.510241
115	0.528401	116	0.546689	117	0.565074	118	0.583517	119	0.601976
120	0.620400	121	1.000000						

Bibliography

Arrow, K. (1971). The theory of risk aversion. *Essays in the theory of risk-bearing*, 90–120.

Bauer, D., A. Kling, and J. Russ (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. *ASTIN Bulletin-Actuarial Studies in Non Life Insurance 38*(2), 621.

Belanger, A., P. Forsyth, and G. Labahn (2009). Valuing the guaranteed minimum death benefit clause with partial withdrawals. *Applied Mathematical Finance 16*(6), 451–496.

Bergman, Y. Z., B. D. Grundy, and Z. Wiener (1996). General properties of option prices. *The Journal of Finance 51*(5), 1573–1610.

Box, G. and M. Muller (1958). A note on the generation of random normal deviates. *The Annals of Mathematical Statistics 29*(2), 610–611.

Chen, Z., K. Vetzal, and P. Forsyth (2008). The effect of modelling parameters on the value of GMWB guarantees. *Insurance: Mathematics and Economics 43*(1), 165–173.

Cont, R. and P. Tankov (2004). *Financial Modelling With Jump Processes*. Chapman & Hall, London.

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

d'Halluin, Y., P. Forsyth, and K. Vetzal (2005). Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis 25*(1), 87–112.

Duffie, D. and L. Epstein (1992). Stochastic differential utility. *Econometrica* 60, 353–394.

Elliott, R. J., L. Chan, and T. K. Siu (2005). Option pricing and Esscher transform under regime switching. *Annals of Finance 1*(4), 423–432.

Forsyth, P. A. and K. Vetzal (2012). A methodology for the valuation of guaranteed lifelong withdrawal benefits. Working paper, University of Waterloo.

Friedman, A. (1964). *Partial differential equations of parabolic type* (1983 ed.), Volume 196. Prentice-Hall Englewood Cliffs, NJ.

Garroni, M. G. and J. L. Menaldi (1992). *Green functions for second order parabolic integro-differential problems*, Volume 275. Longman Scientific & Technical Harlow.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357–384.

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal 5*(2), 41–53.

Holz, D., A. Kling, and J. Russ (2007). GMWB for life an analysis of lifelong withdrawal guarantees. *Zeitschrift für die gesamte Versicherungswissenschaft*, 1–21.

Huang, Y. T. and Y. K. Kwok (2013). Analysis of optimal dynamic withdrawal policies in withdrawal guarantees products.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance 42*(2), 281–300.

Janson, S. and J. Tysk (2004). Preservation of convexity of solutions to parabolic equations. *Journal of Differential Equations 206*(1), 182–226.

Jin, Z., Y. Wang, and G. Yin (2011). Numerical solutions of quantile hedging for guaranteed minimum death benefits under a regime-switching jump-diffusion formulation. *Journal of computational and applied mathematics 235*(8), 2842–2860.

Johnson, C. (2009). *Numerical solution of partial differential equations by the finite element method*. Dover Publications.

Kennedy, J. (2007). *Hedging contingent claims in markets with jumps.* Ph. D. thesis, Applied Mathematics, University of Waterloo.

Kling, A., F. Ruez, and J. Ruß (2009). The impact of stochastic volatility on pricing, hedging, and hedge efficiency of variable annuity guarantees. Working paper, Ulm University.

Kreps, D. M. and E. L. Porteus (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46, 185–200.

Levi, E. E. (1907). Sulle equazioni lineari totalmente ellittiche alle derivate parziali. *Rendiconti del circolo Matematico di Palermo 24*(1), 275–317.

Lin, X. S., K. S. Tan, and H. Yang (2009). Pricing annuity guarantees under a regime-switching model. *North American Actuarial Journal 13*(3), 316.

Matsumoto, M. and T. Nishimura (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS) 8*(1), 3–30.

Merton, R. (1970). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3, 373–413.

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

Moenig, T. and D. Bauer (2011). Revisiting the risk-neutral approach to optimal policyholder behavior: A study of withdrawal guarantees in variable annuities. Working paper, Georgia State University.

Ngai, A. and M. Sherris (2011). Longevity risk management for life and variable annuities: The effectiveness of static hedging using longevity bonds and derivatives. *Insurance: Mathematics and Economics 49*(1), 100–114.

Nishiyama, S. and K. Smetters (2005). Consumption taxes and economic efficiency with idiosyncratic wage shocks. Technical report, National Bureau of Economic Research.

O'Sullivan, C. and M. Moloney (2010). The variance gamma scaled self-decomposable process in actuarial modelling. *Journal of Business* 75, 305–332.

Pasdika, U., J. Wolff, G. Re, and M. Life (2005). Coping with longevity: The new German annuity valuation table DAV 2004 R. The Living to 100 and Beyond Symposium, Orlando Florida.

Perkins, T. (2011). RBC to stop offering defined benefit plan. The Globe and Mail.

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

Pratt, J. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, 122–136.

Rannacher, R. (1984). Finite element solution of diffusion problems with irregular data. *Numerische Mathematik 43*(2), 309–327.

Shah, P. and D. Bertsimas (2008). An analysis of the guaranteed withdrawal benefits for life option. Working paper, MIT.

Siu, T. K. (2005). Fair valuation of participating policies with surrender options and regime switching. *Insurance: Mathematics and Economics 37*(3), 533–552.

Windcliff, H., P. Forsyth, and K. Vetzal (2001). Valuation of segregated funds: shout options with maturity extensions. *Insurance: Mathematics and Economics 29*(1), 1–21.

Windcliff, H., P. A. Forsyth, and K. Vetzal (2004). Analysis of the stability of the linear boundary condition for the Black-Scholes equation. *Journal of Computational Finance* 8, 65–92.

Yuen, F. L. and H. Yang (2010). Pricing Asian options and equity-indexed annuities with regime switching by the trinomial tree method. *North American Actuarial Journal 14*(2), 256–277.

Zhou, X. Y. and G. Yin (2003). Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization 42*(4), 1466–1482.