Statutory financial reporting for variable annuity guaranteed death benefits: Market practice, mathematical modeling and computation

Runhuan Feng  
Department of Mathematics  
University of Illinois at Urbana-Champaign  
rfeng@illinois.edu

Huaxiong Huang  
Department of Mathematics & Statistics  
York University  
hhuang@yorku.ca

November 30, 2015

Abstract

As more regulatory reporting requirements for equity-linked insurance move towards dependence on stochastic approaches, insurance companies are experiencing increasing difficulty with detailed forecasting and more accurate risk assessment based on Monte Carlo simulations. While there is vast literature on pricing and valuations of various equity-linked insurance products, very few has focused on the challenges of financial reporting for regulatory requirement and internal risk management. Most insurers use either simulation-based spreadsheet calculations or employ third-party vendor software packages. We intend to use a basic variable annuity death benefit as a model example to decipher the common mathematical structure of US statutory financial reporting. We shall demonstrate that alternative deterministic algorithms such as partial differential equation (PDE) methods can also be used in financial reporting, and that a fully quantified model allows us to compare alternatives of risk metrics for financial reporting.

Key Words. Guaranteed minimum death benefit; risk measures; statutory financial reporting; individual model; aggregate model; numerical PDE methods; running supremum.

1 Background

The variable annuity (VA) guarantee product is arguably the most complex investment-combined insurance product available to individual investors. Since the introduction of investment guarantees to the market in late 1990s, product features have become more and more sophisticated and made it increasingly difficult to quantify and assess the investment and longevity risks embedded in these variable annuity riders.

All insurers authorized to do business in the United States are required to prepare statutory financial statements in accordance with accounting principles established by the National Association of Insurance Commissioners (NAIC), known as the Statutory Accounting Principles. The American Academy of Actuaries (AAA) is an advisory body representing practicing actuaries in the U.S., which often makes recommendations of regulatory standards that are later adopted by the NAIC. There are two main components of statutory financial reporting for variable annuities that have seen drastic changes of methodology in the recent few decades.

1. Principles-based reserving (PBR)

Reserve calculation is a standard practice of life insurers for setting aside a certain amount of liquid assets in order to cover claims from in-force insurance policies. Since the early dates
of the life insurance business in the 1800s, insurers have always used a formula-based static approach to calculate reserves. However, with the convergence of insurance and capital markets in past few decades, investment-combined insurance products have grown in complexity, which led to the need for a new method for calculating life insurance policy reserves to account for market risks. This new method, PBR, establishes principles upon which reserves are based, rather than specific formulas. It requires insurers to hold reserves that consider a wide range of future economic conditions that closely reflect true risk profiles of their products. In 2008, the NAIC adopted the Variable Annuity Commissioner’s Annuity Reserve Valuation Method (VA-CARVM), known as the Actuarial Guideline XLIII (AG-43), and principles-based approaches for calculating statutory reserves for all life insurers. CARVM applied to fixed annuities and VA guaranteed minimum death benefits was in effect much earlier under Actuarial Guidelines XXXIII (AG-33) and XXXIV (AG-34). Interested readers are referred to [Sharp, 1999b] and [Sharp, 1999a] for details.

2. Risk-based capital requirement (RBC)

The NAIC’s RBC regime was established in the 1990s as an early warning system for US insurance regulators. The RBC requirement specifies the minimum amount of capital an insurer is required to hold in order to support its overall business operation in consideration of its size and risk profile. If an insurer does not meet the RBC requirement, then regulators have the legal authority to take preventive and corrective measures in order to protect policyholders and the stability of the insurance market. The NAIC developed a system of RBC formulas based on three major areas: (1) Asset Risk; (2) Underwriting Risk; and (3) Other Risk, which are called C-1, C-2, C-3 risks by practitioners. The C-3 category includes interest rate and market risks. Due to the complexity of varying exposure of different product designs and funding strategies, the NAIC implemented the RBC requirements for C-3 risk in two phases. Phase I addressed interest rate risk for single premium life insurance and annuities including deferred and immediate annuities, guaranteed investment certificates, etc. Since 2003, the NAIC has adopted several recommendations and revisions from the AAA for Phase II capital standards for variable annuities and other products with equity-related risks.

There is extensive literature on no-arbitrage pricing of various types of variable annuity guaranteed benefits. To name a few, a model for the guaranteed minimum death benefit (GMDB) was first introduced by [Milevsky and Posner, 2001] and later extended in consideration of rollup and ratchet options in [Ulm, 2008] and [Ulm, 2010]. The valuation of guaranteed minimum withdrawal benefits was considered in [Milevsky and Salisbury, 2006] from a policyholder’s perspective. [Feng and Volkmer, 2015] extended their work to consider the valuation from an insurer’s perspective. A recent work by [Huang et al., 2014] investigated optimal initiation of withdrawals for a guaranteed lifetime withdrawal benefit. The valuation of guaranteed minimum income benefits was introduced in [Marshall et al., 2010]. [Ng and Li, 2011] developed risk-neutral pricing of guaranteed benefits under regime-switching models. [Bernard et al., 2014] used American option techniques to derive the optimal policyholder surrender strategy based on risk-neutral prices. [Bauer et al., 2008], [Bacinello et al., 2011], [Pitacco, 2015] and [Ballotta and Haberman, 2006] provided general frameworks under which various guaranteed benefits can be evaluated. However, it should be pointed out that pricing actuaries rarely use no-arbitrage pricing theory to determine fees and charges in practice. Instead, they run stochastic projections of cash flows to determine pathwise pricing metrics such as internal rate of return on capital, net present value of profits, etc. Nevertheless, no-arbitrage pricing in the literature can provide potential solutions to financial reporting with hedging models, for which the nested simulation is a major technical concern in the industry.
The main focus of the existing academic literature has been to address the fundamental question of how much should policyholders be charged for the guaranteed benefits. Very little attention has been paid to an equally important question, which is how much reserve and capital an insurer should hold to cover expected and unexpected losses. [Hardy, 2003] was among the first work in the literature to address such a question, and systematically exploited risk management of equity-linked insurance. As the North American insurance industry has undertaken rapid development of financial reporting standards in the past decade, there has been a wide gap in the actuarial literature with regard to quantitative models of financial reporting. Recently, [Bauer et al., 2012] proposed various Monte Carlo methods to determine the European solvency capital requirements.

In this paper, we investigate quantitative models for insurer's liabilities from variable annuity guaranteed benefits. However, readers should bear in mind that actual statutory reserve and capital calculations mandate detailed accounting standards, which can be complex and tedious with various product lines. As the purpose of this paper is to explore the quantitative structure of financial reporting, we shall take a minimalist approach and focus only on a few essential elements that involve stochastic components. There are significant differences between the calculation of C-3 Phase II risk-based capital requirement (Total Asset Requirement) and that of the AG-43 reserve (Conditional Tail Expectation amount) from the viewpoint of practitioners\(^1\). Reserving is often done on a seriatim (contract by contract) basis whereas the RBC calculation is performed at the aggregate level. However, from a mathematical point of view, both share very similar quantitative structure and dependency on their underlying stochastic processes, which can be summarized in three steps as shown in the left column in Figure 1. In this paper, we only use the AG-43 reserve as a model example of financial reporting.

\begin{itemize}
  \item \textbf{Step 1:} Use either pre-packaged scenarios or internally built stochastic models (called economic scenario generators) for all risk factors driving the insurer's asset and liability portfolio. These stochastic models have to be calibrated to meet regulatory standards. Generate a variety of sample paths of the stochastic models over a projection period.
  \item \textbf{Step 2:} Use spreadsheets or third-party vendor software to determine account values for an individual or aggregate contract. Under each scenario of account values, follow certain accounting standards to determine the accumulated profit/deficiency for the entire projection length.
  \item \textbf{Step 3:} Repeat Step 2 for each scenario many times to generate an empirical distribution of accumulated surplus/deficiency or other performance metrics. Apply order statistics to estimate certain risk metrics, such as quantile/conditional tail expectation, which form the basis of reserve or capital requirement.
\end{itemize}

The first goal of this paper is to formulate the aforementioned standard practice in a quantitative model. Even though the reserving method (CARVM) has become an industrial standard over the past decade, there is no existing research on the mathematical structure and interpretation of such a practice. As spreadsheet calculations in essence are based on pathwise defined recursive relations, we can integrate this information with the underlying stochastic models to determine a stochastic representation of an insurer' accumulated profit/deficiency.

The second goal is to utilize this stochastic representation to find an alternative deterministic method to compute the corresponding risk measures. This goal can be achieved through a technique, known as the Feymann-Kac formula in physics and applied mathematics literature. The

\(^1\)For example, the calculation required by AG 43 is performed on a pre-tax basis whereas that required by C-3 Phase II is done on an after-tax basis, treatment of standard scenarios, the RBC calculation is based on 90\% CTE whereas the AG-43 is based on 70\% CTE, etc. Detailed discussions of the similarities and differences between the two requirements can be found in [Life Practice Note Steering Committee, 2009].
resulting partial differential equations (PDEs) can be solved by either analytical methods or numerical methods. This approach is outlined in the right column of Figure 1. There are several advantages of the Feymann-Kac approach (the right column of Figure 1) over the Monte-Carlo simulations approach (the left column of Figure 1).

1. Stochastic representations of insurers’ liabilities provide intuitive interpretations of all contributing factors. For example, in (4.2), one can easily identify the accumulated values of benefit outgo and fee income.

2. Simulation methods can be very time consuming, due to the repeated sampling of stochastic scenarios. In particular, various risk metrics used in practice only measure rare events and hence their estimators can have large variances. Analytical or analytical solutions to PDEs are deterministic algorithms, which can be much more efficient.

3. Monte Carlo simulation only allows us to find a solution at a single point of initial economic conditions, whereas numerical PDE methods require marching through solutions at multiple time points and multiple grid points of economic conditions. These intermediate solutions can be used to produce very efficient algorithms for sensitivity testing of risk measures for insurers’ liabilities.

Figure 1: Two computational methods of risk measures

In Section 2, we take the guaranteed minimum death benefit (GMDB) as a model example to show the dynamics of an insurer’s liability on an individual contract basis. In Section 3, we use a spreadsheet calculation to illustrate the main principles of the current market practice of AG-43 reserving, which is based on a mixture of seriatim valuation and policy compression. In order to make a clear comparison with the individual model in Section 2, we formulate the reserving problem as a probability model in Section 4. As a result of the comparable quantitative frameworks in Sections 2 and 4, we demonstrate some drawbacks of the AG-43 practice in Section 5 and propose alternative reserving method based on individual models. For the current AG-43 formulation, we shall introduce the PDE solution method in Section 6 as an alternative computational tool.
Numerical examples will be presented in Section 7 and some extensions with common product considerations will be discussed in Section 8.

2 Mathematical formulation: An individual model

It is most convenient to first introduce a plain GMDB model which quantifies an insurer’s liability on a stand-alone basis. This is to be distinguished in the next section from the commonly used “seriatim basis” in practice, which is a representative individual contract with fully diversified mortality.

A typical variable annuity contract offers policyholders a variety of investment options, each of which has a distinct investment objective. In most cases, a policyholder makes a lump sum purchase payment into an investment account, the value of which will be linked to the equity fund of his/her choosing. To protect policyholders from the down side of investment risk, the GMDB is offered as a rider to the base contract. It typically guarantees policyholders a minimum amount, such as the full refund of purchase payments, payable upon the death of the policyholder, regardless of the balance in his/her investment account.

Let us introduce the notation to be used throughout the paper. For better presentation, we shall consider continuous time versions of the model setup whenever possible.

- $S_t$, the market value of the underlying equity fund at $t$. It is a common practice to approximate equity funds by a combination of well-recognized market indices such as the S&P 500 and the Russell 2000. Each of the indices will be modeled and calibrated to a stochastic process. Most popular assumptions on equity funds include geometric Brownian motion (GBM), stochastic volatility models, etc. As we do not intent to get into details on index tracking and fund mapping, we assume that the dynamics of a single equity fund is driven by

$$S_t = S_0 \exp\left\{ (\mu - \frac{1}{2}\sigma^2)t + \sigma B_t \right\},$$

(2.1)

where $\{B_t, t \geq 0\}$ is a Brownian motion with natural filtration $\{F_t, 0 \leq t \leq T\}$. Note, however, that if each fund is modeled by a GBM, then any combination of funds with relatively stable makeup (modeling rebalanced funds) can also be characterized by a GBM. Bear in mind that the PDE technique in this paper allows for more general models.

- $F_t$, the market value of the policyholder’s investment account\(^2\) at $t \geq 0$. $F_0$ is referred to as the initial purchase payment. For simplicity, we assume that no additional purchase payment or withdrawal is allowed.

- $G$, the guarantee level, also called the benefit base, is specified at policy issue. For example, the guarantee level can be a full refund of initial purchase payment, i.e. $G = F_0$. We shall only consider the constant case for illustration, although the time-varying case can also be handled by the same PDE technique used in this paper.

- $m$, the annualized rate at which fees (including rider charges) are deducted from the investment account. Contract fees are typically collected on a daily basis. We consider the continuous-time version where fees are deducted continuously. The charges allocated to fund the guarantees are also called margin offset and are usually split by benefit. We denote the

\(^2\)This is also referred as a separate account, as it is managed by a third-party fund manager. The name is used in contrast with the general account, which is managed by the insurer itself.
annualized rate of charges allocated to the GMDB by $m_d$. Note that $m > m_d$ to allow for other expenses.

- $T$, the length of projection (or maturity date of a finite term death benefit), typically a rider anniversary.

- $r$, the constant annual risk-free interest rate. The rate reflects the overall yield on assets backing up the liabilities. This is a common assumption in practice, although the techniques in this paper also allows for interest rate models with term structure.

- $\tau_x$, the future lifetime of a policyholder of age $x$ at issue. The mortality is assumed to be independent of the performance of investment accounts. We denote by $\tau p_x$ the probability that a life-age $x$ survives $T$ years, and $\tau q_x$ the probability that a life-age $x$ dies within $T$ years, i.e. $\tau p_x = 1 - \tau q_x = P(\tau_x > T)$. In practice, these conditional probabilities are specified by a life table. In this paper, we shall also encounter the continuous time survivorship model in which $\tau q_x$ is the distribution function of the continuous random variable $\tau_x$ and the force of mortality is defined by

$$
\mu_{x+t} = \frac{d\tau q_x}{dp_x}.
$$

- $L$, the present value of the insurer’s liabilities, the distribution of which will be used to determine reserves.

At the end of each trading day, the account value is marked to market according to the performance of funds in which it invests and deducted by mortality and expenses (M&E) fees. Hence, without the effect of investment guarantees, the account value at time $t$ is given by

$$
F_t = F_0 \frac{S_t}{S_0} e^{-mt} = F_0 e^{\mu^* t + \sigma B_t}, \quad 0 \leq t \leq T,
$$

where $\mu^* = \mu - m - \sigma^2/2$, and the margin offset to fund the GMDB at time $t$ is given by

$$
M_t = m_d F_t, \quad 0 \leq t \leq T.
$$

Had the policyholder died prior to maturity, the GMDB rider pays the designated beneficiary at the time of the policyholder’s death $\tau_x$ the greater of a minimum guaranteed amount $G$ and the account value at maturity $F_{\tau_x}$. The VA writer is liable for the difference, called gross liability, should the former exceeds the latter. In consideration of income generated by the collection of margin offset $\{M_s, 0 \leq s \leq T\}$, we can formulate for each contract the present value of the net liability, which is the gross liability net of rider charges, as follows.

$$
L := e^{-\tau \tau_x} (G - F_{\tau_x})_+ I(\tau_x < T) - \int_0^{T \land \tau_x} e^{-rs} M_s \, ds,
$$

where $(x)_+ = \max\{x, 0\}$ and $T \land \tau_x = \min\{T, \tau_x\}$, meaning the M&E fees are taken until the earlier of the death of the policyholder and the maturity. Note that there are two independent sources of randomness in this model, one being the financial return on the equity fund and the other being the uncertainty on the time of death.

The calculations of risk measures for the individual model are discussed in details in [Feng and Volkmer, 2012], [Feng and Volkmer, 2014].
3 Current market practice: actuarial guideline XLIII

For many years, regulators and the insurance industry have struggled with the issue of applying a uniform reserve standard for variable annuity products, which involve significantly higher market risks than traditional life insurance products due to the complex guaranteed benefit riders. After years of deliberation, the American Academy of Actuaries (AAA) proposed in 2008 AG-43, which essentially established a new industrial standard of reserving based on stochastic modeling and simulations, replacing the old deterministic and formula-based calculations.

Under AG-43, the AAA provides a set of pre-packaged economic scenarios for equity prices, yield rates on government and cooperate bonds, etc. Insurance companies are permitted to run their own internal stochastic models to simulate fund performances as long as the models meet the calibration criteria set by the AAA. Examples of such pre-packaged economic scenarios can be found in [Gorski and Brown, 2005a].

While the industrial practice on reserve calculation varies by contract design, and can be highly complex due to various features, we shall take the minimalistic approach and only investigate the core elements and summarize the basic principles outlined in [Gorski and Brown, 2005b] and [Life Practice Note Steering Committee, 2009]. For the ease of illustration, we deliberately exclude expenses from the spreadsheet calculation, which are arguably critical inputs from a practitioner’s point of view. Nevertheless, expenses are often modeled by deterministic projections, which do not significantly alter the stochastic nature of the model. We shall address these practical issues in Section 8, in which more realistic features can also be incorporated.

Let us first consider the so-called Stochastic Scenario Method described by AG-43. In principle, valuation actuaries should run spreadsheets calculations through each of the simulated scenarios of the fund performance on a representative single contract (also called model cell in practice). The calculations can be summarized in three sections.

Section 1: Projected values before decrements

In this example, we consider the AG-43 reserve valuation one year after the policy is issued. To shorten the presentation, we only present the first two quarters of projection in Tables 1, 2, 3 to illustrate the recursive relation in a typical spreadsheet calculation.

As shown in Table 1, the starting account value is $2,350,000 (cell G1) and benefit base is 1,850,000 (cell H1). Annualized equity returns (Column D) are extracted from a simulated scenario for the whole projection period of 10 years. We only show the first two quarters as an example. For each projection period, the percentage of net returns (Column F) is calculated by the percentage of equity return less the percentage of fees and charges (Column E). Then the net returns are added to the projected account value from the previous period to produce the projected account value for the current period. For instance, a Excel function would be used to define Columns F and G:

\[
F_2 = ((1 + D2)^{1/360} - E2/360)^{360/4} - 1, \quad G_2 = IF(C1 >= \text{maturity}, 0, G1 \ast (1 + F1)).
\]

In this case, the GMDB benefit base is assumed to remain constant. In terms of the notation introduced in the previous section, the calculations (from Columns D to G) can be characterized by (2.2) and projected account values are thus described by \( F_t \) on discrete time points \( t = 4.00, 4.25, 4.50, \ldots \) and the GMDB benefit base is denoted by \( G \).

Section 2: Projected values after decrements

Decrement due to deaths, lapses and annuitization are introduced in this part of the projection. Here, we do not consider dynamic lapses due to policyholder behavior (such as low lapse during periods of high growth and high lapse during periods of low growth), although the model can be
extended to incorporate this feature. In Table 2, the calculations in Columns G, F and I can be described as follows.

Projected account value inforce = Survival rate \times Projected account value = t_p x F_t,
Projected benefit base inforce = Survival rate \times Projected benefit base = t_p x G,
Projected net amount at risk = \max\{0, Projected benefit base inforce - Projected account value inforce\} = t_p x (G - F_t)_+.

### Table 2: Projected values after decrements

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proj. Year</td>
<td>Quarterly Contract Duration</td>
<td>Projected Attained Age</td>
<td>Quarterly Death Rate</td>
<td>Quarterly Lapse Annuity</td>
<td>Projected Survivorship</td>
<td>Projected Account Inforce</td>
<td>Projected Benefit Inforce</td>
<td>Projected Net Amt at Risk</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>4.00</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>100%</td>
<td>2,350,000</td>
<td>1,850,000</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4.25</td>
<td>71</td>
<td>0.527%</td>
<td>1.535%</td>
<td>97.946%</td>
<td>2,243,285</td>
<td>1,850,000</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4.50</td>
<td>71</td>
<td>0.527%</td>
<td>1.535%</td>
<td>95.935%</td>
<td>2,391,223</td>
<td>1,812,005</td>
</tr>
</tbody>
</table>

### Section 3: Projected income statement and accumulated surplus

All incoming and outgoing cash flows are considered in the last section in order to generate the income statement based on this particular scenario. As shown in Table 3, the M&E charges income (Column D) is given by a quarter of the annual M&E rate (m_d = 1.435%) times the mid-quarter fund value (the average of fund values from previous quarter and current quarter). The GMDB benefit (Column E) is calculated by the death rate times mid-quarter GMDB net amount at risk (the average of net amounts at risk from previous quarter and current quarter). The pre-tax income (Column G) is determined by the fee income (Column D) less the GMDB benefit (Column E) plus the investment income (Column F). In other words, if we denote the accumulated surplus by U_t, then the incremental change in the accumulated surplus is determined quarter-to-quarter by the following recursive relation:

\[
\Delta U_t = U_{t+\Delta t} - U_t = m_d \Delta t \frac{F_t + F_{t+\Delta t}}{2} - \Delta t q_{x+t} t_p x (G - F_{t+\Delta t})_+ + (G - F_t)_+ + r \Delta t U_t. \tag{3.1}
\]

In this example, the quarterly interest rate on surplus is r = 1.412%. The investment income on surplus is typically generated from post-tax surplus. However, for simplicity, we do not distinguish the pre-tax and post-tax basis here. For the same reason, the maintenance and overhead expenses, and the surrender charges are all omitted.
### Table 3: Projected income statement and accumulated surplus

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proj. Year</td>
<td>Quarterly Contract Duration</td>
<td>Projected Attained Age</td>
<td>M&amp;E Charges Income</td>
<td>GMDB Benefits</td>
<td>Investment Income on Surplus</td>
<td>Adjust Pre-tax Income</td>
<td>Pre-tax Accumulated Surplus</td>
<td>Present Value Surplus</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>4.00</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4.25</td>
<td>71</td>
<td>8,325</td>
<td>0</td>
<td>-</td>
<td>8,325</td>
<td>8,325</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4.50</td>
<td>71</td>
<td>8,400</td>
<td>0</td>
<td>118</td>
<td>8,518</td>
<td>16,961</td>
</tr>
</tbody>
</table>

In most cases such as the one indicated in the example, the accumulated surplus stays positive as the product is designed to be profitable. However, due to the random nature of investment return, there are extreme cases in which the accumulated surplus becomes negative and the insurer may no longer have sufficient funds to cover its liabilities. The measurement on the likelihood and severity of such as extreme events is in fact what the model is designed to capture for the purpose of risk management. After the income statement is generated for each scenario, the worst case present value of surplus (the smallest, sometimes negative, entry in Column I), also known as the greatest present value of accumulated deficiency (GPVAD) in AG-43, is determined and recorded. In other words, the GPVAD for each scenario is defined by

$$L := -\min_{k=0,1,\cdots,n} e^{-rt_k} U_{t_k},$$

(3.2)

where \( \{t_0 = 0, t_1, \cdots, t_n = T\} \) are the time points of projection, which is on a quarterly basis in the spreadsheet example. Then the so-called Scenario Greatest Present Value (SGPV) is the sum of the GPVAD and the starting asset amount\(^3\), \( F_0 \). This concludes the procedure of a single scenario calculation.

**Conditional tail expectation (CTE) amount:**

The single scenario procedure is repeated for 1,000-2,000 randomly generated scenarios for equity returns either from an insurance company’s own internal model or the AAA pre-packaged cases. The SGPV for all scenarios are gathered and ranked to form an empirical distribution of the insurer’s net liability (surplus/deficiency). In order to identify the worst 100\(p\)% scenarios, practitioners first estimate the quantile risk measure, also known as value at risk in finance literature, defined for any loss random variable \( X \) as

$$\text{VaR}_p(X) := \inf\{V : \mathbb{P}(X \leq V) > p\}.$$  

Observe that

$$\text{VaR}_p(\text{SGPV}) = \text{VaR}_p(L + F_0) = \text{VaR}_p(L) + F_0.$$  

As the starting asset is a fixed amount at the valuation date, we shall only consider the GPVAD \( L \) without loss of generality. After obtaining the quantile risk measure, one can further determine the mean of 100(1−\(p\))% worst scenarios by the conditional tail expectation risk measure

$$\text{CTE}_p(L) := \mathbb{E}[L|L > \text{VaR}_p(L)].$$

These risk measures are usually estimated by 100(1−\(p\))% largest order statistics and their averages in practice. Under AG-43, the **CTE amount** is determined by the 70%-CTE of the SGPV.

\(^3\)Here we consider the policyholder invests all purchase payments in separate accounts.
In principle, the reserve calculation should be done on a contract-by-contract basis. However, given the large volume of contracts in their portfolios, it is considered impractical for insurance companies to carry out the procedures for all individual contracts. Therefore, AG-43 stipulated that in the calculation of CTE amount, contracts can be grouped “into representative cells of model plans using all characteristics and criteria having a material impact on the size of the reserve”. A recent study on contract grouping using functional data analysis can be found in [Gan and Lin, 2015].

**AG-43 aggregate reserve:**

Details of the most recent version of the AG-43 can be found in [APP, 2015, Appendix C - Actuarial guidelines - Volume II, C-161]. Here is a summary of the key steps of determining AG-43 reserves. The aggregate reserve for all contracts falling within the scope of the AG-43 is equal to the greater of (1) the standard scenario amount and (2) the CTE amount. The standard scenario amount is the sum of reserves determined by applying a Standard Scenario Method to each individual contract. The Standard Scenario Method is in essence similar to the Stochastic Scenario Method described above, except for some technical treatments of surrender charge amortization and that the calculations in the above-mentioned three sections are done only under a single scenario (typically sharp losses in financial return in early years and modest returns in later years), which is prescribed by the AG-43 regulation. As it does not involve any stochastic component, we do not discuss the Standard Scenario Method here in details. One should keep in mind that the main purpose of the standard scenario amount is for the regulator to put a floor on AG-43 reserves to prevent a understatement of reserves by using the CTE amount under stochastic projections.

For VA contracts with GMDB only, which is generally perceived as a relatively low-risk rider, a company can choose to use the Alternative Methodology described in the AG-43 in place of the stochastic reserving approach. Since the Alternative Methodology is a factor-based formula approach, which does not involve stochastic components, we will not discuss it in this paper.

**Allocation of the aggregate reserves to the contract level:**

If the aggregate reserve is equal to the standard scenario amount, then the reserve allocated to each contract should be the reserve calculated for each contract under the Standard Scenario Method. If the reserve is equal to the CTE amount, then the reserve calculated for each contract should be the reserve calculated for each contract under the Standard Scenario Method, plus an allocation of the excess of the aggregate reserve over the standard scenario amount. The allocation should be made in proportion to the reserve for each contract under the Standard Scenario Method.

**Difficulty with simulations**

As shown in the construction of spreadsheets, carrying out simulations and completing subsequent income statements require intensive computations. The results may sometimes be unreliable due to sampling variability even with the aid of variance reduction techniques. Besides the potential problems with accuracy and efficiency, the simulation approach may also be undesirable from a cost and benefit perspective.

According to the 2008 SOA report [Farr et al., 2008], “76% of the respondents in the survey whose companies have more than $10 billion of annual revenue use a form of the stochastic approach. In contrast, only 27% of the respondents whose companies have less than 1 billion use a form of the stochastic approach.” This clearly indicates that the implementation of simulation approach
imposes prohibitive costs on smaller and medium size insurers. Even the more resourceful major insurers are often “forced to find a balance between accuracy (e.g. number of scenarios, product features, level of aggregation), expenses (e.g. computation hardware, staff), and timeliness of delivery (e.g. run time, time to analyze results and present results)”, according to the same survey.

4 Mathematical formulation: an average model

Although one may not realize this, the above-mentioned spreadsheet calculation is in fact based on a difference equation from a mathematical point of view, which is concisely expressed in the recursive relation (3.1). We shall investigate the continuous-time analogue of (3.1), which defines a differential equation, for several reasons. First, bear in mind that practicing actuaries often make a compromise to use a modestly small valuation period such as a quarter in order to simplify the calculations, even though fee incomes are collected on a daily basis. Therefore, it actually makes the calculation more accurate to investigate smaller valuation periods. Second, we can use the differential equation to identify a stochastic representation of the net liability, enabling us to determine its risk measures by the connection between stochastic differential equations and partial differential equations (PDE). Third, we can make a clear comparison of the spreadsheet calculations and the individual model presented in the previous section, which sheds light on their different treatment of mortality risks.

Dividing both sides of (3.1) by \( \Delta t \) gives

\[
\frac{U_{t+\Delta t} - U_t}{\Delta t} = m_d \tau t + \Delta t p_x F_{t+\Delta t} + \frac{F_t - \Delta t q_{x+t}}{\Delta t} \cdot p_x (G - F_{t+\Delta t})_+ + (G - F_t)_+ + rU_t.
\]

Letting the time difference \( \Delta t \) (a quarter in previous example) shrink to zero yields the differential equation

\[
U'_t = m_d \tau t p_x F_t - \mu_x + t p_x (G - F_t)_+ + rU_t.
\] (4.1)

Using the method of integrating factors, we obtain the following representation of the present value of the accumulated surplus

\[
e^{-rt}U_t = \int_0^t e^{-rs} m_d \tau s p_x F_s \, ds - \int_0^t e^{-rs} \mu_x + s p_x (G - F_s)_+ \, ds,
\]

or in other words, the present value of the accumulated deficiency, which we denote by \( L^*_t \),

\[
L^*_t := \int_0^t e^{-rs} \mu_x + s p_x (G - F_s)_+ \, ds - \int_0^t e^{-rs} m_d \tau s p_x F_s \, ds.
\] (4.2)

Note that in the average model the only source of randomness arises from the financial returns. One should note that the present value of accumulated surplus up to maturity \( L^*_{T} \) in the average model is in fact the conditional expectation of \( L \) from the individual model (2.3) on the natural filtration of asset price process (the market information up to time \( T \)).

\[
\mathbb{E}[L|\mathcal{F}_T] = \mathbb{E} \left[ e^{-r\tau z}(G - F_{\tau z})_+ I(\tau z < T) - \int_0^{\tau z \wedge T} e^{-rs} m_d \tau F_s \, ds \mid \mathcal{F}_T \right]
\]

\[
= \int_0^T e^{-rs}(G - F_s)_+ \, d\mathbb{P}(\tau z \in s) - \int_0^T \int_0^s e^{-ru} m_d \tau F_u \, d\mathbb{P}(\tau z \in s) - \mathbb{P}(\tau z > T) \int_0^T e^{-rs} m_d \tau F_s \, ds
\]

\[
= \int_0^T e^{-rs} s p_x \mu_x + s (G - F_s)_+ \, ds - \int_0^T (u p_x - \tau p_x) e^{-ru} m_d \tau F_u \, du - \tau p_x \int_0^T e^{-rs} m_d \tau F_s \, ds
\]

\[
= \int_0^T e^{-rs} s p_x \mu_x + s (G - F_s)_+ \, ds - \int_0^T \tau p_x e^{-rs} m_d \tau F_s \, ds = L^*_T,
\]
where we exchange the order of integration in the last equality.

We make a brief comment on the key distinction between the individual model and the average model. Suppose there are a total of $n$ policyholders of the same age to which the same death benefit has been issued. Let us assume that the $i$-th individual makes an initial purchase payment of $F_0^{(i)}$ and denote the insurer’s net liability in (2.3) by $L^{(i)}$. Assume that (A1) the future lifetimes of all policyholders are mutually independent and identically distributed and that (A2) all contracts are of equal size, i.e. $F_0^{(i)} = F_0^{(j)}$ for all $i, j = 1, \ldots, n$. Then a true aggregate model for the whole block of the GMDB business is given by

$$
\sum_{i=1}^{n} L^{(i)}.
$$

Note that all policyholders’ accounts $F_t^{(i)} = F_0^{(i)} e^{\mu^* t + \sigma B_t}$ are correlated due to the equity-linking mechanism. It is shown in [Feng and Shimizu, 2015] that, despite the complex dependency of $(L^{(1)}, L^{(2)}, \ldots, L^{(n)})$, the tail risk of the average diminishes as the number of policyholders increases.

In other words, for any integer $n > m$ and any $p \in (0, 1)$

$$
\text{CTE}_p \left( \frac{1}{n} \sum_{i=1}^{n} L^{(i)} \right) \leq \text{CTE}_p \left( \frac{1}{m} \sum_{i=1}^{m} L^{(i)} \right).
$$

Due to the continuity of coherent risk measure, we obtain that as $n \to \infty$,

$$
\frac{1}{n} \sum_{i=1}^{n} L^{(i)} \longrightarrow L^*_T, \quad \text{almost surely.} \tag{4.3}
$$

This means that when the number of i.i.d. policies is large enough, the average of the insurer’s net liability for each contract is roughly the net liability in an average model.

In this continuous-time model, the greatest present value of accumulated deficiencies (GPVAD) with analogy to (3.2) is given by the running supremum of net liabilities,

$$
M_T := - \inf_{0 \leq t \leq T} \{ e^{-rt} U_t \} = \sup_{0 \leq t \leq T} \{ L^*_t \}. \tag{4.4}
$$

Thus, in the continuous-time model, the quantile risk measure is determined by

$$
\text{VaR}_p(M_T) := \inf \{ V : \mathbb{P}(M_T \leq V) = p \}.
$$

After obtaining the quantile risk measure, we can determine the conditional tail expectation by

$$
\text{CTE}_p(M_T) := \mathbb{E}[M_T | M_T > \text{VaR}_p(M_T)].
$$

5 Drawbacks of AG-43 reserve and suggestions

In Sections 2 – 4, we have formulated mathematically both the individual model and the average model, which is the basis of the AG-43 reserve. The key distinction between the two models lies in the economic interpretations of conditional tail expectations.

\footnote{One can replace the assumption of fixed identical amount by a more relaxed assumption that (A2') $F_0^{(i)}$ are i.i.d. In this case, the process $\{F_t, t \geq 0\}$ in the formulation (4.2) should be interpreted as account values of an average account and $F_0$ should be equal to $\mathbb{E}(F_0^{(i)})$. Then the strong law of large numbers (4.3) still holds in this case.}
In the individual model, the risk measure $\text{CTE}_p(L)$ with $L$ defined in (2.3) determines the mean of losses from an individual contract that exceed the $p$-percentile.

In the average model, the risk measure $\text{CTE}_p(L^*_T)$ with $L^*$ defined in (4.2) determines the mean of losses from an average contract for a group of contracts with the same characteristics (i.e. maturity, benefits, the same distribution of initial purchase payments, etc.) that exceed the $p$-percentile.

In the AG-43 regulation, the risk measure $\text{CTE}_p(M_T)$ with $M$ defined in (4.4) determines the mean of maximum losses over time from an average contract for a group of contracts with the same characteristics that exceed the $p$-percentile.

5.1 Diversification of mortality risk

One should be reminded that the validity of the average model relies on two assumptions (A1) and (A2) or (A2'), which in essence guarantees the full diversification of mortality risk. In general, we can use the standard practice of contract grouping to justify the assumption (A1) where contracts with the same benefits and whose policyholders have similar ages are put into the same group. However, unlike exchanges-traded options, VA contracts do not have a standardized contract size and hence (A2) is unlikely to be true in reality. In order to justify the average model underpinning the AG-43 regulation, one would have to resort to the assumption (A2'), which means the chances of choosing from a spectrum of purchase payments are the same for all policyholders.

5.2 Allocation to the contract level

The most noticeable drawback of the AG-43 regulation is the allocation of aggregate reserves to individual contracts. Note that the aggregate reserve is the sum of the CTE amount ($\text{CTE}_{70\%}(M_T)$) and aggregate starting asset amounts for all groups in the whole block of business, unless the standard scenario amount is greater. The allocation of aggregate reserves to individual contracts in proportion to their standard scenario amounts appears to be based on expert opinions, rather than probabilistic reasoning. Because the CTE does not have additive property, this approach of aggregating reserves from all groups and then allocating to individual contracts no longer guarantees that the contract reserves are sufficient to cover the contract liabilities with the confidence level of 70%. If the purpose of reserving is to ensure the confidence level of 70% at the group level, then the allocation of reserves to individual contracts may be unnecessarily complicated. The allocation of reserves based on the seriatim standard scenario amount is just as arbitrary as a simple allocation of reserves made in proportion to the contract size. The extra computational burden arising from the computation of standard scenario amount does not seem to offer any more probabilistic insight than the simple allocation by contract size, which requires virtually no computational efforts.

There are several ways in which one can address the reserving problem at the contract level. One approach is to use the CTE amount from the individual model as the contract reserve. It has been shown in [Feng, 2014] that there are a variety of analytic/semi-closed form solution to the CTE under the individual model, which can be very efficient. Then one can claim the sufficiency of reserves at any confidence level for each individual contract. Another approach is to use the no-arbitrage value of the embedded option with the guaranteed benefits, for which analytic solutions are also available in the literature, as discussed in Section 1. Not only do both approaches better represent the true risks at the contract level, the computations can also be much more efficient than the current AG-43 methods.
5.3 GPVAD - do we really need it?

An argument for using risk measures of the GPVAD in (4.4) to determine statutory reserve is that the line of business may not be able to recover from the most severe loss and hence this adverse economic scenario is best represented and quantified by the greatest present value of accumulated deficiencies. However, in order to find risk measures of the GPVAD, we have to store simulated data for the entire projection period and use additional resources to rank all present values of accumulated deficiency.

Empirical evidence seems to suggest that the process $L^*$ often has monotonic sample paths. For illustration, we generated twenty sample paths in Figure 2, the parameter assumptions of which can be found in Section 7. This can be explained by the structure of $L^*$ in (4.2). Observe that the force of mortality $\mu_{x+s}$ and the rider charge $m_d$ are roughly on the same magnitude. In the extreme cases where the process $\{F_t, t \geq 0\}$ is close to zero, the first integral term in (4.2) is the dominating term, which increases as time lapses. On the other end of the spectrum, where $\{F_t, t \geq 0\}$ exceeds $G$ for a prolonged period of time, the second term in (4.2) is the dominating term, which decreases over time. Hence, we observe the presence of monotonicity on the two ends of the spectrum of sample paths in Figure 2. Bear in mind that statutory reserves are typically determined by risk measures on the right tail of net liabilities (positive liabilities/losses), which are represented by

![Sample paths of GMDB net liabilities](image)
the sample paths at the top part of Figure 2. In the extreme positive cases, sample paths of net liabilities tend to increase over time and hence the greatest present values almost always occur at the end of the projection length.

This raises the question whether it is really necessary to determine the greatest present value of accumulated deficiency over all valuation periods. If the all time high of accumulated deficiency does not lead to significant changes in the CTE amount, it would perhaps be a more efficient use of computational resources to simply focus on the present value of deficiency at the end of the projection. Therefore, we suggest that one use $\text{CTE}_p(L^*_T)$ in place of $\text{CTE}_p(M_T)$ as the basis of reserve. Note that in the case of $\text{CTE}_p(L^*_T)$ only terminal values need to be stored when we perform reserving exercises based on Monte Carlo simulations, for which the computational requirements can be reduced.

Despite our preference of $\text{CTE}_p(L^*_T)$ over $\text{CTE}_p(M_T)$ for computational reasons, we shall present PDE solutions to both risk measures in Section 6. A comparison with distribution functions of $L^*_T$ and $M_T$ is shown by a numerical example in Section 7.

6 Solution Method: PDE

In the current literature, there is no alternative technique available other than Monte-Carlo based statistical methods for computing the AG-43 reserve. While a PDE method for risk measures of guaranteed minimum maturity benefit was presented in [Feng, 2014] under the average model, we provide an extension to the GMDB in this paper. It should be noted, however, such an extension is not trivial due to the complication of the running supremum in (4.4). Here we take the calculation of quantile risk measure VaR as an example; the extension to the conditional tail expectation can be done in a similar manner. Since $M_T$ is a continuous random variable, we consider the inverse problem of finding $V := \text{VaR}_p$ such that

$$\mathbb{P}(M_T \leq V) = p.$$  

This calculation of VaR relies on the evaluation of the distribution function of $M_T$ on any valuation date $t$ after policy issue and before maturity, $0 \leq t \leq T$. Without loss of generality, we assume that $t = 0$, as the time-$t$ valuation can be obtained by shifting the time line. Consider the distribution function of the GPVAD given by

$$P(T, M) := \mathbb{P}(M_T \leq M).$$  

For the ease of presentation, we introduce the function $a(t, x)$, which represents the incremental change in the present value of surplus (benefit outgo less fee income) at time $t$ when the underlying Brownian motion hits level $x$.

$$a(t, x) = e^{-rt} \mu x + tp_x (G - F_0 e^{\sigma x + \mu^* t}) - m_d e^{-rt} tp_x F_0 e^{\sigma x + \mu^* t}.$$  

We have shown in the proof of Proposition 6.1 in Appendix A that the distribution function of the GPVAD can be determined by

$$P(T, M) = h(0, 0, M),$$  

where $h$ satisfies the following PDE. Since the survival probabilities are usually extracted from life tables, we use fractional age assumption to determine the values of $a(t, x)$ in Appendix B. Sample finite difference algorithms for solving the PDEs shall be posted on the authors’ personal webpages.
Proposition 6.1. For $0 < t < T, x \in \mathbb{R}, y \in (0, \infty)$, $h(t, x, y)$ is a solution to the PDE

$$\frac{\partial}{\partial t} h(t, x, y) + a(T - t, x) \frac{\partial}{\partial y} h(t, x, y) = \frac{1}{2} \frac{\partial^2}{\partial x^2} h(t, x, y),$$

(6.1)

subject to the initial and boundary conditions

$$h(0, x, y) = 1;$$  \hspace{1cm} (6.2a)

$$h(t, x, 0) = 0;$$  \hspace{1cm} (6.2b)

$$\lim_{y \to \infty} h(t, x, y) = 1;$$  \hspace{1cm} (6.2c)

$$\lim_{x \to \infty} h_x(t, x, y) = 0;$$  \hspace{1cm} (6.2d)

$$\lim_{x \to -\infty} h_x(t, x, y) = 0.$$  \hspace{1cm} (6.2e)

To illustrate the point we made in Section 5.3 that the tail risk of $M_T$ is very similar to that of $L^*_T$, we can also calculate the distribution function of the present value of accumulated deficiency at the end of projection,

$$Q(T, M) := \mathbb{P}(L^*_T \leq M) = k(0, 0, M),$$

where $k$ satisfies a similar PDE in Proposition 6.2.

Proposition 6.2. For $0 < t < T, x \in \mathbb{R}, y \in \mathbb{R}$, $k(t, x, y)$ is a solution to the PDE

$$\frac{\partial}{\partial t} k(t, x, y) = a(T - t, x) \frac{\partial}{\partial y} k(t, x, y) + \frac{1}{2} \frac{\partial^2}{\partial x^2} k(t, x, y),$$

(6.3)

subject to the initial and boundary conditions

$$k(0, x, y) = I(y < M);$$  \hspace{1cm} (6.4a)

$$\lim_{y \to -\infty} k(t, x, y) = 1;$$  \hspace{1cm} (6.4b)

$$\lim_{y \to \infty} k(t, x, y) = 0;$$  \hspace{1cm} (6.4c)

$$\lim_{x \to \infty} k_x(t, x, y) = 0;$$  \hspace{1cm} (6.4d)

$$\lim_{x \to -\infty} k_x(t, x, y) = 0.$$  \hspace{1cm} (6.4e)

We shall illustrate with a numerical example in Section 7 that $P$ and $Q$ are actually very close to each other for practical purposes.

7 Numerical examples

We first develop a case to test the accuracy of solutions to the PDEs in Proposition 6.1. We consider a special case of the net liability process (4.2), where

$$F_0 = 1, \quad m_d = -1, \quad s_{p_x} \equiv 1, \quad \text{for } 0 \leq s \leq T.$$

In this case, by the scaling property of Brownian motion,

$$e^{-rT} F_T = d e^{2(\nu s + B_s)}; \quad L^*_T = c A_s := c \int_0^s e^{2(\nu t + B_t)} \, dt,$$

16
where \( \{A_s, s \geq 0\} \) is well-studied in financial literature and sometimes referred to as Yor’s process,

\[
\nu := \frac{4(\mu - r)}{\sigma^2}, \quad s := \frac{\sigma^2 T}{4}, \quad c := \frac{4}{\sigma^2}.
\]

Since \( L_t^* \) in this case is an increasing process, it is trivial that \( M_t \equiv L_t^* \) for all \( 0 \leq t \leq T \) given that \( M_0 = L_0 = 0 \). In this case, the function \( h \) defined in (A.1) is given by

\[
h(t, x, y) = \mathbb{P}\left( y + ce^{2\nu \tau} \int_0^{s-\tau} e^{2(\nu u + x + B_u)} \, du < M | B_0 = 0 \right)
\]

\[
= \mathbb{P}\left( \int_0^{s-\tau} e^{2(\nu u + B_u)} \, du < \frac{M - y}{ce^{2(\nu \tau + x)}} \right),
\]

where \( 0 \leq \tau := 4t/\sigma^2 \leq s \). Therefore,

\[
h^*(t, x, y, M) = h(T - t, x, M - y) = \mathbb{P}(A_T < e^{-2(x + \nu(s-\tau))}y/c),
\]

where the distribution function is known (c.f. [Feng and Volkmer, 2015, (2.4)]) in the case of \( \nu > 0 \) to be

\[
\mathbb{P}(A_T \leq z) = \frac{1}{4\pi^2} (2z)^{(1+\nu)/2} \exp\left(-\frac{1}{4z}\right)
\]

\[
\times \int_0^\infty \exp\left(-\frac{(\nu u^2 + p^2)\tau}{2}\right) W_{-\frac{1+\nu}{4}} \left( \frac{1}{2z} \right) \left| \Gamma\left(\frac{\nu + i p}{2}\right) \right|^2 \sinh(\pi p)p \, dp,
\]

where \( W \) is the Whittaker function and \( \Gamma \) is the gamma function. This semi-closed-form solution can be computed very efficiently in most computational software packages, such as Maple and Mathematica, which provide built-in algorithms for special functions. We use two algorithms to solve for the PDE in Proposition 6.1, one of which employs implicit method in \( x \)-dimension and the other employs explicit method in \( x \)-dimension. Both algorithms use the upwind method in \( y \)-dimension. We then benchmark the results from the PDE algorithms against those from the exact solutions from (7.1). Our experiment show that the two methods produce the same results up to three decimal places.

In the second case, we compute the distribution functions of \( M_T \) and \( L_T^* \) using the PDE methods in Propositions 6.1 and 6.2. For comparison with the individual model, we use precisely the same valuation basis as used in [Feng and Volkmer, 2012]. Consider a variable annuity contract issued to a policyholder of age 65 with GMMB and GMDB riders. The length of projection for the variable annuity contract is 10 years, i.e. \( T = 10 \). The valuation is based on the geometric Brownian motion model (2.1) with \( \mu - \sigma^2/2 = 0.09 \) and \( \sigma = 0.3 \) per annum. The discount rate, annualized fees/charges, GMDB rider charges are given by \( r = 0.04, m = 0.01, \) and \( m_d = 0.0035 \) per annum respectively. The guaranteed minimal payment for the GMDB is set at the full refund of initial purchase payment, i.e. \( G = F_0 \). The probability model of survivorship is extracted from the period life table for male and calendar year 2010 published by the U.S. Social Security Administration [Bell and Miller, 2005, page 68], which is reiterated in Table 4.

In Table 5, we provide the results on the distribution functions \( P(T, M) = \mathbb{P}(M_T < M) \) and \( Q(T, M) = \mathbb{P}(L_T^* < M) \). When truncating the state space in (6.1), we set \( b = 10 \) and \( c = 0.4 \). The grid size is given by \( \Delta t = 0.0005, \Delta x = 0.1, \Delta y = 0.001 \). The data suggests that the insurer’s net liability in the average model has a very thin tail. A quick comparison of \( P \) and \( Q \) confirms our earlier observation that the tail probability of \( M_T \) is very close to that of \( L_T^* \), due to the increasing sample paths of \( M_T \) in the extreme cases. Hence, for practical purposes, it may not be worth the
Table 4: Predicted mortality rates of a male at the age of 65

<table>
<thead>
<tr>
<th>x</th>
<th>$1q_x$</th>
<th>$k$</th>
<th>$k\bar{p}_{65}$</th>
<th>x</th>
<th>$1q_x$</th>
<th>$k$</th>
<th>$k\bar{p}_{65}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>0.01753</td>
<td>0</td>
<td>1.00000</td>
<td>71</td>
<td>0.03059</td>
<td>6</td>
<td>0.87275</td>
</tr>
<tr>
<td>66</td>
<td>0.01932</td>
<td>1</td>
<td>0.98246</td>
<td>72</td>
<td>0.03343</td>
<td>7</td>
<td>0.84606</td>
</tr>
<tr>
<td>67</td>
<td>0.02122</td>
<td>2</td>
<td>0.96348</td>
<td>73</td>
<td>0.03633</td>
<td>8</td>
<td>0.81778</td>
</tr>
<tr>
<td>68</td>
<td>0.02323</td>
<td>3</td>
<td>0.94304</td>
<td>74</td>
<td>0.03942</td>
<td>9</td>
<td>0.78807</td>
</tr>
<tr>
<td>69</td>
<td>0.02538</td>
<td>4</td>
<td>0.92113</td>
<td>75</td>
<td>0.04299</td>
<td>10</td>
<td>0.75700</td>
</tr>
<tr>
<td>70</td>
<td>0.02785</td>
<td>5</td>
<td>0.89775</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Distribution functions of GPVAD and PVAD at the end of projection

<table>
<thead>
<tr>
<th>$M$</th>
<th>$0.01$</th>
<th>$0.02$</th>
<th>$0.03$</th>
<th>$0.04$</th>
<th>$0.05$</th>
<th>$0.06$</th>
<th>$0.07$</th>
<th>$0.08$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(T,M)$</td>
<td>0.88733</td>
<td>0.92570</td>
<td>0.94785</td>
<td>0.96279</td>
<td>0.97358</td>
<td>0.98160</td>
<td>0.98760</td>
<td>0.99201</td>
</tr>
<tr>
<td>$Q(T,M)$</td>
<td>0.91735</td>
<td>0.93840</td>
<td>0.95377</td>
<td>0.96581</td>
<td>0.97530</td>
<td>0.98243</td>
<td>0.98831</td>
<td>0.99296</td>
</tr>
</tbody>
</table>

For each of the PDEs in Propositions 6.1 and 6.2, we use two algorithms as described above using both explicit and implicit methods. The algorithms are available at the authors’ personal webpages. Results from both algorithms agree up to at least three decimal places for each test of the distribution functions $P$ and $Q$. Results for both $P$ and $Q$ presented in Table 5 are generated by the algorithm using the implicit method in the $x$-dimension and the upwind method in the $y$-dimension.

8 Expenses, surrenders and withdrawals

The computational method presented in the earlier sections can be extended to include more complex cash flows. For example, we can consider the following items typical in insurer’s bookkeeping.

- $w$, the rate of scheduled withdrawals. It is typical that policyholders withdraw a fixed amount determined by a certain percentage of initial purchase payment.

- $G_t$, the time-varying guarantee level at time $t$. For example, the GMDB with a roll-up option provides the guarantee that accrues compound interest. Therefore, the instantaneous change in the guarantee level is given by $dG_t = \delta G_t \, dt - w \, dt$, which implies that $G_t = e^{\delta t} G_0 - w (e^{\delta t} - 1)/\delta$.

- $E_t$, the rate of expenses at time $t$ such as commissions, policy-related costs, overheads, etc.

- $c_t$, the rate of surrender charges, which is typically scheduled to decline with the time elapsed to discourage early surrenders and are represented as percentages of policyholders’ account values.

- $\mu_t$, the lapse rate, which typically increases with time and peaks in the period when the surrender charge reduces to zero;

Suppose that the equity index/fund is driven by a geometric Brownian motion as in (2.1). The rest of the notation is the same as in Section 2. For example, the rate of M&E fee is $m$ per time
unit and the rate of rider charge allocated to the GMDB is $m_d$ per time unit. The policyholder’s account value is determined by

$$dF_t = [(\mu - m)F_t - w] \, dt + \sigma F_t \, dB_t.$$  

With analogy to (4.1), the instantaneous change of accumulated surplus is determined by

\[
\begin{align*}
U_t' &= rU_t + e^{-rt}m_d \, t p_x F_t \\
&\quad + \mu_l t p_x c_t F_t \quad \text{(interest on surplus & M&E fee income)} \nonumber \\
&\quad - \mu_{x+t} t p_x (G_t - F_t) \quad \text{(death benefit)} \\
&\quad - E_t. \\
\end{align*}
\]

Then it follows that the present value of the accumulated deficiency is given by

$$L^*_t := \int_0^t e^{-rs} \left[ \mu_{x+s} s p_x (G_s - F_s) \right] \text{d}s + E_s - m_d s p_x F_s - \mu_{s} s p_x c_s F_s.$$  

Using arguments similar to those in Appendix 8, similar PDEs can be derived as in Propositions 6.1 and 6.2 with precisely the same boundary conditions. PDE (6.1) should be replaced by

$$\frac{\partial}{\partial t} h(t, x, y) + a(T - t, x) \frac{\partial}{\partial y} h(t, x, y) + \frac{\sigma^2}{2} x \frac{\partial^2}{\partial x^2} h(t, x, y) + [(\mu - m)x - w] \frac{\partial}{\partial x} h(t, x, y),$$

where

$$a(t, z) = e^{-rt} \left[ \mu_{x+t} t p_x (G_t - z) \right] + E_t - m_d t p_x z - \mu_t t p_x z.$$  

PDE (6.3) should be replaced by

$$\frac{\partial}{\partial t} k(t, x, y) = a(T - t, x) \frac{\partial}{\partial y} k(t, x, y) + \frac{\sigma^2}{2} x \frac{\partial^2}{\partial x^2} k(t, x, y) + [(\mu - m)x - w] \frac{\partial}{\partial x} k(t, x, y).$$

**Appendix A**

**Proof of Proposition 6.1:**

Here we shall utilize a standard martingale method (c.f. the pricing of the look-back option in [Shreve, 2004, Section 7.4]) to derive a PDE satisfied by the distribution function. Due to the strong Markov property of the triple $(B_t, L^*_t, M_t)$, there must exist a sufficiently smooth function $v(t, x, y, z) = \mathbb{P}(M_T < M|B_t = x, L^*_t = y, M_t = z)$, such that $v(t, B_t, L^*_t, M_t)$ is a martingale. We observe that

\[
\begin{align*}
v(t, x, y, z) &= \mathbb{E} \left[ I \left( M_t + \left( \sup_{t \leq s < T} L^*_s - M_t \right) \right) < M \right] \bigg| \bigg| B_t = x, L^*_t = y, M_t = z \\
&= \mathbb{E} \left[ I \left( z + \left( \sup_{t \leq s < T} L^*_s - z \right) \right) < M \right] \bigg| \bigg| B_t = x, L^*_t = y \\
&= I(z < M) \mathbb{P} \left( \sup_{t \leq s < T} L^*_s < M \bigg| \bigg| B_t = x, L^*_t = y \right),
\end{align*}
\]
where we used the fact that $I((x)_+ < y) = I(x < y)I(y > 0)$ in the last equality. Note that the conditional probability on the right-hand-side has no dependency on $z$. We define for $x \in \mathbb{R}$ and $y \in (-\infty, M)$

$$h(t, x, y) := \mathbb{P} \left( \sup_{t \leq s < T} L_s^* < M \left| B_t = x, L_t^* = y \right. \right). \quad (A.1)$$

Hence we observe separation of variables by

$$v(t, x, y, z) = I(z < M)h(t, x, y). \quad (A.2)$$

Note that both $L_t^*$ and $M_t$ are processes of finite variation, for which all quadratic variations and cross variations with other processes are zero. In view of (4.2), we apply Itô’s formula to obtain

$$dh(t, B_t, L_t^*) = h_t(t, B_t, L_t^*) \, dt + h_x(t, B_t, L_t^*) \, dB_t + \frac{1}{2} h_{xx}(t, B_t, L_t^*) \, dB_t + h_y(t, B_t, L_t^*) \, dL_t^* + \frac{1}{2} h_{yy}(t, B_t, L_t^*) \, dL_t^* \, dB_t$$

By definition, $h(t, B_t, L_t^*)$ is a martingale. Therefore, we must have

$$\frac{\partial}{\partial t} h(t, x, y) + \frac{1}{2} \frac{\partial^2}{\partial x^2} h(t, x, y) + a(t, x) \frac{\partial}{\partial y} h(t, x, y) = 0,$$

where

$$a(t, x) = e^{-rt} \mu_{x+t} t p_x (G - F_0 e^{\sigma x + \mu^* t}) + m_d e^{-rt} t p_x F_0 e^{\sigma x + \mu^* t}.$$

The terminal condition is given by

$$h(T, x, y) = \mathbb{P} (L_T^* < M | B_T = x, L_T^* = y) = I(y < M). \quad (A.4)$$

We also have

$$h(t, x, M) = \mathbb{P} \left( \sup_{t \leq s < T} L_s^* < M \left| B_t = x, L_t^* = M \right. \right) = 0. \quad (A.5)$$

$$\lim_{y \to -\infty} h(t, x, y) = \lim_{y \to -\infty} \mathbb{P} \left( \sup_{t \leq s < T} L_s^* < M \left| B_t = x, L_t^* = y \right. \right) = 1. \quad (A.6)$$

To develop the corresponding forward PDE, we set $t^* = T - t, y^* = M - y$ and $h^*(t^*, x, y^*) = h(t, x, y)$. We then drop the asterisks for notational brevity. Hence, we obtain the PDE (6.1). The initial and boundary conditions follow immediately from the changes of variables in (A.4), (A.5) and (A.6).

**Proof of Proposition 6.2:**

Owing to the strong Markov property of the pair $(B_t, L_t^*)$, there must exist a sufficiently smooth function

$$u(t, x, y) = \mathbb{P}(L_T^* < M | B_t = x, L_t^* = y),$$

20
such that \( u(t, B_t, L_t^t) \) is a martingale. Using nearly identical arguments as in the proof of Proposition 6.1, we obtain the PDE for \( t \in (0, T), x \in (-\infty, \infty), y \in (-\infty, \infty) \)

\[
\frac{\partial}{\partial t} u(t, x, y) + \frac{1}{2} \frac{\partial^2}{\partial x^2} u(t, x, y) + a(t, x) \frac{\partial}{\partial y} u(t, x, y) = 0,
\]

subject to the conditions

\[
\begin{align*}
    u(T, x, y) &= I(y < M); \\
    \lim_{y \to -\infty} u(t, x, y) &= 1; \\
    \lim_{y \to \infty} u(t, x, y) &= 0; \\
    \lim_{x \to -\infty} u_x(t, x, y) &= 0; \\
    \lim_{x \to \infty} u_x(t, x, y) &= 0.
\end{align*}
\]

We set \( k(t, x, y) = u(T - t, x, y) \) to obtain the PDE (A.7) with the stated conditions.

Appendix B: fractional age assumptions

B.1 Uniform distribution of deaths (UDD)

In practice, the mortality is often modeled by a life table so \( t p_x \) is known for integer-valued \( t \). In the period between two consecutive integers, various assumptions may be made, the most common being the uniform distribution of deaths (UDD). We write \( q_x = p_x \) for short. Under the UDD assumption, the number of deaths in each given period is uniformly spread out across the period, i.e. for any \( 0 \leq s < 1 \),

\[
s q_x = s q_x.
\]

It is easy to show that for \( 0 \leq s < 1 \),

\[
s p_x \mu_{x+s} = q_x.
\]

Let \( \lfloor t \rfloor \) be the integer part of \( t \) and define \( \lceil t \rceil = \lfloor t \rfloor + 1 \). Note that if \( t \) is an integer then \( \lceil t \rceil = t + 1 \). Therefore, for \( t \geq 0 \),

\[
\mu_{x+t} t p_x = \lfloor t \rfloor p_x q_{x+\lfloor t \rfloor}, \quad t p_x = \lfloor t \rfloor p_x \left( 1 - (t - \lfloor t \rfloor) q_{x+\lfloor t \rfloor} \right).
\]

Thus, under the UDD assumption, the function \( a \) is piecewise smooth in \( t \),

\[
a(t, x) = e^{-rt} \lfloor t \rfloor p_x q_{x+\lfloor t \rfloor} (G - F_0 e^{u^\ast t + \sigma x} \right) - m_d e^{-rt} \lceil t \rceil p_x \left( 1 - (t - \lfloor t \rfloor) q_{x+\lfloor t \rfloor} \right) F_0 e^{u^\ast t + \sigma x}.
\]

B.2 Constant force of mortality (CF)

Now we consider the case where the force of mortality is constant between integer ages, i.e. for any \( 0 \leq s < 1 \),

\[
s p_x = p_x^s.
\]

It is easy to show that for \( 0 \leq s < 1 \),

\[
s p_x \mu_{x+s} = -p_x^s \ln(p_x).
\]
Therefore, for $t > 0$,

$$
\mu_{x+t} \cdot t p_x = -[t] p_x p_x^{t-[t]} \ln(p_{x+[t]}), \quad t p_x = [t] p_x p_x^{t-[t]}.
$$

Thus, under the CF assumption, the function $a$ is also piecewise smooth in terms of $t$,

$$
a(t, x) = -e^{-rt} [t] p_x p_x^{t-[t]} \ln([t] p_x)(G - F_0 e^{\mu^* t + \sigma x})_+ - m_d e^{-rt}([t] p_x p_x^{t-[t]}) F_0 e^{\mu^* t + \sigma x}.
$$

B.3 Continuous mortality model

It is also common to model mortality by analytical distributions. For example, Makeham’s law with a simple formula for force of mortality $\mu_x = A + Bc^x$ can provide a very close fit to empirical mortality data from middle age to advanced age. Since this is indeed the range of ages at which variable annuities policies are purchased, we can use Makeham’s law in our model as well.

$$
t p_x = e^{-At} g^{c^x(c-1)}, \quad g := \exp(-B/\ln c).
$$

Therefore, we have a continuous version of $a(t, x)$

$$
a(t, x) = e^{-(A+r)t} g^{c^x(c-1)} \left[(A + Bc^x)(G - F_0 e^{\sigma x + \mu^* t})_+ - m_d F_0 e^{\sigma x + \mu^* t}\right].
$$

Acknowledgement

The authors would like to thank Sun Feng, FSA, CERA, MAAA, a consulting actuary at Towers Watson, for his suggestions and pointing us to the AG-43 regulation and AAA practice notes. Huaxiong Huang is partially supported by the Nature Science and Engineering Research Council (NSERC) of Canada and the Fields Institute in Toronto.

References


Life Practice Note Steering Committee (2009). The application of C3 phase II and actuarial guideline XLIII. A public policy practice note, American Academy of Actuaries.


