Risk based capital for guaranteed minimum withdrawal benefit

Runhuan Feng & Jan Vecer

To cite this article: Runhuan Feng & Jan Vecer (2016): Risk based capital for guaranteed minimum withdrawal benefit, Quantitative Finance, DOI: 10.1080/14697688.2016.1189087

To link to this article: http://dx.doi.org/10.1080/14697688.2016.1189087

Published online: 07 Jun 2016.
Risk based capital for guaranteed minimum withdrawal benefit

RUNHUAN FENG† and JAN VECER*‡§

†Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 W. Green Street (MC-382), Urbana, IL 61801, USA
‡Faculty of Mathematics and Physics, Charles University, Sokolovska 83, 18675 Praha 8, Czech Republic
§Vysoka skola aplikovaneho prava, Chomutovicka 1443, 14900 Praha 4, Czech Republic

(Received 26 June 2015; accepted 5 May 2016; published online 7 June 2016)

The guaranteed minimum withdrawal benefit (GMWB), which is sold as a rider to variable annuity contracts, guarantees the return of total purchase payment regardless of the performance of the underlying investment funds. The valuation of GMWB has been extensively covered in the previous literature, but a more challenging problem is the computation of the risk based capital for risk management and regulatory reasons. One needs to find the tail distribution of the profit–loss function, which differs from its expected payoff required for pricing the GMWB contract. GMWB has embedded two option-like features: Management fees are proportional to the current value of the policyholder’s account which results in an average price of the account. Thus the contract resembles an Asian option. However, the fees are charged only up to the time of the account hitting zero which resembles a barrier option payoff. Thus the GMWB is mathematically more complicated than Asian or barrier options traded on the financial markets. To the authors’ best knowledge, this is the first paper in the literature to formulate and analyse profit–loss distribution using PDE methods of such a product with intricate option-like features. Our approach is much more efficient than the current market practice of rather intensive and expensive Monte Carlo simulations due to the lack of samples for extreme cases.

Keywords: Guaranteed minimum withdrawal benefit; Variable annuity; Risk based capital; Asian options; Partial differential equations; Finite difference

1. Introduction

Variable annuity is an insurance contract which allows policyholders to share financial returns from equity investment. Policyholders are typically offered a variety of options to invest their premium payments. The insurer transfers the funds to third-party vendors, who manage and service the accounts of the policyholders’ choosing. Beside paying up-front fees to initiate the contracts, policyholders’ accounts are subject to the deduction of the account-value-based fees, which are then distributed between the insurer and the third-party vendors. The variable annuity base contract is virtually a mutual fund.

The guaranteed minimum withdrawal benefit (GMWB) is among a number of investment guarantees introduced in the early 2000s. The GMWBs are typically sold as add-ons to variable annuity base contracts. Early versions of GMWBs permitted annual withdrawals of a certain percentage of the benefit base until the benefit base was exhausted, even if the policyholder’s account value itself had already fallen to zero. The benefit base was usually the sum of premium payments.

Later versions enhanced the benefit base to include step-ups or bonuses prior to withdrawals or optional step-ups to reflect investment growth after withdrawals have commenced. Details on market shares, policyholders’ utilizations of the GMWBs can be found in Drinkwater et al. (2013, 2014).

The technical challenges of modelling and pricing GMWBs have been well studied and addressed in the actuarial and finance literature. Milevsky and Salisbury (2006) were first to introduce the mathematical formulation of the GMWB and provided numerical PDE solutions for pricing. Feng and Volkmer (2016 forthcoming) developed analytical solutions to the pricing problem proposed by Milevsky and Salisbury from the policyholder’s point of view as well as the pricing from an insurer’s point of view. Donnelly et al. (2014) used methods of Vecer (2001) and Vecer and Xu (2004) to price GMWBs under stochastic interest rate and volatility models and also provided analytical approximations based on moment-matching. Dai et al. (2008) investigated optimal withdrawal strategies to maximize the option values from the policyholder’s perspective. The same HJB variational inequality was solved by a penalty method in Huang and Forsyth (2012). The work was later extended by Chen and Forsyth (2008) to allow discrete...
withdrawals and more complex reset features, again with the objective to optimize policyholders’ financial returns. Huang and Kwok (2014) developed analytic asymptotic formulas for policyholder’s financial returns and optimal withdrawal boundaries under various limits. A recent work by Peng et al. (2012) considered the pricing of GMWBs under stochastic interest rates. The optimal control of policyholders’ withdrawals was also extended in Forsyth and Vetzal (2014) to the guaranteed lifetime withdrawal benefit (GLWB), an enhanced version of the GMWB, under a regime switching model. Huang et al. (2014) studied the optimal initiation (earliest time of withdrawal) of the GLWB using an American option pricing framework. Kling et al. (2011) investigates how stochastic volatilities affect hedging errors of GMWBs priced under constant volatility assumptions.

As the purpose of this paper is to introduce a new technique to the calculation of risk metrics for GMWBs, we shall primarily focus on the plain-vanilla version of the GMWBs. Nevertheless, the PDE techniques presented in this work can be extended to accommodate more complex GMWBs. To the best of our knowledge, the previous scientific literature have not addressed finding the whole distribution of the profit and loss of the GMWB from the insurer’s point of view required for regulatory financial reporting. A straightforward approach is to use Monte Carlo techniques to simulate various price scenarios in order to obtain an approximate distributional characteristics of the profit and loss distribution. However, since fees are proportional to policyholder’s account values throughout the duration of the contract, one needs to simulate the entire path of the policyholder’s equity fund choices. There are two possible assumptions and their impacts on the riskiness of the insurance liability.

2. Definition of the GMWB risk management problem

2.1. Equity-linking mechanism

Let us formulate the risk management problem of the GMWB mathematically. Assume that the underlying equity index/fund of the policyholder’s choosing is driven by a geometric Brownian motion:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dB_t, \quad t > 0. \tag{2.1}$$

The process $B_t$ is a standard Brownian motion. The rider charges are usually made on a daily basis as a percentage of policyholder’s account values, called mortality and expenses fee (M&E), say $m$ per time unit. Hence it is natural to model the fees and charges by continuous deduction from the account. Note, however, one can also account for payments on discrete time points in this model framework, as done in Forsyth and Vetzal (2014). Without any withdrawals, the policyholder’s investment account evolves in proportion to the equity index/fund. Hence the dynamics of the policyholder’s account values $[F_t, t \geq 0]$ is determined by

$$F_t = F_0 \frac{S_t}{S_0} e^{-mt}, \quad t \geq 0,$$

which satisfies the stochastic differential equation

$$dF_t = (\mu - m) F_t \, dt + \sigma F_t \, dB_t, \quad F_0 > 0. \tag{2.2}$$

We assume that the policyholder chooses to withdraw only at the maximum rate without penalty, which is the majority of cases in practice and has been a commonly used assumption in practitioners’ simulations. Empirical data to support this claim can be found consistently in Drinkwater et al. (2013, 2014). Denote the withdrawal rate per time unit by $w$. Thus, the dynamics of account value is given by

$$dF_t = (\mu - m) F_t \, dt - w \, dt + \sigma F_t \, dB_t, \quad F_0 > 0. \tag{2.2}$$

Note that the withdrawal rate is typically a fixed percentage of the benefit base $G$. In this work, the benefit base $G$ is equal to $F_0$, meaning that the policyholder is guaranteed to receive a full refund of his or her premium payments. Bear in mind that withdrawals are often taken monthly or quarterly in practice. Nevertheless, the continuous-time model can serve as a good approximation.

Let us illustrate the GMWB contract by the following example. A contract starts with an initial purchase payment of $100 and the policyholder elects to withdraw the maximum amount 7% of purchase payment without penalty each year. The GMWB contract guarantees that its holder can withdraw $7 annually until the entire purchase payment is returned, which means the annual withdrawals are protected until the maturity at the end of $100/7 \approx 14.28$ years regardless of the policyholder’s equity fund choices. There are two possible scenarios what may happen to the underlying fund.

1. The fund may have a positive balance at the end of the guaranteed period of withdrawals. The policyholder is entitled to this positive excess and thus has a positive upside of this investment. In this case, the insurer’s profit and loss ends up in profit as all withdrawals are funded by the policyholder’s own account and the insurer merely collects the fees for the entire duration of the contract and has no obligation to pay anything out of pocket at any point of the contract.

2. When the underlying equity index/fund performs poorly, the policyholder’s account may get depleted earlier and the insurer has to pay out the guaranteed withdrawal amount until the total amount reaches the policyholder’s
purchase payment. Moreover, the fund fees are not collected after the fund is depleted. For example, consider that the account is depleted at the end of five years. By this time, the policyholder would have only withdrawn $7 \times 5 = $35 in total. Then the guarantee kicks into sustain the annual withdrawal $7 until the entire purchase payment is returned at the end of $100/7 = 14.28$ years. This may result in a loss for the insurer as the collection of fees prior to the premature account depletion may not be sufficient to cover the cost of sustaining the withdrawals after the account depletion.

2.2. Insurer’s liabilities

The amount of time needed for the policyholder to recoup the original premium payment $G = F_0$ by withdrawing $w$ per time unit is $T = G/w$. Clearly, there is no financial obligation to the insurer until the time at which the account value hits zero, i.e.

$$
\tau_0 = \inf \{t : F_t \leq 0\}. \tag{2.3}
$$

It is only when the account value is depleted prior to the maturity $T$ that the maximum withdrawal rate $w$ is paid at the cost of the insurer. Therefore, the present value of the cost to an insurer of the GMWB rider (gross liability) is thus given by

$$
\int_{\tau_0}^{T} e^{-rs} w \, ds, \tag{2.4}
$$

where $r$ is the yield rate on the assets backing up the liability and $x \wedge y = \min\{x, y\}$. Recall that to compensate for its liability for the GMWB rider, the insurer receives the distribution of fees from the third-party fund manager, which are often a fixed percentage of the policyholder’s account until the account value hits zero. We denote the portion of fees that are used to fund the GMWB rider by $m_w$. Note that the total fee $m$ is in general larger than $m_w$, as the rest goes to cover overheads, commissions and other expenses. Thus the accumulated present value of the fee income is given by

$$
\int_{\tau_0}^{T} m_w e^{-rs} F_t \, ds. \tag{2.5}
$$

From the perspective of risk analysis, we are interested in the insurer’s net liability—the gross liability less fee income

$$
L = \int_{\tau_0}^{T} e^{-rs} w \, ds - m_w \int_{0}^{\tau_0} e^{-rs} F_t \, ds. \tag{2.6}
$$

The net liability is in general expected to be negative in most cases, as the product is designed to be profitable. Nevertheless, it is the likelihood and severity of the positive side of net liability (loss) that is of the most interest from the viewpoint of risk management.

2.3. Risk measures

The general principle of risk management is to determine sufficient funds to set aside in order to absorb unexpected losses in adverse economic scenarios. Here we model the essence of a risk management strategy that determines risk capitals by certain measures of insurance net liabilities. The two most common risk measures used by practitioners and regulatory bodies around the world are quantile risk measure, also known as Value-at-Risk (VaR), and conditional tail expectation (CTE) risk measure, also known as conditional VaR or tail VaR in the literature.

In the US market, the National Association of Insurance Commissioners (NAIC) published model regulations, which are essentially adopted by all insurance regulators at the state level, that requires variable annuity writers to use 70% CTE to determine the reserves and 90% CTE to determine the risk based capital (RBC). Although statutory financial reporting, including reserving and RBC requirement, can be very complex and tedious with detailed accounting rules, we modelled the essence of the insurance liability of the GMWB in Equation (2.6) and hence use the corresponding risk measures to quantify and assess the financial risk of the guaranteed benefit.

The quantile risk measure for $L$ is defined for $\alpha$ ($0 \leq \alpha \leq 1$) as

$$
\text{VaR}_\alpha := \inf \{y : \mathbb{P}[L \leq y] \geq \alpha\}. \tag{2.7}
$$

Since $L$ is modelled by continuous random variables in this model, we can compute the quantile risk measure by $\text{VaR}_\alpha$ using a root search algorithm such that

$$
\mathbb{P}[L > \text{VaR}_\alpha] = 1 - \alpha. \tag{2.8}
$$

The quantile risk measure $\text{VaR}_\alpha$ is interpreted as the minimum capital required to ensure that there is sufficient fund to cover future liability with the probability of at least $\alpha$.

The CTE risk measure for $L$ is also defined for $\alpha$ ($0 \leq \alpha \leq 1$) as

$$
\text{CTE}_\alpha := \mathbb{E}[L | L > \text{VaR}_\alpha]. \tag{2.9}
$$

It is the capital required to cover the average amount of liabilities when they exceed the quantile measure with the probability of at most $1 - \alpha$.

3. Distribution of the net liability $L$

Recall that the net liability defined in Equation (2.6) is a path-dependent functional of the underlying equity index/fund process. The net liability $L$ depends on two random processes, the value of the equity index/fund $F_t$ and the present value of the accumulated capital $A_t = \int_0^t e^{-rs} F_s \, ds$. The hitting time $\tau_0$ is the first time when the process $F_t$ hits zero, after which the insurer’s liability becomes deterministic. The stochastic dynamics are the following:

$$
dF_t = (\mu - m)F_t \, dt - w \, dW_t + \sigma F_t \, dB_t, \quad F_0 = G, \\
daA_t = e^{-rs} F_t \, dt.
$$

Let us define function $v(t, x, y)$ for a fixed number $K$ as

$$
v(t, x, y) = \mathbb{E}[\mathbb{1}[L \leq K]|F_t = x, A_t = y] = \mathbb{P}(L \leq K|F_t = x, A_t = y). \tag{3.1}
$$

We denote the partial derivatives of $v$ with respect to $t$, $x$, $y$ by $v_t, v_x, v_y$, respectively, and the second partial derivative of $v$ with respect to $x$ by $v_{xx}$. In particular, the loss distribution at time $t = 0$ corresponds to $v(0, G, 0) = \mathbb{P}(L \leq K)$. Observe that since the probability $\mathbb{P}(L \leq K)$ can be determined for
Theorem 3.1 (Computation of $\mathbb{P}(L < K)$) The function $v$ is a solution of the following partial differential equation for $t \in (0, T)$, $x \in (0, \infty)$, $y \in (0, \infty)$

$$v_t + ((\mu - m)x - w)v_x + e^{-rt}x v_y + \frac{1}{2} \sigma^2 x^2 v_{xx} = 0$$

with terminal condition

$$v(T, x, y) = 1,$$  \hspace{1cm} (3.3)

and boundary conditions

$$v(t, 0, y) = \frac{u}{(r - \mathbb{E}[\sigma^2])} (e^{-rt} - e^{-rT}) - K,$$  \hspace{1cm} (3.4)

$$\lim_{y \to \infty} v(t, x, y) = 0,$$  \hspace{1cm} (3.5)

$$\lim_{x \to \infty} v(t, x, y) = 1,$$  \hspace{1cm} (3.6)

where $f(t, K) := \frac{u}{(r - \mathbb{E}[\sigma^2])} (e^{-rt} - e^{-rT}) - K.$

Proof The process $v(t, F_t, A_t)$ is a martingale with stochastic dynamics

$$d v(t, F_t, A_t) = v_t dt + v_x dF_t + v_y dA_t + \frac{1}{2} v_{xx} (dF_t)^2$$

$$= v_t + ((\mu - m)x - w)v_x + e^{-rt}x v_y + \frac{1}{2} \sigma^2 x^2 v_{xx} dt.$$

The $dt$ terms must vanish, so the function $v$ must satisfy the partial differential equation (3.2). The partial differential equation applies in the case when none of the boundaries is hit, more specifically when $F_T > 0$ so the fund is not depleted until the end of the contract. When the contract reaches maturity, the insurer in this case makes a profit which corresponds to the accumulated present value of the fee income (the loss is negative):

$$L = -m_w A_T.$$

Should $v$ represent the probability $\mathbb{P}_x(L < K)$, we need $\mathbb{P}(L \leq K)$ as the terminal condition which translates to

$$\mathbb{P}(L \leq K) = \mathbb{P}(-m_w A_T \leq K) = \mathbb{P}(A_T \geq -\frac{K}{m_w}).$$

This gives the terminal condition (3.3) given that $A_T = y > 0$.

Since only the first derivative is involved in the $y$-dimension, we treat $y$ as a time-like variable, which requires only one trivial boundary condition (3.6). A more challenging boundary is when $x = 0$. This corresponds to the absorbing boundary $F_0 = 0$. Given the value of $A_t$, the process $\{F_s, t < s < T\}$ becomes deterministic, i.e.

$$\mathbb{P}(L \leq K) = \mathbb{P} \left( \int_0^T e^{-rs} w \, ds - m_w A_T \leq K \right)$$

$$= \mathbb{P} \left( \frac{u}{r - \mathbb{E}[\sigma^2]} (e^{-rt} - e^{-rT}) - \frac{K}{m_w} \leq A_T \right),$$

which leads to the boundary condition (3.4). The boundary for $x \to \infty$ can be taken anything consistent with other boundaries since $\mathbb{P}(F_T \to \infty) = 0$. In particular, we can impose the Neumann boundary condition (3.5). \hfill $\Box$

For the computation of CTE, we observe that

$$\text{CTE}_u = \frac{1}{1 - \alpha} \mathbb{E}[L \mathbb{I}(L > \text{VaR}_u)].$$

Therefore, we introduce for any fixed $K > 0$,

$$u(t, x, y) = \mathbb{E}[L \mathbb{I}(L > \text{VaR}_u)]|_{F_t = x, A_t = y}.$$

In particular, the CTE is determined by $\text{CTE}_u = u(0, G, 0) / (1 - \alpha)$ with $K = \text{VaR}_u$, which is determined by a solution to the previous PDE.

Theorem 3.2 (Computation of $\mathbb{E}[\|L\|_2(L > K)]$) The function $u$ is a solution of the following partial differential equation for $t \in (0, T)$, $x \in (0, \infty)$, $y \in (0, \infty)$

$$u_t + ((\mu - m)x - w)u_x + e^{-rt}yu_y + \frac{1}{2} \sigma^2 y^2 u_{yy} = 0$$

with terminal condition

$$u(T, x, y) = 0,$$  \hspace{1cm} (3.8)

and boundary conditions

$$u(t, 0, y) = \left[ \frac{w}{r} (e^{-rt} - e^{-rT}) - m_w y \right] \mathbb{P}(f(t, K) > y),$$  \hspace{1cm} (3.9)

$$\lim_{y \to \infty} u_x(t, x, y) = 0,$$  \hspace{1cm} (3.10)

$$\lim_{x \to \infty} u(t, x, y) = 0.$$  \hspace{1cm} (3.11)

Proof The derivation of (3.7) is identical to that of (3.2) due to the strong Markov property. Given that $A_T = y > 0$, we note that

$$u(t, x, y) = L\mathbb{P}(L > K) = L\mathbb{P} \left( y < -\frac{K}{m_w} \right) = 0,$$

which produces the terminal condition (3.8). When the process $F_t$ hits the absorbing boundary 0, we know that

$$L\mathbb{P}(L > K) = \mathbb{E} \left( \int_t^T e^{-rs} w \, ds - m_w A_t \right)$$

$$\times \mathbb{P} \left( \int_t^T e^{-rs} w \, ds - m_w A_t > K \right),$$

which leads to the boundary condition (3.9). The other two boundary conditions (3.10) and (3.11) follow trivially from the definition of $u$. \hfill $\Box$

We can also compute the probability density of the GMWB net liability defined by

$$\mathbb{P}(L_T \in dK|F_t = x, A_t = y) = p(t, x, y) \, dK.$$

Note that

$$p(t, x, y) \approx p^*(t, x, y) := \mathbb{E} \left[ L_T \mathbb{I}(K - \frac{1}{2n}, K + \frac{1}{2n}) \right]$$

$$= \mathbb{E} \left[ L_T \mathbb{I}(K - \frac{1}{2n}, K + \frac{1}{2n}) | F_t = x, A_t = y \right]$$

$$= \int \left[ \mathbb{P} \left( t, x, y; K - \frac{1}{2n} \right) - \mathbb{P} \left( t, x, y; K + \frac{1}{2n} \right) \right] \, dK.$$  \hspace{1cm} (3.12)

It follows immediately that $p^*$ satisfies the same PDE as the function $v$ (see equation (3.2)) for $t \in (0, T)$, $x \in (0, \infty)$, $y \in (-\infty, \infty)$

$$p^*_t + ((\mu - m)x - w)p^*_x + e^{-rt}x p^*_y + \frac{1}{2} \sigma^2 x^2 p^*_{xx} = 0$$

with terminal condition

$$p^*(T, x, y) = 0.$$  \hspace{1cm} (3.14)
and boundary conditions
\[ p^*(t, 0, y) = n \left( f \left( t, K + \frac{1}{2n} \right) \leq y \leq f \left( t, K - \frac{1}{2n} \right) \right). \]

\[
\lim_{x \to \infty} p^*_x(t, x, y) = 0, \quad \lim_{y \to \infty} p^*(t, x, y) = 0. \tag{3.16}\]

In other words, the density \( p \) satisfies the PDE (3.13) subject to the terminal condition (3.14), the boundary conditions (3.16), (3.17) and

\[ p(t, 0, y) = \delta_{\tilde{f}(t, K)}(y), \]

where \( \delta_x \) is the Dirac delta function that assigns probability mass one to the point \( x \). We use (3.12) as an approximation of the Dirac delta function in the boundary condition (3.15).

The desired quantity, the probability density of the GMWB net liability, is thus determined by \( p(0, G, 0) \).

4. Numerical examples

Monte Carlo simulation has a great advantage of easy implementation with merely the knowledge of the underlying stochastic process, which explains its popularity in practice. However, one has to bear in mind that positive net liabilities are very rare observations for two reasons.

(1) It is rare that the account is depleted prior to full refund of premium \((t_0 < T)\).

(2) In the rare cases where it does happen, the time of depletion has to arrive early enough so that the insurer’s benefit payments exceed the accumulation of fee income prior to depletion, i.e. on the set \((t_0 < T)\), the following has to become positive

\[
\int_{t_0}^{T} e^{-r s} w \, ds - m_w \int_{0}^{t_0} e^{-r s} F_s \, ds.
\]

It is a known fact that estimators of risk measures based on rare events generally have large sampling errors. In order to achieve a satisfactory level of accuracy, simulation methods for risk measures of the GMWB would often require intensive computation, as we shall demonstrate in the numerical example.

Here we use plain Monte Carlo simulations as the benchmark to test the accuracy and efficiency of the computing algorithm proposed in earlier sections. In this numerical example, we set the valuation assumptions as follows. All parameters are provided on per annum basis.

- Drift parameter \( \mu = 0.09 \) and volatility parameter \( \sigma = 0.3 \);
- M&E fee rate \( m = 0.01 \) and the GMWB fee \( m_w = 0.0035 \);
- Rate of return on insurer’s assets backing up the GMWB liability, \( r = 0.05 \);
- Fixed rate of withdrawals per annum, \( w = 0.07 G \).

The guarantee base is chosen to be \( G = F_0 = 1 \) so that the resulting risk measures are represented as percentages of the guarantee base amount. Note that in this example all parameters including the management fee \( m_w \) were arbitrarily chosen in ranges of values commonly observed in practice. As the focus of this paper is on the computation of risk measures, we did not attempt to find a ‘fair value’ of \( m_w \) under risk neutral measure, although such an exercise can be done as shown in Feng and Volkmer (2016 forthcoming). We have intentionally chosen the volatility parameter, which is larger than that typically observed in practice, in order to make enough observations for Monte Carlo estimation. All computations are performed on HP SL390G7 1U Servers with two Intel HP X5650 2.66Ghz 6C Processors, and 12–96GB RAM.

We run simulations of account values from the policy issue to the earlier of the time of depletion and maturity on equally space time points \( \{t_0 = 0, \ldots, t_K = k \Delta t, \ldots, t_n = T\} \). Then we compute each scenario of the net liability according to the discretized version of equation (2.6). In each experiment, we generate three million scenarios \((N = 3,000,000)\) of \( L \) and rank them to form an empirical distribution. As the 90%-CTE is often used for the RBC calculation in practice, we set \( \alpha = 0.90 \) in this example. The risk measure VaR\(\alpha\) is estimated by the order statistic \( L(\alpha N) \) and CTE\(\alpha\) is estimated by the arithmetic average of the largest \( \alpha N \) order statistics. In order to achieve a smaller sampling variation, we calculate all estimates of risk measures by the average of 100 experiments. In table 1, we report the estimates of survival probabilities with various values of \( \Delta t \). The standard deviations of the estimates are reported in brackets.

For comparison, we compute the corresponding probability distribution function according to the computing algorithm described in section 1. In this algorithm, we set \( \Delta t = \Delta x = \Delta y = 0.01, b = 6, c = 12 \). It is shown in table 1 that the PDE algorithm can reach an accuracy up to three to four decimals within roughly 5 minutes. Note that the Monte Carlo simulations suffer from a noticeable discretization error for coarser \( \Delta t \)’s. The precision improves by taking a much finer \( \Delta t \), however the simulation takes several orders of magnitude longer time in comparison to the presented PDE approach.

We can adapt the PDE algorithm from section 4 to determine and plot the probability density function of the GMWB net liability. This is shown in figure 1 for \( K > 0 \). For computational purposes, we set \( n = 100 \) for the approximation for the Dirac
This explains why in practice most policyholders withdraw at unimodal and reaches a maximum at around \( K \) wealth MC \( (\Delta t = 0.1) \). One should note that the net refund of policyholder’s initial deposit, can be viewed as some complex ‘put option’ on the initial deposit. The more withdrawals, the more benefits can be realized from the ‘put option’. This explains why in practice most policyholders withdraw at the maximal allowed amount without incurring a penalty. On the other hand, since the policyholder has the ownership of the policyholder’s account can actually exacerbate the insurer’s liability, as it increases the chance of severe losses. The innate structure of this risk management problem seems to encourage reducing exogenous costs, thereby pushing the insurer to use policyholders’ account values, and hence increase the chance of fund depletion prior to maturity, thereby increasing the net liabilities. This appears to be a conundrum from an insurer’s point of view.

The problems of optimal withdrawal strategies were discussed in full details in Dai et al. (2008) and Huang and Kwok (2014). Although we do not investigate the optimal strategy in this paper our focus is on the computation of risk capitals, we can see the impact of the withdrawal rate from the perspective of an insurer. Unlike the policyholder’s competing interests with withdrawals, it is clear from table 2 that the more policyholder withdraws, the higher the insurance liability. This is not surprising as the relation can be explained by the formulation of net liability in equation (2.6). The first term in equation (2.6) is an increasing function of \( w \). In the second term, the account value \( F \) decreases with \( w \) and so does the fee income, represented by the second integral in equation (2.6). Hence, in general, the net liability increases with \( w \).

It is also interesting to see in table 3 how the fee distribution can affect the insurer’s net liability. It is typical that the total M&E fee \( m \) is distributed between the third party fund manager with whom the policyholder’s purchase payments are invested and the variable annuity writer. Only a certain portion \( m_w \) goes to cover the cost of the GMWB. For each fixed total fee \( m \), the higher rider charge, the more income for the insurer to compensate for its liabilities and hence the lower net liabilities. For each fixed rider charge \( m_w \), higher M&E fees lead to lower account values, and hence increase the chance of fund depletion prior to maturity, thereby increasing the net liabilities. This appears to be a conundrum from an insurer’s point of view.

On the one hand, the insurer wants to charge more than the net cost of guaranteed benefits to cover administrative and other costs. On the other hand, if only channelling the same amount of fees to fund the GMWB, then charging more fees from the policyholders’ account can actually exacerbate the insurer’s liability, as it increases the chance of severe losses. The innate structure of this risk management problem seems to encourage reducing exogenous costs, thereby pushing the insurer to use policyholder’s investment in a more efficient manner.

<table>
<thead>
<tr>
<th>( K )</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>Time (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC ( (\Delta t = 0.1) )</td>
<td>0.640570</td>
<td>0.697059</td>
<td>0.756167</td>
<td>0.815808</td>
<td>0.872781</td>
<td>126</td>
</tr>
<tr>
<td>MC ( (\Delta t = 0.001) )</td>
<td>0.63529</td>
<td>0.691627</td>
<td>0.750476</td>
<td>0.810114</td>
<td>0.867442</td>
<td>1175</td>
</tr>
<tr>
<td>MC ( (\Delta t = 0.001) )</td>
<td>0.634963</td>
<td>0.691042</td>
<td>0.749874</td>
<td>0.809524</td>
<td>0.866930</td>
<td>11600</td>
</tr>
<tr>
<td>PDE</td>
<td>0.635076</td>
<td>0.691099</td>
<td>0.749851</td>
<td>0.809359</td>
<td>0.866573</td>
<td>5</td>
</tr>
<tr>
<td>( \int_0^T w e^{-rt} dt = \frac{w}{r} (1 - e^{-rT}) \approx 0.7146, ) which corresponds to the extreme case where the policyholder’s account is exhausted immediately after the policy issue, i.e. ( \tau_0 = 0 ). Figure 1 indicates that the density function of ( L ) is unimodal and reaches a maximum at around ( K = 0.06 ).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the second example, we investigate the impact of withdrawal rate on the insurer’s net liability, which is measured by both VaR\(_{0.9}\) and CTE\(_{0.9}\) in table 2. Both risk measures are computed using the PDE algorithms developed in section 1. From the perspective of a policyholder, there are two competing factors that may affect his or her rate of withdrawal. On the one hand, the GMWB itself, which is a guaranteed death function in equation (3.12). One should note that the net liability increases with \( w \) as \( L \) is bounded above by \( \int_0^T w e^{-rt} dt = \frac{w}{r} (1 - e^{-rT}) \approx 0.7146, \) which corresponds to the extreme case where the policyholder’s account is exhausted immediately after the policy issue, i.e. \( \tau_0 = 0 \). Figure 1 indicates that the density function of \( L \) is unimodal and reaches a maximum at around \( K = 0.06 \). In the second example, we investigate the impact of withdrawal rate on the insurer’s net liability, which is measured by both VaR\(_{0.9}\) and CTE\(_{0.9}\) in table 2. Both risk measures are computed using the PDE algorithms developed in section 1. From the perspective of a policyholder, there are two competing factors that may affect his or her rate of withdrawal. On the one hand, the GMWB itself, which is a guaranteed death benefit, can be viewed as some complex ‘put option’ on the initial deposit. The more withdrawals, the more benefits can be realized from the ‘put option’. This explains why in practice most policyholders withdraw at the maximal allowed amount without incurring a penalty. On the other hand, since the policyholder has the ownership of the policyholder’s account can actually exacerbate the insurer’s liability, as it increases the chance of severe losses. The innate structure of this risk management problem seems to encourage reducing exogenous costs, thereby pushing the insurer to use policyholder’s investment in a more efficient manner.

5. Conclusion and future works

This paper develops a quantitative framework of profit–loss analysis of an insurer’s net liabilities from GMWBs with intricate option-like features. A novel PDE approach is presented for the computation of risk measures of GMWB net liabilities, filling in a gap in the literature where Monte Carlo method is
the only available tool. Our numerical examples show that the PDE approach is more efficient than Monte Carlo simulations under the current model setting.

Future research is warranted on more complex model structures of particular interest to practitioners. For example, this paper simply assumes that all assets are invested in bonds with deterministic and flat term structure of yield rates. One may extend the model to include stochastic interest rates. Practitioners may also be concerned about dynamic policyholder behaviour under which lapse rates increase when the GMWB is out-of-the-money and decrease when the GMWB is in-the-money.

An anonymous reviewer also pointed out that the model framework works under the assumption of a stand-alone contract. Our approach provides a means of determining risk based capital at the contract level. This is indeed the first step of stochastic reserving exercise in practice where only representative contracts for each group of contracts with similar characteristics are considered. Further computations of reserves and capital requirements at the level of a block of business require aggregate and allocation methods. A discussion of individual and aggregate models for similar variable annuity contracts can be found in Feng and Huang (2016) and Feng and Shimizu (2016).

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
The work of Jan Vecer was supported in part by GACR [grant number 16-21216S], [grant number 13-34480S].

Appendix 1. Computational algorithm
In this section, we develop numerical algorithm to solve the PDE (3.2) with the terminal and boundary conditions (3.3)-(3.6). In this algorithm, we use the implicit method in the x-dimension and hence the algorithm is unconditionally stable.

For the numerical purposes, we truncate the state space of X to (0, b) and the state space of Y to (0, c) for large enough b, c > 0. Consider the following conditions

\[ v(x, y, t) = 1, \quad x \in (0, b), y \in (0, c), \quad (\text{terminal condition}), \]

\[ v(x, c, t) = 0, \quad x \in (0, b), t \in (0, T), \quad (\text{boundary condition 1}), \]

\[ v(x, y, t) = h(t, x, y), \quad t \in (0, T), y \in (0, c), \quad (\text{x-boundary condition 1}), \]

\[ v(x, y, t) = 0, \quad t \in (0, T), y \in (0, c), \quad (\text{boundary condition 2}). \]

For the ease of presentation, we introduce \( h(t, x, y) = v(T-t, x, c-y) \) defined on \( (0, T) \times (0, b) \times (0, c) \), which satisfies the forward PDE

\[ h_t + e^{-r(T-t)} h_y = \{[\mu - m]x - w\} h_x + \frac{1}{2} \sigma^2 x^2 h_{xx}, \quad (A1) \]

subject to the conditions

\[ h(0, x, y) = 1, \quad x \in (0, b), y \in (0, c); \quad (A2) \]

\[ h(t, x, 0) = 1, \quad x \in (0, b), t \in (0, T); \quad (A3) \]

\[ h(t, 0, y) = 0, \quad t \in (0, T), y \in (0, c); \quad (A4) \]

\[ h_x(t, b, y) = 0, \quad t \in (0, T), y \in (0, c). \quad (A5) \]

Observe that the Neumann condition (x-boundary condition 2) has a natural probabilistic interpretation in the truncated space. The boundary \( x = b \) is considered instantaneously reflecting, i.e. the underlying Brownian motion bounces back immediately after it hits the boundary. Although this is not the intended problem as described in (2.1), we can choose \( b \) sufficiently large so that the probability of the Brownian motion ever reaching the reflecting boundary is negligible.

References
We set up the grids \((t_i, y_j, x_k)\) as follows and denote the solution
\(h(t_i, x_k, y_j)\) on the grid points by \(h^{i,j}_k\), where
\[
t_i = (i - 1)\Delta t, \quad \text{for } i = 1, \ldots, N_t + 1, \quad N_t := T / \Delta t;
\]
\[
y_j = (j - 1)\Delta y, \quad \text{for } j = 1, \ldots, N_y + 1, \quad N_y := c / \Delta y;
\]
\[
x_k = x_0, \quad \text{for } k = 0, \ldots, N_x, \quad N_x := b / \Delta x.
\]
Discretization of (A1) gives
\[
\frac{h^{i+1,j+1}_k - h^{i,j+1}_k}{\Delta t} + g^t_k h^{i+1,j+1}_k h^{i,j+1}_k + \frac{\Delta y}{\Delta t} = ((\mu - m)k\Delta x - w) h^{i+1,j+1}_k h^{i,j+1}_k.
\]
\[
\frac{\Delta x}{\Delta x}^2 h^{i+1,j+1}_k + h^{i+1,j+1}_k - 2h^{i+1,j+1}_k, \quad \Delta x^2.
\]
where
\[
g^t := e^{-(T - (i - 1)\Delta t)}.
\]
For brevity, we use the notation
\[
\alpha = \frac{1}{\Delta t}, \quad \beta = \frac{\Delta x}{\Delta y}, \quad \gamma = \frac{w}{\Delta x}.
\]
For \(k = 1, \ldots, N_x\), the recursive relation (A6) can be rewritten as
\[
-\frac{1}{2} \left[ \sigma^2 k^2 - (\mu - m)k + \gamma \right] h^{i+1,j+1}_k
+ \frac{\Delta x}{\Delta x} h^{i+1,j+1}_k + \frac{\Delta y}{\Delta t} h^{i+1,j+1}_k
- \frac{1}{2} \left[ \sigma^2 k^2 + (\mu - m)k - \gamma \right] h^{i+1,j+1}_k = \alpha h^{i,j}_k + \beta g^t h^{i+1,j}_k.
\]
The Dirichlet condition (A4) gives
\[
h^{i,j}_0 = \left\{ \begin{array}{ll}
(j - 1)\Delta y & c \leq \frac{K}{m_w} \frac{w}{r m_w} e^{-rT} (e^{r(1-1)\Delta t} - 1)
\end{array} \right.
\]
The Neumann condition (A5) yields
\[
h^{i,j}_{N_t+1} = h^{i,j}_{N_t-1}.
\]
In view of (A7) and (A8), we obtain
\[
-\sigma^2 N_x h^{i+1,j+1} + \sigma^2 N_x g^t + \alpha h^{i+1,j+1} = \alpha h^{i,j} + \beta g^t N_x h^{i+1,j+1}.
\]
Then
\[
B^t h^{i+1,j+1} = \alpha h^{i,j+1} + \beta h^{i+1,j} + \left( \frac{1}{2} \sigma^2 - \mu + m + \gamma \right) h^{i+1,j+1}, 0, \ldots, 0 \right)^T.
\]
where \(C\) is a diagonal matrix with diagonal elements \((\beta g^t, 2\beta g^t, \ldots, N_x, \beta g^t)\) and \(B^t\) is an \(N_x\) by \(N_x\) matrix
\[
B^t = \begin{pmatrix}
\sigma^2 + \beta g^t + \alpha & -\frac{1}{2} (\sigma^2 + \mu - m) & \ldots & 0 & 0 & 0 \\
-\frac{1}{2} (\sigma^2 - 2 - \mu - m) & \sigma^2 + 2\beta g^t + \alpha & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & \sigma^2 N_x^2 & \sigma^2 N_x^2 + \beta N_x g^t + \alpha & \ldots
\end{pmatrix}.
\]
On the \(k\)th row \((k = 2, \ldots, N_x - 1)\), the tridiagonal elements are
\[
-\frac{1}{2} \left[ \sigma^2 k^2 - (\mu - m)k + \gamma \right], \quad \sigma^2 k^2 + \beta g^t + \alpha,
-\frac{1}{2} \left[ \sigma^2 k^2 + (\mu - m)k - \gamma \right].
\]
We can determine \(h\) by
\[
h^{i+1,j+1} = (B^t)^{-1} \left[ \alpha h^{i,j+1} + Ch^{i+1,j} + \left( \frac{1}{2} (\sigma^2 - \mu + m + \gamma) h^{i+1,j+1}, 0, \ldots, 0 \right)^T \right].
\]
Marching alternatively in the \(i\) and \(j\) directions, starting from \(i = 1\) and \(j = 1\), where we have the initial conditions (A2) and (A3) written in vector form
\[
h^{1,j} = 1, \quad h^{i,1} = 1.
\]
The above-described algorithm can be easily adapted to compute \(L[L[L \cdots (\Delta t)] \cdots] \), which is determined by the PDE (3.7) with boundary conditions (3.8)-(3.11). In this case, we have the same forward PDE (A1) for \(h(t, x, y) = u(T - t, x, c - y)\) subject to the following conditions
\[
h(0, x, y) = 0, \quad x \in (0, b), \quad y \in (0, c); \quad (A9)
\]
\[
h(t, 0) = 0, \quad x \in (0, b), \quad t \in (0, T); \quad (A10)
\]
\[
h(t, 0) = \left\{ \begin{array}{ll}
\frac{w}{r} (e^{-rT} - e^{-rT}) - mw(c - y) \times \left\{ f(T - t, K) \right. & c - y), \\
\end{array} \right.
\]
\[
t \in (0, T), \quad y \in (0, c); \quad (A11)
\]
\[
h(t, b, y) = 0, \quad t \in (0, T), \quad y \in (0, c). \quad (A12)
\]
With the same discretization scheme, we obtain the boundary conditions for (A9) and (A10)
\[
h^{1,j} = 0, \quad h^{i,1} = 0.
\]
The Dirichlet condition (A11) translates to
\[
h^{1,j}_0 = \left( \frac{w}{r} e^{-rT} (e^{r(1-1)\Delta t} - 1) - mw(c - (j - 1)\Delta y) \right) \times \left\{ (j - 1)\Delta y > c \right. \quad \frac{K}{m_w} \quad \frac{w}{r m_w} e^{-rT} (e^{r(1-1)\Delta t} - 1) \right.
\]
The rest of the algorithm is precisely the same as in the previous case.