Actuarial Research Centre
Institute and Faculty of Actuaries

Actuarial Research Centre (ARC)

PhD studentship output

The Actuarial Research Centre (ARC) is the Institute and Faculty of Actuaries' network of actuarial researchers around the world.
The ARC seeks to deliver research programmes that bridge academic rigour with practitioner needs by working collaboratively with academics, industry and other actuarial bodies.
The ARC supports actuarial researchers around the world in the delivery of cutting-edge research programmes that aim to address some of the significant challenges in actuarial science.

CALCULATING VARIABLE ANNUITY LIABILITY "GREEKS" USING MONTE CARLO SIMULATION

BY

MARK J. CATHCART, HSIAO YEN LOK, ALEXANDER J. MCNEIL AND STEVEN MORRISON

ABSTRACT

The implementation of hedging strategies for variable annuity products requires the calculation of market risk sensitivities (or "Greeks"). The complex, path-dependent nature of these products means that these sensitivities are typically estimated by Monte Carlo methods. Standard market practice is to use a "bump and revalue" method in which sensitivities are approximated by finite differences. As well as requiring multiple valuations of the product, this approach is often unreliable for higher-order Greeks, such as gamma, and alternative pathwise (PW) and likelihood-ratio estimators should be preferred. This paper considers a stylized guaranteed minimum withdrawal benefit product in which the reference equity index follows a Heston stochastic volatility model in a stochastic interest rate environment. The complete set of first-order sensitivities with respect to index value, volatility and interest rate and the most important second-order sensitivities are calculated using PW, likelihood-ratio and mixed methods. It is observed that the PW method delivers the best estimates of first-order sensitivities while mixed estimation methods deliver considerably more accurate estimates of second-order sensitivities; moreover there are significant computational gains involved in using PW and mixed estimators rather than simple BnR estimators when many Greeks have to be calculated.

KEYWORDS

Stochastic simulation, Monte Carlo estimation, variable annuity, Greeks, sensitivities, Heston stochastic volatility model, pathwise method, likelihood-ratio method, stochastic interest rates.

1. INTRODUCTION

Many issuers of variable annuity (VA) contracts use a hedging strategy to mitigate some of the risk involved in writing these products. To design a hedging strategy it is essential to be able to calculate reliable estimates of the sensitivities

Astin Bulletin 45(2), 239-266. doi: 10.1017/asb.2014.31

of VA liabilities to their key risk drivers. In this paper, we will focus on the guaranteed minimum withdrawal benefit (GMWB) product and compare different Monte Carlo methods for estimating first-order and second-order sensitivities (or "Greeks"). We do this in the context of an extended version of the Heston stochastic volatility model (Heston, 1993) incorporating both stochastic volatility and stochastic interest rates.

There is a growing literature on VA products and their valuation and hedging; Ledlie *et al.* (2008) is a good introductory article to the product class. Milevsky and Salisbury (2006) were the first to study the GMWB product and they showed that it can be broken down into a term-certain annuity component plus an Asian quanto option. Due to the long-term investment in a VA product, stochastic volatility and interest rates can have a significant effect on valuation. Peng *et al.* (2012) studied the valuation of the GMWB product under stochastic interest rates and derived lower and upper bounds for the value of the option component. Jaimungal *et al.* (2014) developed valuation methods that incorporate both stochastic volatility and interest rates, based on a PDE approach, whereas Bacinello *et al.* (2011) have investigated the effects of stochastic volatility, interest rates and force of mortality using a Monte Carlo and least squares Monte Carlo approach for various VA products. Kling *et al.* (2011) show that stochastic volatility is important especially for hedging purposes. A number of papers also show that assumptions about policyholder lapsation behaviour can have a significant effect on the valuation and hedging of VA products (Milevsky and Salisbury, 2006; Chen *et al.*, 2008; Dai *et al.*, 2008).

The main contributions of our paper are as follows. Building on the work of Broadie and Kaya (2004), the likelihood-ratio (LR) method, a standard approach for evaluating option price sensitivities by Monte Carlo simulation, is extended to work under a more realistic market model with stochastic volatility and stochastic interest rates. The PW method is developed in the context of VA liabilities, following a similar approach to Hobbs *et al.* (2009), but extending this to obtain PW estimates of sensitivities to interest rates and volatility (rho and vega). Finally, the PW and the LR approaches are combined to construct a new and efficient mixed estimator for second-order sensitivities, such as gamma, vanna and the sensitivity of delta with respect to the interest rate. This is relevant for practitioners hedging their VA exposures, since accurate and unbiased estimates of higher-order sensitivities allow the adaptation of hedging strategies to take account of the convexity of liabilities with respect to key market exposures.

Indeed conversations with practitioners suggest that the use of second-order sensitivities in hedging VA products is becoming more widespread, particularly with the increasing availability of high-performance computing solutions. A strategy that has been described to us involves the active management of a portfolio of hedging instruments to match, as far as possible, first and second-order Greeks on both the asset and liability side. The hedging instruments are futures, forwards and interest rate swaps as well as, occasionally, more exotic instruments like repos. Particular attention is devoted to matching delta and rho (sensitivity to the equity index and interest rate) and dynamic rebalancing on an intra-day basis is used for these Greeks when necessary. For second-order Greeks the rebalancing is less frequent but also considered important due to the long-term nature of the guarantees and the sensitivity to changes in interest rates and volatility that has been reported in the literature.

The paper is organized as follows. In Section 2, an overview of the standard methods for estimating option price sensitivities by Monte Carlo simulation is given. In Section 3, the asset price model incorporating stochastic volatility and interest rates is presented and a nested approach to implementing the LR method in the context of the model is described. A stylized VA product of GMWB type is presented in Section 4 and the details of the various estimation methods for the Greeks are developed for this product. Section 5 compares the accuracy and efficiency of all the estimators developed for the stylized VA product in Section 4 under a number of different scenarios for the model parameters and Section 6 concludes.

2. Option price sensitivity estimators

We briefly review three standard approaches to estimating option price sensitivities by Monte Carlo: the "bump and revalue" (BnR) method, a natural finite-difference approach that is often used by practitioners; the PW method; the LR method. These methods were developed in the context of option pricing by Broadie and Glasserman (1996); see also the textbook by Glasserman (2003).

Let $Y(\theta)$ denote the discounted payoff of a path-dependent option expressed as a function of a sensitivity parameter or risk factor θ. In the context of an option on an equity index (which is effectively the situation with a VA product) this sensitivity parameter could, for example, be the initial equity index level or the initial interest rate. The price or value of the option is given by $\alpha(\theta) = \mathbb{E}(Y(\theta))$, where the expectation is taken with respect to a risk–neutral or pricing measure. We are interested in estimating $\alpha'(\theta)$.

2.1. Bump and revalue approach

Under this approach we simulate the discounted payoff of the option under some base scenario for the risk factor and then again under a so-called "bumped scenario", where the sensitivity parameter θ is increased by some small perturbation $\Delta\theta$. We form a forward difference estimate in the sensitivity parameter by first simulating n independent payoffs under the base scenario $(Y_1(\theta), \ldots, Y_n(\theta))$ and n independent payoffs under the bumped scenario $(Y_1(\theta + \Delta\theta), \ldots, Y_n(\theta + \Delta\theta))$ and then calculating

$$\hat{\Delta}^{\mathrm{B}} = \frac{\sum_{i=1}^{n} Y_i(\theta + \Delta\theta) - \sum_{i=1}^{n} Y_i(\theta)}{n\Delta\theta}.$$

It follows, by the strong law of large numbers, that, as $n \to \infty$, we have

$$\hat{\Delta}^{\mathrm{B}} \longrightarrow \frac{E(Y(\theta + \Delta\theta)) - E(Y(\theta))}{\Delta\theta} = \frac{\alpha(\theta + \Delta\theta) - \alpha(\theta)}{\Delta\theta} \approx \alpha'(\theta).$$

A similar approach involving second-order finite-difference approximations can obviously be applied to estimate second-order sensitivities.

The variance of the estimator $\hat{\Delta}^{\mathrm{B}}$ can be significantly reduced if the same random number stream is used to estimate the bumped and base option prices, as opposed to using independent random number streams for each. Nevertheless, the BnR approach can incur fairly large sampling errors and it may also be biased if the difference approximation to the derivative is poor. These problems are particularly acute for second-order sensitivity estimates. Note that the bias can be reduced by using a central difference rather than the forward difference described above, although this introduces a further computational cost; see Glasserman (2003) (from p. 378) for discussion.

2.2. Pathwise estimator

Under this approach we assume that the option payoff $Y(\theta)$ may be analytically differentiated with respect to θ to obtain $Y'(\theta)$. If an interchange of differentiation and the taking of expectations is justified, then we have that

$$\mathbb{E}(Y'(\theta)) = \frac{d}{d\theta}\mathbb{E}(Y(\theta)) = \alpha'(\theta), \tag{2.1}$$

and it follows that $n^{-1}\sum_{i=1}^{n} Y_i'(\theta)$ is an unbiased and strongly consistent estimator of $\alpha'(\theta)$ where $(Y_1'(\theta), \ldots, Y_n'(\theta))$ denote the values of the derivatives of the payoff with respect to θ for each Monte Carlo simulation of the path of the underlying asset price.

Sufficient conditions for the interchanging of expectations and derivatives are derived in Broadie and Glasserman (1996) (see also Glasserman, 2003, p. 393–396). The most important condition is that the payoff function should be Lipschitz continuous with respect to θ. This is satisfied by many, but not all, payoff functions that are common in option pricing and Glasserman suggests that the "rule of thumb" that the payoff should be continuous in the parameter of interest gives good guidance in most problems (see Glasserman, 2003, p. 396). The PW method is certainly not applicable in the presence of discontinuous payoffs.

2.3. Likelihood ratio estimator

In this method, we assume that the discounted payoff may be expressed as a function $Y = f(S_1, \ldots, S_m)$ of a vector of asset prices at a series of ordered times which, for the purposes of this paper, may be taken to be the regular times $t = 1, \ldots, m$ representing the policy anniversaries.

To begin with, we also assume that the function f does not depend explicitly on θ and that the dependence of Y on θ is through a parameter of the joint density g_θ of $(S_1, \dots, S_m)$ so that

$$\alpha(\theta) = \mathbb{E}(Y(\theta)) = \int_{\mathbb{R}^m} f(x_1, \dots, x_m) g_\theta(x_1, \dots, x_m) dx_1 \cdots dx_m. \tag{2.2}$$

Since the densities that arise in applications are typically continuous functions of θ, differentiation under the integral presents fewer problems than in the PW case

$$\begin{aligned}\alpha'(\theta) &= \int_{\mathbb{R}^m} f(x_1, \dots, x_m) \frac{d}{d\theta} g_\theta(x_1, \dots, x_m) dx_1 \cdots dx_m \\ &= \int_{\mathbb{R}^m} f(x_1, \dots, x_m) \frac{\frac{d}{d\theta} g_\theta(x_1, \dots, x_m)}{g_\theta(x_1, \dots, x_m)} g_\theta(x_1, \dots, x_m) dx_1 \cdots dx_m \\ &= \mathbb{E}\left(f(S_1, \dots, S_m) \frac{d}{d\theta} \ln g_\theta(S_1, \dots, S_m) \right). \end{aligned} \tag{2.3}$$

Thus an unbiased estimator of $\alpha'(\theta)$ may also be obtained by taking

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \frac{d}{d\theta} \ln g_\theta(S_{i,1}, \dots, S_{i,m}),$$

where $(S_{i,1}, \dots, S_{i,m})$, $i = 1, \dots, n$, represent Monte Carlo simulations of the asset value vector and $Y_i = f(S_{i,1}, \dots, S_{i,m})$, $i = 1, \dots, n$, are the corresponding discounted payoffs. The term $\frac{d}{d\theta} \ln g_\theta(x_1, \dots, x_m)$ is known in statistics as the "score function" of the density g_θ. In the Monte Carlo context it is referred to as the LR weight, since it multiplies the discounted payoff function to give the sensitivity estimator.

If the underlying asset value process is first-order Markovian, then, conditional on a fixed starting value $S_0 = x_0$, the joint density is given by a product of the form

$$g_\theta(x_1, \dots, x_m) = g_1(x_1|x_0) g_2(x_2|x_1) \cdots g_m(x_m|x_{m-1}),$$

and the LR weight is given by

$$\frac{d}{d\theta} \ln g_\theta(S_1, \dots, S_m) = \sum_{t=1}^{m} \frac{d}{d\theta} \ln g_t(S_t|S_{t-1}),$$

where we suppress the θ dependence of the conditional densities for notational simplicity.

For later purposes we note that, if the function f in (2.2) has an explicit dependence on θ then the expression for $\alpha'(\theta)$ in (2.3) has to be amended accordingly and we obtain

$$\alpha'(\theta) = \mathbb{E}\left(\frac{\partial}{\partial\theta} f(\theta, S_1, \ldots, S_m)\right) + \mathbb{E}\left(f(\theta, S_1, \ldots, S_m)\frac{d}{d\theta}\ln g_\theta(S_1, \ldots, S_m)\right). \tag{2.4}$$

This will be relevant to the estimation of certain sensitivities in our application; see also Example 7.3.7 in Glasserman (2003).

3. The asset price model and the nested approach

We will apply the estimators introduced in the previous section in the context of an asset price model with stochastic volatility and stochastic interest rates. This presents problems, in particular, for the LR method, since we need to obtain the conditional distributions $g_t(x_t|x_{t-1})$ of the asset prices to calculate the LR weights. Broadie and Kaya (2004) provided a solution to this problem in the context of Heston's stochastic volatility model (Heston, 1993) and also considered the case of a stochastic volatility jump diffusion. By conditioning on the volatility path, they developed a conditional LR method, which relies on the fact that the LR weights are available in closed form for the basic Black–Scholes model (see Glasserman, 2003, pages 403–405). We extend this approach to also incorporate stochastic interest rates via the Cox–Ingersoll–Ross (CIR) model.

3.1. The model

The system of stochastic differential equations that governs our asset price dynamics is

$$\begin{aligned} dV_t &= \kappa_V(\theta_V - V_t)dt + \sigma_V\sqrt{V_t}dW_t^V, \\ dr_t &= \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_t^r, \\ dS_t &= r_t S_t dt + \sqrt{V_t}S_t dW_t^S, \end{aligned} \tag{3.1}$$

where W_t^S, W_t^V and W_t^r are Wiener processes under the risk–neutral measure. Thus, the variance and interest rate both follow the same kind of stochastic process. We note that, it would be possible to further extend the methodology we describe below to consider an interest rate process with a larger number of risk-factors; this could be useful for incorporating a more advanced interest rate term structure.

We assume that the multivariate Wiener process $(W_t^V, W_t^r, W_t^S)^\top$ has instantaneous correlation matrix ρ, where ρ is a positive-definite matrix admitting the Cholesky decomposition $\rho = AA^\top$ for a lower-triangular matrix $A = (a_{ij})$.

We can express the asset price dynamics in terms of the elements of the matrix A by

$$\frac{dS_t}{S_t} = r_t dt + \sqrt{V_t}\Big(a_{31} dW_{t,1} + a_{32} dW_{t,2} + a_{33} dW_{t,3}\Big),$$

where $W_{t,1}$, $W_{t,2}$ and $W_{t,3}$ are three independent Brownian motions and an application of Ito's lemma yields

$$d \ln S_t = r_t dt + \sum_{k=1}^{3} \left(\sqrt{V_t} a_{3k} dW_{t,k} - \frac{1}{2} V_t a_{3k}^2 dt \right).$$

Following the approach of Broadie and Kaya (2004), we observe that the log asset price S_t may be written as

$$\ln\left(\frac{S_t}{S_u}\right) = \int_u^t r_s ds + a_{33} \int_u^t \sqrt{V_s} dW_{s,3} - \frac{1}{2} a_{33}^2 \int_u^t V_s ds + Y_{u,t},$$

where

$$Y_{u,t} = \sum_{k=1}^{2} \left(a_{3k} \int_u^t \sqrt{V_s} dW_{s,k} - \frac{1}{2} a_{3k}^2 \int_u^t V_s ds \right).$$

Hence we may write

$$S_t = S_u \xi_{u,t} \exp\left(\left(\bar{r}_{u,t} - \frac{1}{2} \bar{\sigma}_{u,t}^2 \right) \Delta_{u,t} + a_{33} \int_u^t \sqrt{V_s} dW_{s,3} \right), \qquad (3.2)$$

where $\Delta_{u,t} = t - u$ and where

$$\xi_{u,t} = \exp\left(Y_{u,t}\right),$$

$$\bar{\sigma}_{u,t} = \sqrt{\frac{a_{33}^2}{\Delta_{u,t}} \int_u^t V_s ds},$$

$$\bar{r}_{u,t} = \frac{1}{\Delta_{u,t}} \int_u^t r_s ds.$$

3.2. Conditional estimators

Recall that the payoff is a function $Y = f(S_1, \ldots, S_m)$ of the asset value at the times $t = 1, \ldots, m$. Suppose, we condition on volatility and interest rates and write H_m for the sigma algebra generated by the volatility and interest rate paths up to time m.

For any interval $[u, t] \subset [0, m]$, the quantities $\bar{r}_{u,t}$, $\bar{\sigma}_{u,t}$ and $\xi_{u,t}$ in (3.2) are known functions of the information in H_m. In practice we will approximate them in terms of the volatility and interest rates sampled at a series of discrete times in $[u, t]$. We will differ from Broadie and Kaya (2004) in that we will not generate

$\int_u^t V_s ds$ randomly given the endpoints V_u and V_t but will instead approximate the integral by quadrature based on a sufficiently fine discretization; this approach generalizes more easily to the complete set of integrals we require in our more general model.

Conditional on H_m, we have $a_{33} \int_u^t \sqrt{V_s} dW_{s,3} \sim N(0, \bar{\sigma}_{u,t}^2 \Delta_{u,t})$ so that S_t has a conditional lognormal distribution given by

$$\ln S_t \sim N\left(\ln(S_u \xi_{u,t}) + \Delta_{u,t}\left(\bar{r}_{u,t} - \frac{1}{2}\bar{\sigma}_{u,t}^2\right), \bar{\sigma}_{u,t}^2 \Delta_{u,t}\right).$$

Thus S_t can be sampled given S_u and H_m using the equation

$$S_t = S_u \xi_{u,t} \exp\left(\left(\bar{r}_{u,t} - \frac{1}{2}\bar{\sigma}_{u,t}^2\right)\Delta_{u,t} + \bar{\sigma}_{u,t}\sqrt{\Delta_{u,t}} Z_{u,t}\right), \tag{3.3}$$

where $Z_{u,t}$ is an independent standard normal shock or innovation variable.

In the LR method, the expectation we are trying to estimate can be written as

$$\mathbb{E}\left(\mathbb{E}\left(f(S_1, \ldots, S_m) \sum_{t=1}^{m} \frac{d}{d\theta} \ln g_t(S_t | S_{t-1}, H_m)\right)\right), \tag{3.4}$$

where $g_t(S_t | S_{t-1}, H_m)$ denotes the lognormal density of S_t given S_{t-1} and H_m. The inner expectation is taken with respect to the multivariate distribution of $(S_1, \ldots, S_m)$ given H_m and can be estimated by a conditional application of the LR method where the LR weights are straightforward to compute as explained below. The outer expectation is estimated by averaging over volatility and interest rate paths. The simulation setup is illustrated in Figure 1.

The conditional LR weights are easily computed by observing that

$$g_t(S_t | S_{t-1}, H_m) = \frac{1}{S_t \bar{\sigma}_{t-1,t}\sqrt{\Delta_{t-1,t}}} \phi\left(\frac{\ln\left(\frac{S_t}{S_{t-1}\xi_{t-1,t}}\right) - \left(\bar{r}_{t-1,t} - \frac{1}{2}\bar{\sigma}_{t-1,t}^2\right)\Delta_{t-1,t}}{\bar{\sigma}_{t-1,t}\sqrt{\Delta_{t-1,t}}}\right), \tag{3.5}$$

where ϕ represents the standard normal density function. Derivatives of $\ln g_t(S_t | S_{t-1}, H_m)$ with respect to key sensitivity parameters are mostly straightforward to compute.

The nested setup may also be used to implement both the PW and BnR methods. It is not necessary to do this but it facilitates a comparison between methods. In the case of the former, we express the expectation (2.1) as $\mathbb{E}(\mathbb{E}(Y'(\theta) | H_m))$ and estimate the inner expectation by conditioning on volatility and interest rate paths and averaging over asset value paths. A similar approach may be taken to estimating both the bumped value $\mathbb{E}(\mathbb{E}(Y(\theta + \Delta\theta) | H_m))$ and the base value $\mathbb{E}(\mathbb{E}(Y(\theta) | H_m))$ in the BnR method.

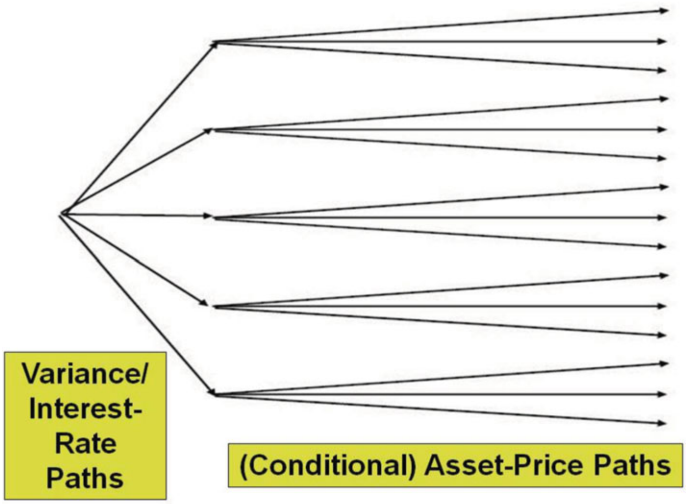

FIGURE 1: The Greeks are estimated using a nested application of Monte Carlo. For each outer volatility and interest rate scenario, multiple inner asset price paths are generated. Averaging over the inner scenarios gives a conditional estimate of the sensitivity for a fixed volatility and interest rate scenario; averaging over outer scenarios gives the unconditional estimate of the sensitivity. (Color online)

3.3. Generating model paths

There are various different numerical schemes for generating the volatility process and the interest rate process. In this paper we will consider the case when $\rho_{V,r} = 0$. This means that the two processes can be sampled independently and allows us to use exact simulation algorithms for this purpose. In practice it is often assumed that there is no correlation between volatility and interest rate changes, since the correlations $\rho_{S,V}$ (between asset-price changes and volatility changes) and $\rho_{S,r}$ (between asset-price changes and interest rate changes) are viewed as more important. To sample dependent paths for volatility and interest rates we could use simple bivariate Euler or Milstein schemes but these generally introduce more biases into the calculations; an exact bivariate scheme for generating volatility and interest rates jointly is not available.

We consider the following two univariate sampling schemes.

3.3.1. *GQE: the gamma quadratic exponential scheme of Chan and Joshi (2013).* We will describe the process for the equity volatility path. The process for the interest rate path follows similarly. Given V_u, V_t is sampled so that it satisfies

$$(V_t | V_u = v_u) \stackrel{d}{=} \Gamma(\alpha_V, \beta_V) + \sum_{i=1}^{N_{\lambda_t}} \mathrm{Exp}_i(\beta_V), \tag{3.6}$$

where

$$\alpha_V = \frac{2\kappa_V\theta_V}{\sigma_V^2}, \quad \beta_V = \frac{\sigma_V^2}{2\kappa_V}\left(1 - e^{-\kappa_V(t-u)}\right), \quad \lambda_t = \frac{2\kappa_V}{\sigma_V^2(e^{\kappa_V(t-u)} - 1)} v_u,$$

and where $\Gamma(\alpha_V, \beta_V)$ is a gamma random variable with mean $\alpha_V\beta_V$ and variance $\alpha_V\beta_V^2$, N_{λ_t} is a Poisson random variable with mean λ_t, and $\text{Exp}_i(\beta_V)$, $i = 1, \ldots, N_{\lambda_t}$, are independent and identically distributed exponential random variables with mean β_V. The second term in (3.6) is approximated using a scheme similar to the QE scheme of Andersen (2008). More details of the GQE scheme can be found in Chan and Joshi (2013).

3.3.2. *GL: a scheme proposed in Glasserman (2003), see p. 123.* This is a simple method in which we sample V_t given V_u using the equation

$$(V_t | V_u = v_u) \stackrel{d}{=} k_1\left(\left(Z_t^{(GL)} + \sqrt{k_2 v_u}\right)^2 + \chi^2_{2\alpha_V - 1}\right), \tag{3.7}$$

where $Z_t^{(GL)}$ is a standard normal variable, $k_1 = \sigma_V^2 \frac{1-e^{(-\kappa_V(t-u))}}{4\kappa_V}$, $k_2 = \frac{e^{-\kappa_V(t-u)}}{k_1}$, and α_V is as previously defined. The method is only possible under the restriction that $2\alpha_V > 1$.

3.4. Quadrature approximation of integrated processes

To calculate terms like $\bar{r}_{u,t}$ and $\bar{\sigma}_{u,t}$ in (3.2) we require accurate estimates of the integrated processes $\int_u^t V_s ds$ and $\int_u^t r_s ds$ for the volatility and interest rate paths. We also need to be able to approximate integrals like $\int_u^t \sqrt{V_s} dW_{s,k}$ for $k = 1, 2$ to calculate $\xi_{u,t}$. For example, for $k = 1$ we need to compute an estimate for

$$\int_u^t \sqrt{V_s} dW_{s,1} = \frac{1}{\sigma_V}\left(V_t - V_u - \kappa_V\theta_V\Delta_{u,t} + \kappa_V \int_u^t V_s ds\right), \tag{3.8}$$

which again involves $\int_u^t V_s ds$. In this paper, we will use simple quadrature approximations for all such integrals. For these approximation to be accurate, we require fine time steps in the discretization of the volatility and interest rate process. Suppose, we have p time steps between time u and time t, then the approximations we use are of the form

$$\int_u^t V_s ds \approx \sum_{i=1}^{p} (0.5 V_{u+(i-1)\Delta_{u,t}^{(p)}} + 0.5 V_{u+i\Delta_{u,t}^{(p)}}) \Delta_{u,t}^{(p)}, \tag{3.9}$$

where $\Delta_{u,t}^{(p)} = \frac{t-u}{p}$. The method for estimating $\int_u^t r_s ds$ is analogous.

4. Variable annuity liability sensitivities

4.1. Example variable annuity product

The methods for estimating option price sensitivities will now be applied to the problem of estimating the sensitivities of a stylized VA liability. The idea behind this stylized example product is that it should be simple enough both in terms of tractability and ease of exposition, yet retain some of the key features which make these products so popular in many markets. It should also have liabilities which are path-dependent, as is typical of many VA products on the market. The example product, based on a product described in Ledlie *et al.* (2008), will be of the GMWB type, where the policyholder is entitled to annual withdrawals from the underlying fund throughout the product lifetime, even if poor returns mean that the fund diminishes to zero (at which point the insurance company must provide this income out of its own reserves). The policyholder is entitled to the remaining fund value, if any, at product maturity.

More details of this VA product and the policyholder will now be given. Firstly the policyholder is a male who has just turned 65 and the GMWB option (or so-called guarantee rider) is initially active. This means that the policyholder pays an extra η^{guar} of the guarantee base (which is introduced below) in charges on top of the annual fund management charge; both are deducted from the VA fund value. If the policyholder wishes to turn off this option, he is entitled to do so and this cancels the extra guarantee charge, an event referred to as a customer “lapse”. The modelling of the rate of policyholder lapsation is by no means a trivial issue and we will make the simple assumption that there is a constant annual lapse rate ℓ. Incorporating dynamic policyholder lapsation will typically increase the value of the guarantee, even if a relatively high penalty charge is imposed (see Dai *et al.*, 2008). However, embedding dynamic policyholder behaviour within a Monte Carlo framework can be computationally inefficient and Bauer *et al.* (2008) have proposed a multidimensional discretization approach to address this issue.

With the GMWB option active the policyholder is guaranteed to receive income at a fixed percentage level w of the guarantee base each year after his 65th birthday, until the contract expires after T years. The contract expires early if the policyholder lapses or dies. The guarantee base is initially set at the amount of the policyholder premium, but can increase in value with an increasing VA fund level for the first α years after annuitization. After the end of this period, the guarantee base remains at the same level for the remainder of the product’s lifetime. This ratchet feature, where the guarantee base steps up to the fund value if this is greater at the rebalancing date, is capped at a maximum year-on-year increase c in the guarantee base. This will become clearer when the cash flows are described mathematically shortly.

The underlying VA fund, which is initially funded by the policyholder premium, is invested wholly in a single equity index (S_t) with dynamics governed by the Heston–CIR model described by (3.1). Possible extensions of this analysis include investing in a mixture of equities and bonds or in a portfolio of two different equity indices.

To model the cash flows on this policy we denote the fund value and guarantee base at the end of year t by F_t and G_t, where time is measured in years since annuitization. An income I_t is paid at time t in respect of year t. Note that F_t refers to the fund value immediately before an income payment.

Also, let R_t denote the equity index return from time $t-1$ to time t minus the management fees. That is, $R_t = S_t/S_{t-1} - \eta^{mc}$, where η^{mc} is the fund management charge, quoted as an annual percentage.

The policyholder invests in ψ units of the equity index with initial value S_0 so that the initial fund value is $F_0 = \psi S_0$. The guarantee base is initially set at the level $G_0 = F_0$. At time $t = 1$ the fund value is given by

$$F_1 = \psi S_0 R_1. \tag{4.1}$$

The first income I_1 is then paid as a percentage of the guarantee base as described below. Thereafter we can track the fund value throughout the lifetime of the policy using the equation

$$F_t = \max((F_{t-1} - I_{t-1}) R_t, 0), \quad 2 \le t \le T. \tag{4.2}$$

The guarantee base at the end of year t after annuitization can be expressed as

$$G_t = \begin{cases} \min(\max(G_{t-1}, F_t), (1+c)G_{t-1}), & 1 \le t \le \alpha, \\ G_{t-1}, & \alpha < t \le T. \end{cases} \tag{4.3}$$

The income level the policyholder withdraws from the policy fund value at the end of year t is given by

$$I_t = w G_t. \tag{4.4}$$

Here, w is a fixed parameter dictating the proportion of the guarantee base that is withdrawn by the policyholder at each annual re-balancing date.

The liability the insurer faces from issuing this VA contract on the market, measured at annuitization, can be expressed as

$$L = \sum_{t=1}^{T} D_t\, p_t^{surv} \max(I_t - F_t, 0) \tag{4.5}$$

where D_t is the discounting factor, p_t^{surv} is the probability of the policy remaining in force until year t after annuitization (encompassing both mortality and lapsation) and T is the maximum contract term. We have used the same mortality table as used in Ledlie *et al.* (2008), i.e. the probability of death in any year is based on the RMC00 tables (medium cohort) with a minimum improvement rate of 1% a year. Clearly, the insurer only faces a liability when the policyholder income cannot be met by the VA fund level. In other words, the shortfall an insurer faces is greater than zero only if F_t drops below I_t at any time $t \leq T$.

There are many ways for the insurer to charge for this liability. For example, the insurer can impose a guarantee charge η^{guar} taken as a percentage of the guarantee base each year. Suppose this charge is deducted directly from the income paid to the policyholder, i.e. the policyholder receives net income of $(w - \eta^{guar})G_t$ at the end of each year, and the insurer sets aside the guarantee charge in an account A. The value of this asset can be expressed as

$$A = \sum_{t=1}^{T} D_t \, p_t^{surv} \eta^{guar} G_t,$$

and the fair value of the contract would be the value η^{guar} that sets $\mathbb{E}(L) = \mathbb{E}(A)$.

In the numerical analysis that follows, we will simply specify illustrative values of w and η^{mc}, and report the resulting level of $\mathbb{E}(L)$. We choose the following values for the parameters. The maturity of the contract is $T = 30$. The income withdrawal level is set at $w = 4\%$. The lapse rate is assumed to be $\ell = 4\%$. The management charge is set at $\eta^{mc} = 1.25\%$. The annual cap on the guarantee base ratchet feature is $c = 15\%$ and the ratchet terminates at time $\alpha = 10$.

4.2. Pathwise delta

We begin by developing a PW method for estimating the VA liability sensitivities, for the stylized VA product of the previous section. This method, proposed for a simple VA product by Hobbs *et al.* (2009), is just the natural extension of the PW approach for option sensitivities to the case of a VA product. The liability of a VA product to the insurer is equivalent to the sum of a series of options of increasing maturity. The payoff function in Equation (4.5) is a continuous function of the initial asset price S_0 and fulfils the conditions of Broadie and Glasserman (1996) permitting the interchange of the order of differentiation and integration in (2.1) and the application of the PW approach.

The payoff is

$$\Delta_{\mathrm{PW}} := L'(S_0) = \sum_{t=1}^{T} D_t \, p_t^{surv} \frac{d}{dS_0} \max(I_t - F_t, 0), \tag{4.6}$$

$$= \sum_{t=1}^{T} D_t \, p_t^{surv} I(I_t > F_t) \cdot \left(\frac{dI_t}{dS_0} - \frac{dF_t}{dS_0} \right). \tag{4.7}$$

The problem of estimating the delta sensitivity of the VA liability is now one of estimating the derivative of the fund value, F_t, and income level, I_t, at each time t after annuitization. Appealing to the structure of the product's cash flows these derivatives must be calculated recursively.

Using (4.1), and noting that $\frac{dR_t}{dS_0} = 0$ for all t, the recursion is initialized by

$$\frac{dF_1}{dS_0} = \psi R_1,$$

and then continued by differentiating (4.2) to obtain

$$\frac{dF_t}{dS_0} = I(F_{t-1} > I_{t-1})\Big(\frac{dF_{t-1}}{dS_0} - \frac{dI_{t-1}}{dS_0}\Big)R_t, \quad 2 \le t \le T. \tag{4.8}$$

In view of (4.4) we have that $\frac{dI_t}{dS_0} = w\frac{dG_t}{dS_0}$ and using Equation (4.3) we obtain

$$\frac{dG_t}{dS_0} = I(A_t)\frac{dF_t}{dS_0} + (I(B_t) + (1+c)I(C_t))\,\frac{dG_{t-1}}{dS_0}, \quad 1 \le t \le \alpha. \tag{4.9}$$

Here, A_t, B_t and C_t are events given by

$$A_t = \{G_{t-1} \le F_t \le (1+c)G_{t-1}\}, \quad B_t = \{F_t \le G_{t-1}\}, \quad C_t = \{F_t > (1+c)G_{t-1}\}.$$

Clearly, for $t > \alpha$ we simply have $\frac{dG_t}{dS_0} = \frac{dG_{t-1}}{dS_0}$.

4.3. Pathwise vega and rho

The PW approach to estimating vega and rho is similar to the approach for delta but the calculation of the derivatives is more intricate. The estimators, which we denote by υ_{PW} and ϱ_{PW}, are given by

$$\upsilon_{\mathrm{PW}} = \sum_{t=1}^{T} D_t\, p_t^{surv} I(I_t > F_t)\cdot\Big(\frac{dI_t}{dV_0} - \frac{dF_t}{dV_0}\Big),$$

$$\varrho_{\mathrm{PW}} = \sum_{t=1}^{T} D_t\, p_t^{surv}\bigg(I(I_t > F_t)\cdot\Big(\frac{dI_t}{dr_0} - \frac{dF_t}{dr_0}\Big) - \max(I_t - F_t, 0)\Delta_{0,t}\frac{d\bar{r}_{0,t}}{dr_0}\bigg).$$

To compute these we need to derive recursions for $\frac{dF_t}{dV_0}$, $\frac{dF_t}{dr_0}$, $\frac{dG_t}{dV_0}$ and $\frac{dG_t}{dr_0}$. We concentrate here on the derivatives of the fund value F_t; similar arguments apply to G_t.

Recursions are initiated by using (4.1) to obtain

$$\frac{dF_1}{dV_0} = \psi S_0\frac{dR_1}{dV_0}, \quad \frac{dF_1}{dr_0} = \psi S_0\frac{dR_1}{dr_0}.$$

We note that, in contrast to delta, the return variables $R_t = S_t/S_{t-1} - \eta^{mc}$ do depend on initial volatility and interest rates. The equivalent recursion to (4.8) now takes the form

$$\frac{dF_t}{dV_0} = I(F_{t-1} > I_{t-1})\Big(\Big(\frac{dF_{t-1}}{dV_0} - \frac{dI_{t-1}}{dV_0}\Big)R_t + (F_{t-1} - I_{t-1})\frac{dR_t}{dV_0}\Big), \quad 2 \le t \le T,$$

for V_0 and a similar equation applies for the sensitivity with respect to r_0. Thus we need to be able to compute the sensitivities for the returns R_t. It follows from (3.3) that

$$R_t = \xi_{t-1,t} \exp\left(\left(\bar{r}_{t-1,t} - \frac{1}{2}\bar{\sigma}^2_{t-1,t}\right)\Delta_{t-1,t} + \bar{\sigma}_{t-1,t}\sqrt{\Delta_{t-1,t}}Z_{t-1,t}\right) - \eta^{mc},$$

and hence, by applying the product and chain rules, we have

$$\begin{aligned}\frac{dR_t}{dV_0} &= \frac{\partial R_t}{\partial\bar{\sigma}_{t-1,t}}\frac{d\bar{\sigma}_{t-1,t}}{dV_0} + \frac{\partial R_t}{\partial\xi_{t-1,t}}\frac{d\xi_{t-1,t}}{dV_0}\\ &= \frac{S_t}{S_{t-1}}\left(\sqrt{\Delta_{t-1,t}}Z_{t-1,t} - \bar{\sigma}_{t-1,t}\Delta_{t-1,t}\right)\frac{d\bar{\sigma}_{t-1,t}}{dV_0} + \frac{S_t}{S_{t-1}\xi_{t-1,t}}\frac{d\xi_{t-1,t}}{dV_0}.\end{aligned}$$

where

$$\frac{d\bar{\sigma}_{t-1,t}}{dV_0} = \frac{a_{33}^2}{2\bar{\sigma}_{t-1,t}\Delta_{t-1,t}}\frac{d}{dV_0}\int_{t-1}^{t} V_s ds,$$

and

$$\frac{1}{\xi_{t-1,t}}\frac{d\xi_{t-1,t}}{dV_0} = \sum_{k=1}^{2}\left(a_{3k}\frac{d}{dV_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,k} - \frac{1}{2}a_{3k}^2\frac{d}{dV_0}\int_{t-1}^{t} V_s ds\right).$$

Similarly

$$\begin{aligned}\frac{dR_t}{dr_0} &= \frac{\partial R_t}{\partial\bar{r}_{t-1,t}}\frac{d\bar{r}_{t-1,t}}{dr_0} + \frac{\partial R_t}{\partial\xi_{t-1,t}}\frac{d\xi_{t-1,t}}{dr_0}\\ &= \frac{S_t}{S_{t-1}}\left(\frac{d}{dr_0}\int_{t-1}^{t} r_s ds + a_{32}\frac{d}{dr_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,2}\right).\end{aligned}$$

From (3.8) we see that

$$\frac{d}{dV_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,1} = \frac{1}{\sigma_V}\left(\frac{dV_t}{dV_0} - \frac{dV_{t-1}}{dV_0} - \kappa_V\theta_V\Delta_{t-1,t} + \kappa_V\frac{d}{dV_0}\int_{t-1}^{t} V_s ds\right),$$

and $\frac{d}{dV_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,2}$ and $\frac{d}{dr_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,2}$ follow similarly.

Clearly we need to be able to evaluate or approximate terms like $\frac{d}{dV_0}\int_u^t V_s ds$, $\frac{d}{dr_0}\int_u^t r_s ds$, $\frac{dV_t}{dV_0}$ and $\frac{dr_t}{dr_0}$. Using the quadrature scheme in (3.9) and assuming a discretization of p time steps in the period $t-1$ to t we approximate $\frac{d}{dV_0}\int_u^t V_s ds$ using

$$\frac{d}{dV_0}\int_{t-1}^{t} V_s ds \approx \sum_{i=1}^{p}\left(0.5\frac{dV_{t-1+(i-1)\Delta^{(p)}_{t-1,t}}}{dV_0} + 0.5\frac{dV_{t-1+i\Delta^{(p)}_{t-1,t}}}{dV_0}\right)\Delta^{(p)}_{t-1,t},$$

with $\Delta^{(p)}_{t-1,t} = 1/p.\ \frac{d}{dr_0}\int_{t-1}^{t} r_s ds$ follows similarly. Calculation of $\frac{dV_t}{dV_0}$ and $\frac{dr_t}{dr_0}$ at different times t can be carried out recursively. For ease of exposition consider a single discretization step in the period $t-1$ to t. For the GL scheme of (3.7) we have that

$$\frac{dV_t}{dV_0} = k_1 k_2 \left(1 + \frac{Z_t^{(GL)}}{\sqrt{k_2 V_{t-1}}}\right)\frac{dV_{t-1}}{dV_0},$$

initialising at $\frac{dV_0}{dV_0} = 1$. The calculation of $\frac{dV_t}{dV_0}$ for the GQE scheme is described in Appendix 6. Computation of $\frac{dr_t}{dr_0}$ follows similarly.

4.4. Conditional LR method for delta, gamma, vega and rho

We first note that the loss function in (4.5) can be written as $L = \sum_{t=1}^{T} L_t$, where each L_t represents a path dependent payoff function. The LR weights required to implement an estimator of (3.4) for each of the L_t are the quantities

$$W_t = \sum_{j=1}^{t} \frac{d}{d\theta} \ln g_j(S_j | S_{j-1}, H_t), \tag{4.10}$$

and the form of the conditional density is given in (3.5).

For delta it is clear that the only term in the sum in (4.10) that involves S_0 is the first term $\ln g_1(S_1|S_0, H_t)$ and the derivative may be computed to be

$$\frac{d}{dS_0} \ln g_1(S_1|S_0, H_t) = \frac{\ln\left(\frac{S_1}{S_0 \xi_{0,1}}\right) - \left(\bar{r}_{0,1} - \frac{1}{2}\bar{\sigma}^2_{0,1}\right)\Delta_{0,1}}{S_0 \bar{\sigma}^2_{0,1}\Delta_{0,1}}.$$

Using (3.3), this may be expressed in terms of a standard normal shock $Z_{0,1}$ to obtain the weight

$$W_t^{\text{delta}} = \frac{Z_{0,1}}{S_0 \bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}.$$

The unbiased quantity on which the estimator is based is

$$\Delta_{\text{LR}} = \sum_{t=1}^{T} L_t W_t^{\text{delta}} = \sum_{t=1}^{T} D_t\, p_t^{surv} \max(I_t - F_t, 0) \frac{Z_{0,1}}{S_0 \bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}. \tag{4.11}$$

This is a standard calculation for a path-dependent option and we also use the standard conditional LR estimator Γ_{LR} which is derived in similar way. For more details see Glasserman (2003) and Broadie and Kaya (2004).

For the vega and rho estimators we again apply the chain rule and product rule. The LR weight for vega is given by

$$W_t^{\text{vega}} = \sum_{j=1}^{t}\left(\frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \bar{\sigma}_{j-1,j}}\frac{d\bar{\sigma}_{j-1,j}}{dV_0} + \frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \xi_{j-1,j}}\frac{d\xi_{j-1,j}}{dV_0}\right),$$

where

$$\frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \bar{\sigma}_{j-1,j}} = \frac{Z_{j-1,j}^2 - 1}{\bar{\sigma}_{j-1,j}} - Z_{j-1,j}\sqrt{\Delta_{t-1,t}},$$
$$\frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \xi_{j-1,j}} = \frac{Z_{j-1,j}}{\xi_{j-1,j}\bar{\sigma}_{j-1,j}\sqrt{\Delta_{t-1,t}}}.$$

For rho, the LR weight is given by

$$W_t^{\text{rho}} = \sum_{j=1}^{t}\left(\frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \bar{r}_{j-1,j}}\frac{d\bar{r}_{j-1,j}}{dr_0} + \frac{\partial \ln g_j(S_j|S_{j-1}, H_t)}{\partial \xi_{j-1,j}}\frac{d\xi_{j-1,j}}{dr_0}\right)$$
$$= \frac{Z_{j-1,j}}{\bar{\sigma}_{j-1,j}\sqrt{\Delta_{t-1,t}}}\left(\frac{d}{dr_0}\int_{t-1}^{t} r_s ds + a_{32}\frac{d}{dr_0}\int_{t-1}^{t}\sqrt{V_s}dW_{s,2}\right).$$

However, the payoff also has an explicit dependence on r_0 through the discount factors $D_t = \exp(-\int_0^t r_s ds)$ and so this is a situation where formula (2.4) is used to estimate the sensitivity. We require the partial derivative of the payoff with respect to r_0, which is given by

$$\frac{\partial L}{\partial r_0} = -\sum_{t=1}^{T} D_t\, p_t^{surv}\max(I_t - F_t, 0)\Delta_{0,t}\frac{d\bar{r}_{0,t}}{dr_0}.$$

Putting everything together, the conditional LR estimators for vega and rho, denoted by υ_{LR} and ϱ_{LR}, are given by

$$\upsilon_{\text{LR}} = \sum_{t=1}^{T} D_t\, p_t^{surv}\max(I_t - F_t, 0)\, W_t^{\text{vega}},$$
$$\varrho_{\text{LR}} = \sum_{t=1}^{T} D_t\, p_t^{surv}\max(I_t - F_t, 0)\left(W_t^{\text{rho}} - \Delta_{0,t}\frac{d\bar{r}_{0,t}}{dr_0}\right).$$

4.5. Second-order mixed estimators

Now that we have the PW and LR estimators for the first-order sensitivities, we can combine them to obtain second-order sensitivities. For example, in the case of the gamma sensitivity to the initial stock price, applying the PW method to the conditional LR estimator in (4.11) yields

$$\begin{aligned}\Gamma_{\text{LR–PW}} &= \frac{d}{dS_0}\left(\sum_{t=1}^{T} D_t\, p_t^{surv} \max(I_t - F_t, 0)\left(\frac{Z_{0,1}}{S_0\bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}\right)\right)\\ &= \left(\frac{Z_{0,1}}{S_0\bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}\right)\sum_{t=1}^{T} D_t\, p_t^{surv}\, I(I_t > F_t)\cdot\left(\frac{dI_t}{dS_0} - \frac{dF_t}{dS_0}\right)\\ &\quad -\left(\frac{Z_{0,1}}{S_0^2\bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}\right)\sum_{t=1}^{T} D_t\, p_t^{surv} \max(I_t - F_t, 0).\end{aligned}$$

Applying the LR weight for S_0 to the PW estimator in (4.6) is trickier because the PW estimator has an explicit dependence on S_0. Moreover, in the context of estimating gamma for a vanilla European option it is reported in Glasserman (2003) (Table 7.2) that the mixed estimator that is obtained by applying the LR method followed by the PW approach (LR–PW) is more accurate than the converse approach in which LR weights are applied to the PW estimator (PW–LR).

For this reason we concentrate on the LR–PW strategy and also apply it to the other second-order Greeks. Thus $\frac{d\upsilon_{\text{LR}}}{dS_0}$ represents the vanna sensitivity estimator derived by applying the PW method for the sensitivity to S_0 to the LR estimator of vega; we also consider $\frac{d\Delta_{\text{LR}}}{dV_0}$, which is the vanna sensitivity estimator calculated by applying the PW method for the sensitivity to V_0 to the LR estimator of delta. Clearly, the estimates obtained in practice should agree closely.

This should also be the case for $\frac{d\varrho_{\text{LR}}}{dS_0}$ and $\frac{d\Delta_{\text{LR}}}{dr_0}$ which represent the two alternative mixed estimators for the second order sensitivity with respect to S_0 and r_0. The full set of estimators of this kind that we consider is

$$\begin{aligned}\frac{d\upsilon_{\text{LR}}}{dS_0} &= \sum_{t=1}^{T} D_t\, p_t^{surv}\, I(I_t > F_t)\cdot\left(\frac{dI_t}{dS_0} - \frac{dF_t}{dS_0}\right) W_t^{\text{vega}},\\ \frac{d\Delta_{\text{LR}}}{dV_0} &= \left(\frac{\upsilon_{\text{PW}} Z_{0,1}}{S_0\bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}\right) - \frac{\Delta_{\text{LR}} a_{33}^2}{2\bar{\sigma}_{0,1}^2\Delta_{0,1}}\frac{d}{dV_0}\int_0^1 V_s\,ds,\\ \frac{d\varrho_{\text{LR}}}{dS_0} &= \sum_{t=1}^{T} D_t\, p_t^{surv}\, I(I_t > F_t)\cdot\left(\frac{dI_t}{dS_0} - \frac{dF_t}{dS_0}\right)\left(W_t^{\text{rho}} - \Delta_{0,t}\frac{d\bar{r}_{0,t}}{dr_0}\right),\\ \frac{d\Delta_{\text{LR}}}{dr_0} &= \frac{\varrho_{\text{PW}} Z_{0,1}}{S_0\bar{\sigma}_{0,1}\sqrt{\Delta_{0,1}}}.\end{aligned}$$

TABLE 1

FIVE SPECIFICATIONS (LABELLED A–E) FOR THE ASSET PRICE MODEL ARE CONSIDERED. REFER TO (3.1) FOR EXPLANATION OF PARAMETERS.

Case	κ_V	θ_V	σ_V	κ_r	θ_r	σ_r	$\rho_{S,V}$	$\rho_{S,r}$	$\rho_{V,r}$
A	2	0.04	0.15	1.5	0.04	0.1	−0.7	−0.3	0
B	1	0.04	0.3	1.5	0.04	0.1	−0.7	−0.3	0
C	2	0.04	0.15	0.75	0.04	0.2	−0.7	−0.3	0
D	1	0.04	0.3	0.75	0.04	0.2	−0.7	−0.3	0
E	1	0.04	0.3	0.75	0.04	0.2	−0.9	−0.3	0

5. NUMERICAL COMPARISON OF ESTIMATORS

In our experiments we use the GQE and GL simulation methods as described in Section 3.3. Five specifications for the asset-price model are considered, labelled A–E in Table 1. These give different parameter settings for the Heston and CIR processes and different correlations between the normal shocks driving the variance, interest rate and equity processes.

The parameter choices are based on discussions with users of these models in the economic scenario generation context. Since the volatility and mean reversion rate parameters in the dynamics of volatility and interest rates are the hardest parameters to calibrate, we experiment with different values.

Case A represents our base case. In case B we half the mean reversion rate and double the instantaneous volatility of the equity volatility process. In case C we half the mean reversion rate and double the instantaneous volatility of the interest rate process. In case D we half the mean reversion rate and double the instantaneous volatility for both processes. Case E is similar to case D except that we use a higher value for $\rho_{S,V}$. Note that we have set $\rho_{V,r}$ to be equal to zero, which imposes the constraint that $\rho^2_{S,V} + \rho^2_{S,r} < 1$ to guarantee a positive-definite correlation matrix ρ. In all cases we use the starting values $V_0 = \theta_V$ and $r_0 = \theta_r$.

We analyse the stylized VA product described in Section 4.1. The initial policyholder premium is set to be \$100, invested in 1 unit of equity index, with $S_0 = 100$. In the BnR method the bump perturbation size is set at 0.5%. There is a trade-off to be made here because smaller perturbations reduce the bias in the estimator but increase its variance, particularly for the second-order sensitivities. We use the conditional sampling setup described in Section 3.2 for all methods to facilitate comparison.

5.1. Effect of varying annual time steps in discretization

First, we focus on case A where the parameters are relatively benign. Figure 2 shows how the expected liability $\mathbb{E}(L)$, the PW delta estimate Δ_{PW} and the mixed gamma estimate $\Gamma_{\text{LR–PW}}$ vary as we increase the number of time steps per year

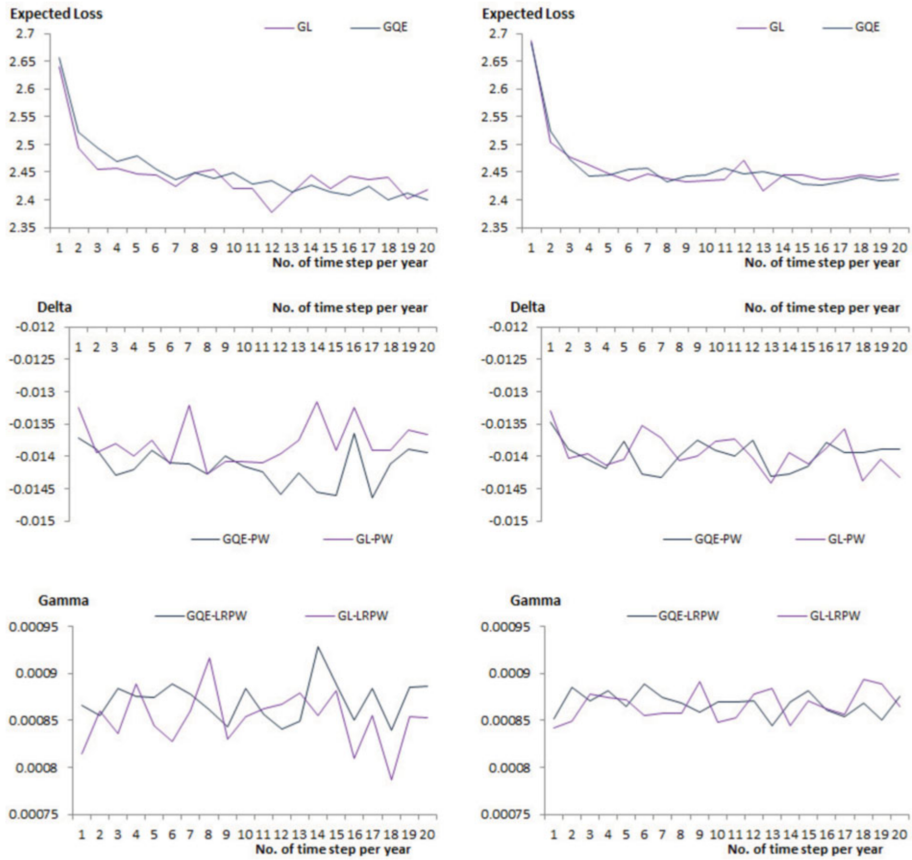

FIGURE 2: Estimates of expected liability $E(L)$, PW delta Δ_{PW} and mixed gamma $\Gamma_{\text{LR-PW}}$ using the GQE and GL schemes with different numbers of time steps per year. Left pictures show results for 10 inner paths and 10,000 outer paths; right pictures show results for 20 inner paths and 20,000 outer paths. (Color online)

in the discretization for both the GQE and GL schemes. We give results for different numbers of inner and outer paths.

We see that both schemes produce similar result. The estimated expected liability requires a larger number of time steps before the value stabilizes, due to the use of quadrature approximation. The estimated delta and gamma seem to be less affected by the number of time steps. The volatility of the estimator can be reduced by increasing the number of paths, as can be seen by comparing the graphs on the left with those on the right.

5.2. Comparison of estimation methods

We now turn to the comparison between the various estimation methods. Figures 3 and 4 show delta estimates obtained by the BnR, PW and LR methods and gamma estimates obtained by the BnR, LR and mixed LR–PW

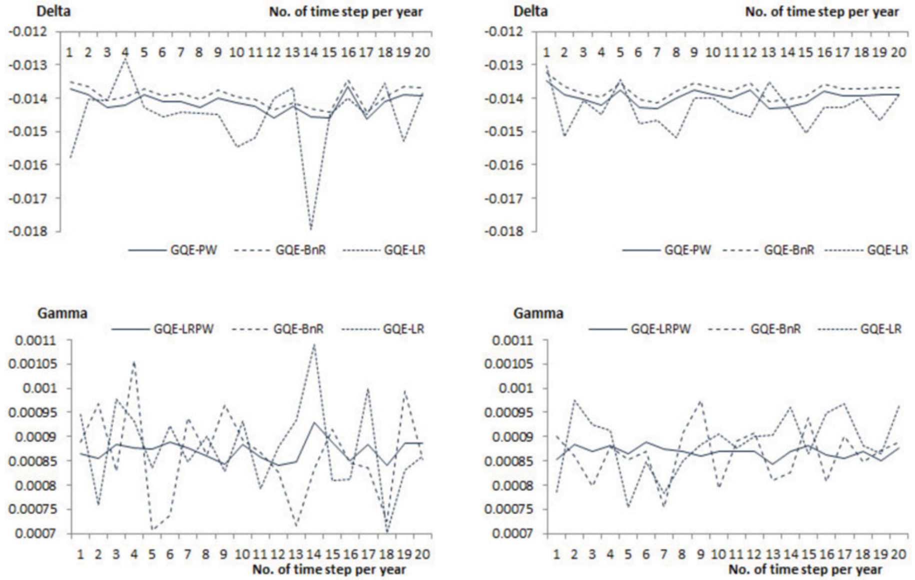

FIGURE 3: Estimates of delta and gamma using different estimation methods and different numbers of time steps in the GQE scheme. The methods for delta are BnR, PW and LR; for gamma the methods are BnR, LR and LR-PW. Left pictures show results for 10 inner paths and 10,000 outer paths; right pictures show results for 20 inner paths and 20,000 outer paths. (Color online)

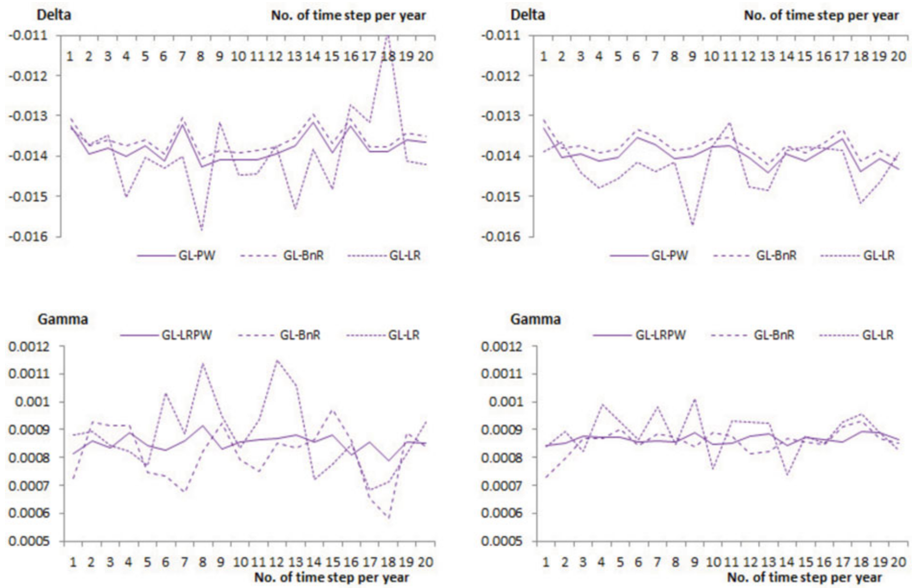

FIGURE 4: Estimates of delta and gamma using different estimation methods and different numbers of time steps in the GL scheme. The methods for delta are BnR, PW and LR; for gamma the methods are BnR, LR and LR-PW. Left pictures show results for 10 inner paths and 10,000 outer paths; right pictures show results for 20 inner paths and 20,000 outer paths. (Color online)

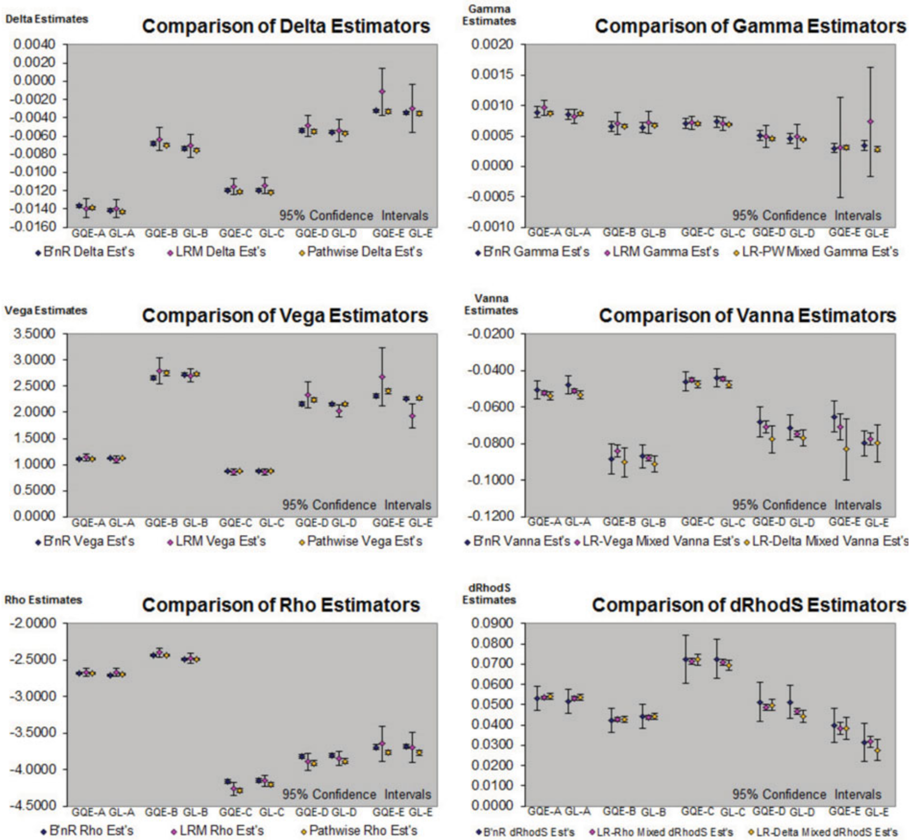

FIGURE 5: Boxplots of the estimates of the Greeks for the different estimation methods based on 20 inner paths and 20,000 outer paths. Results are shown for the five different parameterisations A–E and both the GQE and GL sampling schemes with 20 time steps per year. (Color online)

methods. The estimates are obtained for different numbers of time steps in the discretization. In Figure 3 the GQE scheme is used while in Figure 4 the GL scheme is used.

Consistent with the previous analysis, both the delta and gamma estimates seem to be rather robust with respect to the number of time steps used in the discretization.

The delta estimates obtained using the LR method are much more volatile than those calculated using the PW and BnR methods. Also, there is a small but noticeable bias in the BnR when compared with the unbiased PW method.

For gamma, both the LR and BnR methods produce estimates that are much more volatile than the mixed LR–PW method. The volatility can be reduced by using larger number of inner and outer paths. The bias in BnR values for gamma is less evident as the estimates are dominated by the large volatility.

The differing variances of the various estimators can be seen more clearly in the boxplots of Figure 5. For delta the BnR estimator has approximately the same standard error as the PW estimator, whereas the LR estimator has a much larger standard error. However, for gamma, the mixed LR–PW estimator is clearly less volatile than both the BnR and LR estimators. For the cross gammas, the estimators $\frac{dv_{\text{LR}}}{dS_0}$ and $\frac{d\varrho_{\text{LR}}}{dS_0}$ seem to be more efficient. Figure 5 shows results for the full set of Greeks considered in this paper. In general we conclude that the PW method is preferable for first-order sensitivities (although the BnR method is also quite accurate) whereas the mixed LR–PW estimators are preferred for second-order sensitivities.

Point estimates and standard errors for the expected liability $\mathbb{E}(L)$ and the Greeks in all five cases are shown in Table 2 for reference. These results are for the case where the sampling scheme is GQE with 20 inner paths and 20,000 outer paths and where 20 time steps are used in the discretization of each year. As indicated in Figure 5, very similar results are obtained when the GL sampling method is used.

Note that the expected liability $\mathbb{E}(L)$ varies as we move from case A to case E. In general, increasing the volatility of the equity volatility process or increasing $\rho_{S,V}$ will increase $\mathbb{E}(L)$, whereas increasing volatility of the interest rate process will decrease $\mathbb{E}(L)$ (due to greater discounting).

It is natural to consider the question of how best to choose the number of outer and inner simulations. Our experience suggests that using multiple inner and outer paths is more efficient than using a single inner path. The optimal choice depends on the quantity of interest, since more inner paths correspond to more asset price paths, and more outer paths correspond to more interest rate and volatility paths. Thus, if we are interested in computing the price and the delta, we should generate more inner paths, and if we are more interested in vega and rho, we should generate more outer paths.

5.3. Comparison of computational time

Finally we look at the computational time required for each of the schemes. Table 3 shows the time taken in minutes to compute increasing number of Greeks in both the GQE and GL schemes. The first-order Greeks are calculated using the PW method and the second-order Greeks by the mixed LR–PW method and this approach is compared to the use of BnR for all Greeks. In general, as the number of Greeks to be calculated increases, the more advanced approach requires increasingly less time than the BnR method. The difference in computing time is due to the difference in extra time required to compute an additional path versus the extra time required to compute the estimator. The separation is much clearer for the GQE scheme, as it is computationally more expensive than the GL scheme.

6. Conclusion

With the increasing popularity of VA products it is essential that insurers can employ an effective hedging strategy for mitigating the risks inherent in the

TABLE 2

POINT ESTIMATES OF THE EXPECTED LIABILITY AND THE GREEKS FOR CASES A TO E. SIMULATIONS ARE CARRIED OUT USING THE GQE SCHEME WITH 20 INNER PATHS AND 20,000 OUTER PATHS. 20 TIME STEPS PER YEAR ARE USED IN THE DISCRETIZATION.

	Case A	Case B	Case C	Case D	Case E
$\mathbb{E}(L)$	2.438 (0.00502)	2.733 (0.00620)	1.861 (0.00463)	2.170 (0.00590)	2.262 (0.00622)
Δ_{PW}	−0.0139 (0.00010)	−0.0070 (0.00010)	−0.0121 (0.00009)	−0.0055 (0.00009)	−0.0033 (0.00009)
Δ_{LR}	−0.0139 (0.00052)	−0.0064 (0.00063)	−0.0116 (0.00045)	−0.0049 (0.00061)	−0.0012 (0.00131)
Δ_{BnR}	−0.0137 (0.00010)	−0.0068 (0.00010)	−0.0120 (0.00009)	−0.0054 (0.00009)	−0.0032 (0.00009)
Γ_{LR}	0.00096 (0.00006)	0.00071 (0.00009)	0.00072 (0.00005)	0.00049 (0.00009)	0.00031 (0.00042)
$\Gamma_{\text{LR-PW}}$	0.00088 (0.00001)	0.00066 (0.00001)	0.00070 (0.00001)	0.00045 (0.00001)	0.00031 (0.00002)
Γ_{BnR}	0.00089 (0.00005)	0.00066 (0.00004)	0.00071 (0.00004)	0.00051 (0.00004)	0.00030 (0.00004)
υ_{PW}	1.107 (0.00654)	2.754 (0.02628)	0.869 (0.00629)	2.239 (0.02221)	2.407 (0.02342)
υ_{LR}	1.136 (0.03344)	2.801 (0.12736)	0.858 (0.02912)	2.335 (0.12487)	2.683 (0.28393)
υ_{BnR}	1.107 (0.00654)	2.660 (0.02285)	0.868 (0.00629)	2.168 (0.01942)	2.312 (0.02147)
ϱ_{PW}	−2.675 (0.00733)	−2.440 (0.00755)	−4.287 (0.01437)	−3.908 (0.01490)	−3.765 (0.01505)
ϱ_{LR}	−2.667 (0.02677)	−2.397 (0.03372)	−4.264 (0.04520)	−3.888 (0.05922)	−3.645 (0.12324)
ϱ_{BnR}	−2.679 (0.00740)	−2.438 (0.00758)	−4.161 (0.01459)	−3.817 (0.01494)	−3.69 (0.01499)
$\frac{d}{dS_0}\upsilon_{\text{LR}}$	−0.0524 (0.00056)	−0.0838 (0.00168)	−0.0449 (0.00054)	−0.0710 (0.00167)	−0.0708 (0.00354)
$\frac{d}{dV_0}\Delta_{\text{LR}}$	−0.0536 (0.00111)	−0.0903 (0.00393)	−0.0473 (0.00100)	−0.0776 (0.00379)	−0.0832 (0.00848)
$\frac{d}{dS_0}\upsilon_{\text{BnR}}$	−0.0504 (0.00249)	−0.0882 (0.00422)	−0.0460 (0.00260)	−0.0681 (0.00420)	−0.0651 (0.00441)
$\frac{d}{dS_0}\varrho_{\text{LR}}$	0.0534 (0.00045)	0.0427 (0.00047)	0.0714 (0.00079)	0.0488 (0.00084)	0.0384 (0.00161)
$\frac{d}{dr_0}\Delta_{\text{LR}}$	0.0543 (0.00073)	0.0430 (0.00073)	0.0722 (0.00133)	0.0498 (0.00143)	0.0383 (0.00271)
$\frac{d}{dS_0}\varrho_{\text{BnR}}$	0.0531 (0.00311)	0.0423 (0.00298)	0.0724 (0.00594)	0.0514 (0.00496)	0.0397 (0.00430)

TABLE 3

TIME TAKEN IN MINUTES TO COMPUTE DIFFERENT NUMBERS OF GREEKS USING THE GQE AND GL SCHEMES, FOR PW AND LR METHOD AND BNR METHOD IN BRACKETS, SIMULATED USING 10 INNER PATHS AND 10,000 OUTER PATHS, AND 20 TIME STEPS PER YEAR.

	GQE	GL
Δ	0.586 (0.559)	0.335 (0.332)
Δ, ϱ	0.999 (1.059)	0.635 (0.613)
$\Delta, \varrho, \upsilon$	1.248 (1.553)	0.889 (0.912)
$\Delta, \varrho, \upsilon, \Gamma$	1.322 (1.591)	0.948 (0.951)
$\Delta, \varrho, \upsilon, \Gamma, \frac{d\upsilon}{dS_0}$	1.368 (1.669)	1.001 (1.019)
$\Delta, \varrho, \upsilon, \Gamma, \frac{d\upsilon}{dS_0}, \frac{d\varrho}{dS_0}$	1.437 (1.732)	1.033 (1.048)

liabilities arising from such products. The recent financial crisis has demonstrated that under turbulent market conditions a hedging portfolio can require much more frequent rebalancing. The widely adopted BnR approach for estimating the Greeks has some shortcomings, such as an inherent bias and low computational efficiency due to the necessity of running additional perturbed simulations. The latter problem is particularly relevant when large numbers of Greeks need to be estimated to implement hedging strategies based on both first-order and second-order sensitivities. Improvements in computational efficiency could be of great benefit to practitioners in managing a hedging strategy for their VA books.

In this article alternative estimators for the VA Greeks based on the PW and LR methods have been developed and implemented for a stylized VA product in the context of an extended Heston model for the underlying asset price. The model incorporates stochastic volatility and stochastic interest rates.

We conclude that for first-order derivatives, the PW method is preferable to the BnR method because it is unbiased and the computational time required is similar to the BnR method. For second-order sensitivities the mixed estimator gives estimates with much smaller standard errors compared to the BnR method in most cases. Furthermore, the bias-variance trade-off in choosing the perturbation size in the BnR method is avoided. The computational gain in using PW and mixed estimators increases as the required number of Greeks increases.

In future work it would be interesting to extend the setup to include a non-zero correlation between volatility and interest rate shocks. While model paths from such a model could be sampled in a simple manner using Euler and related schemes, it would be more accurate to use an exact bivariate sampling method for arbitrary time steps that extends the univariate GQE and GL approaches; to our knowledge, such a scheme is not currently available. A further topic of interest would be the incorporation into our framework of more sophisticated assumptions and models for lapsation behaviour, which may have a large effect on the value of the expected liability and the sensitivities. The optimization of the trade-off between the number of inner and outer paths is also a potential topic for future research.

REFERENCES

ANDERSEN, L. (2008) Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance*, **11**, 1–42.

BACINELLO, A.R., MILLOSSOVICH, P., OLIVIERI, A. and POTACCO, E. (2011) Variable annuities: A unifying valuation approach. *Insurance: Mathematics and Economics*, **49**, 285–297.

BAUER, D.K., KLING, A. and RUSS, J. (2008) A universal pricing framework for guaranteed minimum benefits in variable annuities. *Astin Bulletin*, **38**, 621–650.

BROADIE, M. and GLASSERMAN, P. (1996) Estimating security price derivatives using simulation. *Journal of Economic Dynamics and Control*, **21**, 1323–1352.

BROADIE, M. and KAYA, O. (2004) Exact simulation of option Greeks under stochastic volatility and jump diffusion models. In *Proceedings of the 2004 Winter Simulation Conference*, Washington DC.

CHAN, J.H. and JOSHI, M.S. (2013) Fast and accurate long-stepping simulation of the Heston stochastic volatility model. *Journal of Computational Finance*, **16**, 47–97.

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

DAI, M., KWOK, Y.K. and ZONG, J. (2008) Guaranteed minimum withdrawal benefit in variable annuities. *Mathematical Finance*, **18**, 595–611.

GLASSERMAN, P. (2003) *Monte Carlo Methods in Financial Engineering*. New York: Springer.

HESTON, S.L. (1993) A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, **6**, 327–343.

HOBBS, C., KRISHNARAJ, B., LIU, Y. and MUSSELMAN, J. (2009) Calculation of variable annuity market sensitivities using a pathwise methodology. *Life and Pensions*, September, 40–44.

JAIMUNGAL, S., DONNELLY, R. and RUBISOV, D.H. (2014) Valuing GWBs with stochastic interest rates and volatility. *Quantitative Finance*, **14**, 369–382.

KLING, A., RUEZ, F. and RUSS, J. (2011) The impact of stochastic volatility on pricing, hedging, and hedge efficiency of withdrawal benefit guarantees in variable annuities. *Astin Bulletin*, **41**, 511–545.

LEDLIE, M.C., CORRY, D.P., FINKELSTEIN, G.S., RITCHIE, K.S. and WILSON, D.C.E. (2008) Variable annuities. *British Actuarial Journal*, **14**, 327–389.

MILEVSKY, M.A. and SALISBURY, T.S. (2006) Financial valuation of guaranteed minimum withdrawal benefits. *Insurance: Mathematics and Economics*, **38**, 21–38.

PENG, J., LEUNG, K.S. and KWOK, Y.K. (2012) Pricing guaranteed minimum withdrawal benefits under stochastic interest rate. *Quantitative Finance*, **12**, 933–941.

MARK J. CATHCART
Standard Life Group
Edinburgh EH1 2DH, UK
E-mail: mark.j.cathcart@gmail.com

HSIAO YEN LOK
Heriot-Watt University, Edinburgh
Maxwell Institute for Mathematical Sciences
Edinburgh EH14 4AS, UK
E-mail: yhl30@hw.ac.uk

ALEXANDER J. MCNEIL (Corresponding author)
Heriot-Watt University, Edinburgh
Maxwell Institute for Mathematical Sciences
Edinburgh EH14 4AS, UK
E-mail: a.j.mcneil@hw.ac.uk

STEVEN MORRISON
Moody's Analytics
Edinburgh EH3 8RD, UK
E-mail: steven.morrison@moodys.com

APPENDIX A. CALCULATING $\frac{dV_t}{dV_0}$ IN THE GQE SCHEME

We assume a discretization of one time step per year and define

$$K_0 = \frac{2\kappa_V}{\sigma_V^2(e^{\kappa_V/p} - 1)},$$

$$K_{t,1} = K_0(\omega_t - \lambda_t^2 e^{-\lambda_t})\frac{dV_{t-1}}{dV_0},$$

$$K_{t,2} = \lambda_t(1 - \omega_t),$$

$$K_{t,3} = (1 - P_{Y_t^{(2)}})(1 + b_t^2),$$

$$K_{t,4} = K_{t,2} - (1 - P_{Y_t^{(2)}})(2 + K_{t,2}),$$

$$K_{t,5} = \Phi^{-1}\left(\frac{U_{Y_t^{(2)}}}{1 - P_{Y_t^{(2)}}}\right),$$

$$K_{t,6} = \Phi^{-1}\left(\frac{U_{Y_t^{(2)}} - P_{Y_t^{(2)}}}{1 - P_{Y_t^{(2)}}}\right).$$

Then we have

$$\frac{dV_t}{dV_0} = \frac{dY_t^{(1)}}{dV_0} + \frac{dY_t^{(2)}}{dV_0},$$

where

$$\frac{dY_t^{(1)}}{dV_0} = \left(\frac{\mu_t}{P_{Y_t^{(1)}} - 1}\frac{dP_{Y_t^{(1)}}}{dV_0} + \ln\left(\frac{1 - P_{Y_t^{(1)}}}{1 - U_{Y_t^{(1)}}}\right)\frac{d\mu_t}{dV_0}\right) I(P_{Y_t^{(1)}} < U_{Y_t^{(1)}} \leq 1),$$

$$\frac{dY_t^{(2)}}{dV_0} = \left((b_t + K_{t,5})^2\frac{da_t}{dV_0} + 2a_t(b_t + K_{t,5})\left(\frac{db_t}{dV_0} + \frac{dK_{t,5}}{dV_0}\right)\right) I(U_{Y_t^{(2)}} \leq c_{t,1})$$
$$+ \left((b_t + K_{t,6})^2\frac{da_t}{dV_0} + 2a_t(b_t + K_{t,6})\left(\frac{db_t}{dV_0} + \frac{dK_{t,6}}{dV_0}\right)\right) I(U_{Y_t^{(2)}} > c_{t,2}),$$

$$\frac{dK_{t,2}}{dV_0} = K_0\frac{dV_{t-1}}{dV_0} - K_{t,1}, \quad \frac{dK_{t,3}}{dV_0} = (-1 - b_t^2)\frac{dP_{Y_t^{(2)}}}{dV_0} + 2b_t(1 - P_{Y_t^{(2)}})\frac{db_t}{dV_0},$$

$$\frac{dK_{t,4}}{dV_0} = P_{Y_t^{(2)}}\frac{dK_{t,2}}{dV_0} + (2 + K_{t,2})\frac{dP_{Y_t^{(2)}}}{dV_0},$$

$$\frac{dK_{t,5}}{dV_0} = \frac{\sqrt{2\pi}\ e^{0.5K_{t,5}^2} U_{Y_t^{(2)}}}{(1 - P_{Y_t^{(2)}})^2} \frac{dP_{Y_t^{(2)}}}{dV_0}, \quad \frac{dK_{t,6}}{dV_0} = \frac{\sqrt{2\pi}\ e^{0.5K_{t,6}^2} (U_{Y_t^{(2)}} - 1)}{(1 - P_{Y_t^{(2)}})^2} \frac{dP_{Y_t^{(2)}}}{dV_0},$$

$$\frac{dP_{Y_t^{(1)}}}{dV_0} = \frac{-4K_{t,1}}{(2 + \lambda_t \omega_t)^2}, \quad \frac{dP_{Y_t^{(2)}}}{dV_0} = \frac{e^{\lambda_t}}{P_{Y_t^{(1)}}^2} \left(P_{Y_t^{(1)}} K_0 \frac{dV_{t-1}}{dV_0} - \frac{dP_{Y_t^{(1)}}}{dV_0} \right),$$

$$\frac{d\mu_t}{dV_0} = \frac{\beta_V \left((1 - P_{Y_t^{(1)}}) K_{t,1} + \omega_t \lambda_t \frac{dP_{Y_t^{(1)}}}{dV_0} \right)}{(1 - P_{Y_t^{(1)}})^2},$$

$$\frac{dc_t}{dV_0} = \frac{2K_{t,4} \frac{dK_{t,2}}{dV_0} - 2K_{t,2} \frac{dK_{t,4}}{dV_0}}{K_{t,4}^2}, \quad \frac{db_t}{dV_0} = \frac{1}{2b_t} \left(\frac{2c_t + 1}{2\sqrt{c(c+1)}} - 1 \right) \frac{dc_t}{dV_0},$$

$$\frac{da_t}{dV_0} = \frac{\beta_V (K_{t,3} \frac{dK_{t,2}}{dV_0} - K_{t,2} \frac{dK_{t,3}}{dV_0})}{K_{t,3}^2}.$$