Variable annuities are often sold with guarantees to protect investors from downside investment risk. The majority of variable annuity guarantees are written on more than one asset, but in practice, single-asset (univariate) stochastic investment models are mostly used for pricing and hedging these guarantees. This practical shortcut may lead to problems such as basis risk. In this article, we contribute a multivariate framework for pricing and hedging variable annuity guarantees. We explain how to transform multivariate stochastic investment models into their risk-neutral counterparts, which can then be used for pricing purposes. We also demonstrate how dynamic hedging can be implemented in a multivariate framework and how the potential hedging error can be quantified by stochastic simulations.

1. INTRODUCTION

The deep financial crisis in 2000 revealed that investment guarantees sold with variable annuities are highly valuable. Since then, insurance companies have been pursuing different risk management strategies for variable annuity guarantees. Because the underlying financial risk is not diversifiable, it cannot be managed with traditional deterministic methods, which rely entirely on margins in financial assumptions. In North America, national organizations of the actuarial profession strongly recommend the use of stochastic approaches to manage the financial risk associated with variable annuity guarantees. We refer interested readers to the American Academy of Actuaries (2005) and the Canadian Institute of Actuaries (2002) for further details.

Unlike fixed annuities, variable annuities allow investors to allocate their assets among various separate account funds, which may be viewed as the functional equivalent of mutual funds. Sellers typically offer more than a dozen separate account funds for investors to choose from. Investors can allocate their assets however they like, usually on a percentage basis. Although the majority of variable annuity guarantees are linked to more than one asset, practitioners mostly use single-asset (univariate) stochastic investment models for pricing and risk management purposes.

To illustrate, let us consider a two-asset portfolio with \( w_1 \) units of asset 1, which has a time-\( t \) price of \( S_{1,t} \), and \( 1 - w_1 \) units of asset 2, which has a time-\( t \) price of \( S_{2,t} \). Assuming no management charge for simplicity, the payoff from a variable annuity guarantee with a guarantee level of \( g \) (expressed as a percentage of the initial value of the policyholder’s portfolio) is given by

\[
(P_0(w_1)g - P_T(w_1))^+,
\]

where \( P_t(w_1) = w_1S_{1,t} + (1 - w_1)S_{2,t} \) is the time-\( t \) value of the policyholder’s portfolio, and \( T \) is the time at which the guarantee benefit is paid. As a practical shortcut, one may model the proxy fund \( P_t \) directly using a single-asset model without modeling \( S_{1,t} \) and \( S_{2,t} \) explicitly. This shortcut somewhat simplifies the actuarial work involved, but it has several significant drawbacks, which we now detail.

- Basis risk: The distribution of the proxy fund \( P_t \) can be complicated and may not necessarily have the same form as that of the constituent assets \( S_{1,t} \) and \( S_{2,t} \). In the simplest case where \( S_{1,t} \) and \( S_{2,t} \) follow geometric Brownian motions, it is well...
known that $P_t(w_1)$ does not follow a geometric Brownian motion unless $w_1$ is 0 or 1. Moreover, if the volatilities of the geometric Brownian motion are very different, then $P_t(w_1)$ may be far from being lognormal. In extreme circumstances, we may not be able to identify a satisfying univariate model for the dynamics of $P_t$. The problem of basis risk has recently attracted considerable attention in the insurance industry (see, e.g., Hadley 2010).

- A lack of flexibility with respect to asset allocations: The distribution of the proxy fund $P_t$ depends on $w_1$, which typically varies among policyholders. The use of a univariate shortcut would therefore require different univariate asset models for different policyholders. The use of a large number of models does not seem to be efficient.

The same problem arises even when we consider only one single policyholder, who may change his or her asset allocation from time to time. Changes in asset allocation over the duration of a policy are not uncommon because, for instance, a policyholder may become more risk adverse as he or she gets older. If the univariate shortcut is used, then the asset model has to be changed every time when the policyholder alters his or her asset allocation.

- Hedging error: In forming a hedge, dynamic or static, for the guarantee liability, we need to know the relevant Greeks, in particular, deltas. If the univariate shortcut is used, we would only be able to compute the delta with respect to the proxy fund $P_t$, which is not actually traded in the market. Although we may approximate the deltas with respect to $S_{1,t}$ and $S_{2,t}$ by using the portfolio weights $w_1$ and $1 - w_1$, the resulting deltas could be inaccurate, because the approximation imposes the constraint that the ratio of the two deltas is constant at any time point and for any values of $S_{1,t}$ and $S_{2,t}$. This inaccuracy may lead to a high hedging error.

The purpose of this article is to put forward a multivariate framework for pricing and hedging variable annuity guarantee liabilities. This multivariate framework is based on a stochastic process that explicitly models $S_{1,t}$, $S_{2,t}$, and the correlation between them, thereby avoiding the problems described above. The proposed framework can work with a wide range of stochastic processes; for example, multivariate generalizations of the regime-switching lognormal (RSLN) model (Hardy 2001) and the generalized autoregressive conditional heteroskedastic (GARCH) model.

The challenging part of the multivariate framework is the identification of a risk-neutral probability measure, which is required for pricing purposes. If one assumes a simple multivariate lognormal model without any serial correlation and regime switching, then the risk-neutral probability measure is unique, and the risk-neutral dynamics of the asset prices can be specified quite easily by replacing the expected rate of appreciation for each asset with the risk-free interest rate. However, more sophisticated stochastic processes, including those analyzed recently by Boudreault and Panneton (2009), would imply market incompleteness. In this situation, there would be more than one risk-neutral probability measure, which in turn means that there are a range of no-arbitrage prices for a variable annuity guarantee. Also, since a guarantee cannot be replicated by the underlying assets and cash perfectly, there are multiple ways to construct a hedge portfolio.

What we need is a method to identify, for common multivariate stochastic investment models, a risk-neutral probability measure that is economically justifiable. We consider the Esscher transform (Gerber and Shiu 1994, 1996), which can be justified within an expected utility framework. The original Esscher transform is applicable to one-period models only. To overcome this problem, Bühmann et al. (1996) introduced the (univariate) conditional Esscher transform, which can be applied to stochastic investment models involving multiple time steps. In a univariate setting, the conditional Esscher transform has been used successfully for pricing stock options (Siu et al. 2004), reverse mortgages (Li et al. 2010), and investment guarantees (Ng et al. 2011). A multivariate extension of the conditional Esscher transform would be a good fit to our proposed framework.

The multivariate Esscher transform was first proposed as a mathematical concept by Gerber and Shiu (1994, Section 7). It is subsequently considered by Kijima (2006), Rombouts and Stentoft (2011), and Ng and Li (2011) for various pricing purposes. We notice that these previous studies are presented in a rather theoretical manner, which could possibly shy practitioners away from using a multivariate framework. In this article, we describe in more detail how this tool can be applied to each type of multivariate stochastic investment model, demonstrating to practitioners that pricing under the multivariate framework can in fact be implemented in practice.

The last part of our proposed framework is dynamic hedging. To our knowledge, this is the first article to consider multivariate dynamic hedging for variable annuity guarantees. We shall explain how, by using a Greek approximation, to ensure that the liability being hedged and the portfolio of hedging instruments have similar price sensitivities to changes in the underlying fund values. Because continuous trading is impossible in practice, the hedge portfolio may not be self-financing, which means additional funds may be required when the hedge is rebalanced. We shall also illustrate how the potential hedging error may be quantified by using stochastic simulations.

The remainder of this article is organized as follows. In Section 2 we review some common multivariate stochastic investment models. In Section 3 we transform the models, which are originally defined in the real-world probability measure, into their risk-neutral counterparts by using the multivariate conditional Esscher transform. In Section 4 we present the pricing formulas for two types of variable annuity guarantees and apply the proposed pricing method to hypothetical variable annuity guarantees that...
are linked to multiple subaccount funds. In Section 5 we demonstrate how a guarantee liability can be hedged dynamically under the multivariate framework. Finally, in Section 6 we discuss our proposed framework and conclude the article with suggestions for future research.

2. COMMON MULTIVARIATE INVESTMENT MODELS

Many actuaries working on stochastic investment modeling are familiar with univariate regime-switching lognormal models and GARCH models. For each of these univariate models, there are several multivariate extensions, which differ in, for example, the way that correlations between log returns are specified. In this article, we focus only on extensions that are relatively easy to interpret and estimate in practice. Readers are referred to Silvennoinen and Terasvirta (2009) and Bauwens et al. (2006) for more comprehensive surveys of multivariate investment models.

Let $S_{i,t}, i = 1, \ldots, N$, be the value of the $i$th equity index (adjusted for dividend income) at time $t$, and let $y_{i,t} = \ln(S_{i,t}/S_{i,t-1})$ be the log total return on the index from time $t-1$ to time $t$. The vector of log total returns on the $N$ equity indices is denoted by $y_t = (y_{1,t}, \ldots, y_{N,t})'$, where $w'$ denotes the transpose of $w$.

2.1. Lognormal Models

In a univariate lognormal model, $y_{1,t} = \mu + a_{1,t}$, where $a_{1,t}$ is normally distributed with zero mean. To extend the above univariate model to a multivariate set up, we let $y_t = \mu + a_t$.

In the above, $a_t \sim \text{MVN}(0, H)$, where MVN stands for a multivariate normal distribution, $0$ is a vector of $N$ zeros, and $H$ is the $N \times N$ covariance matrix for the log total returns.

2.2. GARCH Models

In a univariate GARCH model (Bollerslev 1986; Taylor 1986), $y_t$ is modeled by two submodels, one for the conditional mean and the other for the conditional variance. In particular, an autoregressive moving average specification is used to model the conditional mean:

$$y_t = \phi_0 + \sum_{i=1}^{R} \phi_i y_{t-i} - \sum_{j=1}^{M} \theta_j a_{t-j} + a_t,$$

where $\phi_0$ is the drift term. The coefficients $\phi_i, i = 1, \ldots, R$, and $\theta_j, j = 1, \ldots, M$, govern the relative emphasis on previous log returns and innovations, respectively. Let $\{F_t\}$ be the natural filtration generated by the total return process. It is assumed that $a_t|F_{t-1} \sim \text{N}(0, h_t)$. We assume that $y_t$ is $F_{t-1}$-measurable for each $t$.

The serial dependence of $h_t$ is then captured by the following conditional variance model:

$$h_t = k + \sum_{i=1}^{Q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{P} \beta_j h_{t-j},$$

where $k, \alpha_i, i = 1, \ldots, Q$, and $\beta_j, j = 1, \ldots, P$, are constants.

There are many multivariate extensions of the univariate GARCH model. In this article we consider three extensions, including BEKK-GARCH, CCC-GARCH, and DCC-GARCH. They all share the same submodel for the conditional mean vector of log returns:

$$y_t = c + \sum_{i=1}^{R} \Phi_i y_{t-i} - \sum_{j=1}^{M} \Theta_j a_{t-j} + a_t,$$

\[1\] By natural filtration we mean the information up to and including time $t$. 

where $c$ is a vector of $N$ constants, $a_i$ is a vector of $N$ innovations, $\Phi_i$, $i = 1, \ldots, R$, and $\Theta_j$, $j = 1, \ldots, M$, are $N \times N$ matrices that determine the dependence of $y_t$ on previous log returns and innovations, respectively. It is assumed that

$$a_t | F_{t-1} \sim \text{MVN}(0, H_t).$$

The difference between the three multivariate GARCH models lies in the way in which the conditional covariance matrix $H_t$ is specified. There are different specifications of $H_t$, because specifying $H_t$ is a tradeoff between flexibility and parsimony. On one hand, the submodel should be flexible enough to realistically represent the dynamics of the conditional variances and covariances. On the other hand, it should be parsimonious enough so that accurate estimation of model parameters is feasible in practice.

### 2.2.1. BEKK-GARCH Models

Engle and Kroner (1995) considered the Baba-Engle-Kraft-Kroner (BEKK-) GARCH model, which is defined as follows:

$$H_t = \mathbf{K} \mathbf{K}^\prime + \sum_{i=1}^{Q} \sum_{k=1}^{K} \mathbf{A}_{k,i} a_{t-i} a_{t-i}^\prime \mathbf{A}_{k,i}^\prime + \sum_{j=1}^{P} \sum_{k=1}^{K} \mathbf{B}_{k,j} H_{t-j} \mathbf{B}_{k,j}^\prime,$$

where $\mathbf{A}_{k,i}$, $\mathbf{B}_{k,j}$, and $\mathbf{K}$ are $N \times N$ matrices and, in particular, $\mathbf{K}$ is a lower triangular matrix.

Although the BEKK-GARCH model automatically ensures positive definiteness of $H_t$, estimation may be computationally demanding because its submodel for $H_t$ contains a large number of parameters. There are several ways to simplify the BEKK-GARCH model (see, e.g., Kroner and Ng 1998).

### 2.2.2. CCC-GARCH Models

One way to simplify the submodel for $H_t$ is to decompose the conditional covariance matrix into conditional standard deviations and correlations. Bollerslev (1990) introduced the Constant Conditional Correlation (CCC-) GARCH model, in which conditional standard deviations are time varying but the conditional correlations are not. Mathematically, the model can be expressed as

$$H_t = \mathbf{D}_t \mathbf{P} \mathbf{D}_t,$$

where $\mathbf{D}_t = \text{diag}(\sqrt{h_{1,t}}, \ldots, \sqrt{h_{N,t}})$; $h_{i,t}$, $i = 1, \ldots, N$, is the conditional variance of the log total return on the $i$th equity index from time $t-1$ to time $t$; and $\mathbf{P} = [\rho_{ij}]$ is the conditional correlation matrix with $\rho_{ii} = 1$ for $i = 1, \ldots, N$.

Let $h_t = \{h_{1,t}, \ldots, h_{N,t}\}$. The stochastic nature of $h_t$ is modeled by

$$h_t = k + \sum_{i=1}^{Q} \mathbf{A}_i a_{t-i} \circ a_{t-i} + \sum_{j=1}^{P} \mathbf{B}_j h_{t-j},$$

where $k$ is a vector of $N$ constants, $\mathbf{A}_i$ and $\mathbf{B}_j$ are diagonal $N \times N$ parameter matrices, and $\circ$ denotes the Hadamard (elementwise) product operator. The CCC-GARCH model implies positive definiteness of $H_t$ if $\mathbf{P}$ is positive definite and the elements of $k$ and the diagonal elements of $\mathbf{A}_i$ and $\mathbf{B}_j$ are positive.

### 2.2.3. DCC-GARCH Models

The CCC-GARCH model is relatively parsimonious. However, its simple specification may not be able to capture the relevant dynamics in the correlation structure. There are other sophisticated conditional correlation models, such as the extended CCC-GARCH model (Jeanntheau 1998), the Varying Correlation (VC-) GARCH model (Tse and Tsui 1999), and the Dynamic Conditional Correlation (DCC-) GARCH model (Engle 2002). In what follows, we discuss the DCC-GARCH model.

The DCC-GARCH model has the same equation for the vector of conditional variances $h_t$. However, the conditional covariance matrix is modeled by $H_t = \mathbf{D}_t \mathbf{P} \mathbf{D}_t$, where the conditional correlation matrix $\mathbf{P}_t$ is modeled by

$$[\mathbf{P}_t]_{ij} = [\mathbf{Q}_t]_{ij} / \sqrt{[\mathbf{Q}_t]_{ii} [\mathbf{Q}_t]_{jj}},$$

where $\mathbf{Q}_t$ is a vector of $N$ innovations, $\Phi_i$, $i = 1, \ldots, R$, and $\Theta_j$, $j = 1, \ldots, M$, are $N \times N$ matrices that determine the dependence of $y_t$ on previous log returns and innovations, respectively. It is assumed that

$$a_t | F_{t-1} \sim \text{MVN}(0, H_t).$$

It would be numerically very difficult to ensure that $H_t$ is a covariance matrix if its structure is specified arbitrarily. It is typical to use an eigenvalue–eigenvector decomposition to ensure the positive definiteness of a matrix. As the dimension of the matrix increases, such a procedure becomes increasingly time-consuming and unstable.
with

\[
Q_t = \left( 1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j \right) S + \sum_{i=1}^{q} \alpha_i a_{t-i} a'_{t-i} + \sum_{j=1}^{p} \beta_j Q_{t-j}.
\]

In the above, \( S \) is the unconditional correlation matrix of \( a_t \), and \( \alpha_i, i = 1, \ldots, p \), and \( \beta_j, j = 1, \ldots, q \), are constant scalars.

### 2.3. Regime-Switching Lognormal Models

The class of univariate regime-switching lognormal models was proposed by Hamilton (1989) and then introduced to the actuarial profession for modeling long-term equity returns by Hardy (2001).

Define by \( J_t = 1, \ldots, K \) the regime between time \( t-1 \) to time \( t \). In a regime-switching model, the switching between the \( K \) regimes is driven by a Markov process. Specifically, the probability of switching from regime \( i \) to regime \( j \) in a single time step is given by

\[
P[J_{t+1} = j | J_t = i] = p_{i,j}.
\]

Let \( F_t^J \) be the natural filtration generated by the discrete-time Markov chain \( J_t \). In a univariate setting, a regime-switching lognormal model with \( K \) regimes can be expressed as

\[
y_t = \mu_{J_t} + a_t,
\]

\[
a_t | F_t^J \sim N(0, h_{J_t}),
\]

where \( a_t \) is the innovation at time \( t \), and \( \mu_{J_t} \) and \( h_{J_t} \) are the mean and variance of the log total return given knowledge about regime \( J_t \), respectively.

Hardy (2001) found that a model with \( K = 2 \) regimes gives an adequate fit to the TSE 300 index from 1956 to 1999. The two-regime model (RSLN-2) captures not only stochastic volatility but also the potential association between poor returns and high volatility.

It is rather straightforward to derive a multivariate regime-switching lognormal model. Assuming that a change in regime affects all \( N \) equity indices, we have the following multivariate regime-switching lognormal model:

\[
y_t = \mu_{J_t} + a_t,
\]

\[
a_t | F_t^J \sim MVN(0, H_{J_t}),
\]

where \( \mu_{J_t} \) and \( H_{J_t} \) are the conditional mean vector and the covariance matrix for \( y_t \) given knowledge about regime \( J_t \), respectively. Same as the univariate version, a Markov process is used to model the switching between regimes. The transition probabilities are specified by equation (1).

The above is not the only possible multivariate extension. As Boudreault and Panneton (2009) point out, one may use the same correlation matrix for all regimes to reduce the number of parameters to estimate. One may also add regimes in the Markov process to permit transitions that do not affect all equity indices being modeled.

### 2.4. Numerical Examples

We now apply the models to real financial data. Here we consider monthly total returns from S&P600 and S&P500 indices over the period of October 1994 to December 2011. The former index is linked to small-cap U.S. common stocks, while the latter is linked to large-cap U.S. common stocks. We use data since October 1994, because S&P600 was first published in that month. In what follows, we denote the log total return series for S&P600 and S&P500 by \( \{y_{1,t}\} \) and \( \{y_{2,t}\} \), respectively. All models are estimated by maximum likelihood.

#### 2.4.1. Bivariate Lognormal

The estimated model is

\[
\begin{align*}
y_{1,t} & = 0.0071184 + a_{1,t} \\
y_{2,t} & = 0.0047536 + a_{2,t}.
\end{align*}
\]
where \((a_{1,t}, a_{2,t})'\) follows a bivariate normal distribution with zero means, standard deviations of 0.057569 and 0.046467, and a correlation of 0.81839.

2.4.2. GARCH Models

We consider the simplest mean structure (i.e., \(R = M = 0\)). The estimated model for the conditional mean is

\[
\begin{align*}
    y_{1,t} &= 0.0071184 + a_{1,t} \\
    y_{2,t} &= 0.0047536 + a_{2,t}
\end{align*}
\]

The above applies to all three GARCH models we consider.

The covariance structure of \(a_t\) for the three GARCH models is as follows:

- **BEKK-GARCH model** with \(P = Q = K = 1\). The estimated model is
  \[
  H_t = \begin{bmatrix} h_{1,t} & 0.80103 \sqrt{h_{1,t} h_{2,t}} \\
  0.80103 \sqrt{h_{1,t} h_{2,t}} & h_{2,t} \end{bmatrix},
  \]
  where

  \[
  C = \begin{bmatrix} 0.017041 & 0 \\
  0.011808 & 0.001882 \end{bmatrix}, \quad A = \begin{bmatrix} 0.38106 & -0.0040462 \\
  -0.0031425 & 0.39337 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 0.90178 & -0.012811 \\
  -0.021157 & 0.91377 \end{bmatrix}.
  \]

- **CCC-GARCH model** with \(P = Q = 1\). The estimated model is
  \[
  H_t = \begin{bmatrix} h_{1,t} & \rho_t \sqrt{h_{1,t} h_{2,t}} \\
  \rho_t \sqrt{h_{1,t} h_{2,t}} & h_{2,t} \end{bmatrix},
  \]
  where

  \[
  \begin{align*}
  h_{1,t} &= 0.00038685 + 0.10509a_{1,t-1}^2 + 0.78663h_{1,t-1} \\
  h_{2,t} &= 0.000065917 + 0.19153a_{2,t-1}^2 + 0.80329h_{2,t-1}.
  \end{align*}
  \]

- **DCC-GARCH model** with \(P = Q = p = q = 1\). The estimated model is
  \[
  H_t = \begin{bmatrix} h_{1,t} & \rho_t \sqrt{h_{1,t} h_{2,t}} \\
  \rho_t \sqrt{h_{1,t} h_{2,t}} & h_{2,t} \end{bmatrix},
  \]
  where \(\rho_t = [Q_t]_{12}/\sqrt{[Q_t]_{11}[Q_t]_{12}}\); the process for \(Q_t\) is given by

  \[
  Q_t = \begin{bmatrix} 0.000002 & 0.00000164 \\
  0.00000164 & 0.000002 \end{bmatrix} + 0.10584a_{t-1} \otimes a_{t-1} + 0.89416Q_{t-1}
  \]

  and the process for \(h_t\) is the same as that for the estimated CCC(1,1)-GARCH model.

2.4.3. Regime-Switching Lognormal Model

We suppose that there are two regimes and that the correlation between the innovation terms \(a_{1,t}\) and \(a_{2,t}\) is the same under both regimes in order to reduce the number of parameters.
In the estimated model, the probability of switching from regime 1 to regime 2 is 0.035248, while the probability of switching from regime 2 to regime 1 is 0.029042. In regime 1, the model is
\[
\begin{align*}
    y_{1,t} &= 0.014338 + a_{1,t} \\
    y_{2,t} &= 0.011437 + a_{2,t}
\end{align*}
\]
where \((a_{1,t}, a_{2,t})\)' follows a bivariate normal distribution with zero means, standard deviations of 0.036578 and 0.024391, and a correlation of 0.81150. In regime 2, the model is
\[
\begin{align*}
    y_{1,t} &= 0.0017346 + a_{1,t} \\
    y_{2,t} &= -0.00023164 + a_{2,t}
\end{align*}
\]
where \((a_{1,t}, a_{2,t})\)' follows a bivariate normal distribution with zero means, standard deviations of 0.067888 and 0.056422, and a correlation of 0.81150.

3. IDENTIFYING A RISK-NEUTRAL PROBABILITY MEASURE IN A MULTIVARIATE FRAMEWORK

In this section we transform the models introduced in Section 2 into their risk-neutral counterparts by using the conditional Esscher transform. We start with lognormal models without regime switching. Since the market is complete under these models, the conditional Esscher transform would result in a unique risk-neutral measure. We then proceed to multivariate GARCH models and regime-switching lognormal models. Since the market is incomplete under these two classes of models, the conditional Esscher transform would only result in one of the many possible risk-neutral measures. A justification of the use of the Esscher transform is provided in the Appendix.

3.1. The Esscher Transform

In this section we present the idea of Esscher transform in the simplest one-period securities market model: a market with a riskless asset and a single risky asset that follows a lognormal distribution. The constant continuously compounded risk-free interest rate is \(r\) per unit of time. The time-1 price of the risky asset is
\[
S_1 = S_0 \exp\left(\alpha - \delta - \frac{\sigma^2}{2} + \sigma a_1\right),
\]
where \(\alpha\) is the mean total return of the asset, \(\delta\) is the dividend yield, \(\sigma\) is the volatility, and \(a_1\) is a standard normal random variable under the real-world probability measure \(P\). It is easy to see that under the real-world measure, the mean of \(S_1\) is
\[
E[S_1] = S_0 e^{\alpha - \delta}.
\]
The log return, which can be expressed as
\[
y_1 = \ln \frac{S_1}{S_0} = \alpha - \delta - \frac{\sigma^2}{2} + \sigma a_1,
\]
is normally distributed with mean \(\alpha - \delta - \frac{\sigma^2}{2}\) and variance \(\sigma^2\).

The idea of the Esscher transform is to find a probability measure under which the risky asset earns a mean total return of \(r\) per period, so that arbitrage opportunities in the market are precluded. A mean total return of \(r\) is equivalent to \(e^r S_1\) having a mean of \(S_0 e^r\). This is because an investor can reinvest all dividends received by purchasing extra shares, implying that one share at time 0 grows to \(e^r\) shares at time 1. The resulting probability measure, which we call the risk-neutral Esscher measure, must be equivalent to the real-world probability measure. This loosely means that an event that is probable in the risk-neutral Esscher measure is also probable in the real-world measure and vice versa.

To introduce the Esscher transform, we let \(\lambda\) (the Esscher parameter) be a constant and define a random variable \(Z_1\) as follows:
\[
Z_1 = \frac{e^{\lambda y_1}}{E[e^{\lambda y_1}]}. \tag{2}
\]
By using the moment generating function for a normal random variable, we see that
\[ Z_1 = \frac{e^{\lambda y_1}}{e^{(\alpha - \delta - \sigma^2/2)\lambda + \frac{1}{2} \lambda^2 \sigma^2}}. \]
Since \( Z_1 \) is a positive random variable and \( \mathbb{E}(Z_1) = 1 \), we can use \( Z_1 \) as a Radon-Nikodym derivative to generate a probability measure \( \mathbb{P}^{\lambda} \):
\[ \mathbb{P}^{\lambda}(A) = \int_A Z_1 d\mathbb{P} \]
for any measurable event \( A \).

Now we consider the behavior of the risky asset under this derived measure. Let \( v \) be a fixed constant and consider the moment generating function of \( y_1 \):
\[ \mathbb{E}^{\lambda}[e^{v y_1}] = \mathbb{E}[e^{v y_1} Z_1] = \frac{\mathbb{E}[e^{(\lambda+1) y_1}]}{e^{(\alpha - \delta - \sigma^2/2)\lambda + \frac{1}{2} \lambda^2 \sigma^2}} = e^{(\alpha - \delta - \sigma^2/2 + \lambda \sigma^2 + \frac{1}{2} \nu^2 \sigma^2)}. \]
In particular, by putting \( v = 1 \), we obtain
\[ \mathbb{E}^{\lambda}[S_1] = S_0 e^{\alpha - \delta + \lambda \sigma^2}. \]
As a result, by setting \( \lambda \) to
\[ \lambda = \frac{r - \alpha}{\sigma^2}, \]
we have \( \mathbb{E}^{\lambda}[S_1] = S_0 e^{-\delta} \), which means that the risky asset earns a mean total return of \( r \) per period. The measure induced by \( \lambda \) is the risk-neutral Esscher measure, which we denote by \( \tilde{\mathbb{P}} \). It can be shown that under the risk-neutral Esscher measure,
\[ S_1 = S_0 \exp(r - \delta - \sigma^2/2 + \sigma \tilde{a}_t), \]
where \( \tilde{a}_1 \) is a standard normal random variable under \( \tilde{\mathbb{P}} \).

3.2. Lognormal Models

In this case the market model is complete, and there is a unique risk-neutral measure, which is the same as the risk-neutral Esscher measure. Under the risk-neutral measure,
\[ y_t = \mu_t + a_t, \quad a_t \sim \text{MVN}(0, \mathbf{H}), \]
where \( r \) is the continuously compounded risk-free interest rate per period, \( \mathbf{1} \) is a column vector of ones, \( \boldsymbol{\delta} \) is the vector of the dividend yields of the indices per period, and \( \mathbf{h} \) is the column vector of the variances of the innovations; that is, \( h_i = [\mathbf{H}]_{i,i} \).

3.3. GARCH Models

Recall that if \( y_t \) follows a multivariate GARCH model, the dynamics of \( y_t \) is given by
\[ y_t = \mu_t + a_t, \]
We refer readers to Appendix A of Panjer et al. (1998) for definitions of a Radon-Nikodym derivative and a measurable event.
where $\mu_i$ is $\mathcal{F}_{t-1}$-measurable and $\mathbf{a}_i|\mathcal{F}_{t-1}$ is multivariate normal with zero mean and $\mathcal{F}_{t-1}$-measurable covariance matrix $\mathbf{H}_t$ under the real-world measure $\mathbb{P}$; that is,

$$
y_{i|\mathcal{F}_{t-1}} \sim \text{MVN}(\mu_i, \mathbf{H}_t).
$$

(3)

We let $\delta_{i,t}$ be the continuous dividend yield on the $i$th asset from time $t-1$ to time $t$ and let $\delta_t$ be the column vector formed by $\delta_{i,t}$.

For all three multivariate GARCH models we consider, the procedure for identifying the risk-neutral Esscher measure is the same. We now present the general derivation.

For $t = 1, 2, \ldots, T$, let $\lambda_t = (\lambda_{1,t}, \lambda_{2,t}, \ldots, \lambda_{k,t})'$. We assume that $\{\lambda_t\}$ is $\mathcal{F}$-predictable and define

$$
Z_t = \prod_{j=1}^{t} \frac{e^{\lambda_j y_j}}{\mathbb{E}[e^{\lambda_j y_j}|\mathcal{F}_{j-1}]}.
$$

It is not difficult to prove that $\{Z_t\}$ is a $(\mathbb{P}, \mathcal{F})$ martingale. We can then define a new probability measure $\tilde{\mathbb{P}}_t$ that is locally absolutely continuous with respect to $\mathbb{P}$ by using $Z_t$:

$$
\tilde{\mathbb{P}}_t(A) = \int_A Z_t d\mathbb{P}
$$

for any $A \in \mathcal{F}_t$. Since $\{\tilde{\mathbb{P}}_t\}$ is consistent, there is a probability measure $\tilde{\mathbb{P}}$ so that $\tilde{\mathbb{P}}_t(A) = \tilde{\mathbb{P}}(A)$ for $A \in \mathcal{F}_t$. We denote the expectation under $\tilde{\mathbb{P}}$ by $\tilde{\mathbb{E}}$.

We then obtain the distribution of $y_{i|\mathcal{F}_{t-1}}$ under the new measure $\tilde{\mathbb{P}}$ by deriving its conditional moment generating function. Let $\mathbf{v}_t = (v_{1,t}, v_{2,t}, \ldots, v_{k,t})'$. Then

$$
\tilde{\mathbb{E}}(e^{\mathbf{v}' y_{i|\mathcal{F}_{t-1}}})
= \frac{\mathbb{E}[e^{(\lambda_i y_i + \mathbf{v})_{i|\mathcal{F}_{t-1}}}]}{\mathbb{E}[e^{\lambda_i y_i}|\mathcal{F}_{t-1}]}
= e^{(\lambda_i y_i + \mathbf{v})_{i|\mathcal{F}_{t-1}}/\mathbb{E}[y_i|\mathcal{F}_{t-1}]}
= e^{(\mu_i + \mathbf{H}_t \lambda_i + \mathbf{v})/\mathbb{E}[y_i|\mathcal{F}_{t-1}]}
$$

which implies that

$$
y_{i|\mathcal{F}_{t-1}} \sim \text{MVN}(\mu_i + \mathbf{H}_t \lambda_i, \mathbf{H}_t).
$$

(4)

Comparing equations (3) and (4), we see that the conditional mean of $y_i$ changes from $\mu_i$ to $\mu_i + \mathbf{H}_t \lambda_i$, while the conditional covariance matrix is unchanged.\(^4\)

Now we find the values of $\lambda_t$ such that $\tilde{\mathbb{P}}$ is a risk-neutral Esscher measure. Suppose that the continuously compounded risk-free interest rate per period is $r$. By substituting $\mathbf{v}_t$ with $\hat{\mathbf{v}}_t$, the $i$th unit column vector, we have

$$
\mathbb{E}[e^{\mathbf{v}_t y_{i|\mathcal{F}_{t-1}}}] = e^{\hat{\mathbf{v}}_t \cdot (\mu_i + \mathbf{H}_t \lambda_i + \mathbf{H}^\dagger [r - \delta_t - \mu_t - \frac{1}{2} \mathbf{h}_t])}
$$

But since each of the $N$ assets must earn a risk-free interest rate under the risk-neutral measure, we obtain

$$
\delta_{i,t} + \hat{\mathbf{v}}_t \cdot (\mu_i + \mathbf{H}_t \lambda_i) + \frac{1}{2} \mathbf{h}_t \cdot \lambda_t = r, \quad i = 1, 2, \ldots, N.
$$

The $N$ equations above form a linear simultaneous system of equations to which the solution is

$$
\lambda_t = \mathbf{H}_t^{-1} \left( r 1 - \delta_t - \mu_t - \frac{1}{2} \mathbf{h}_t \right).
$$

\(^4\)Siu et al. (2004) used similar arguments in deriving a univariate risk-neutral GARCH model.
where \( h_t \) is the column vector formed by the main diagonal of \( H_t \). As a result, under the risk-neutral Esscher measure, the conditional distribution of \( a_t \) given \( \mathcal{F}_{t-1} \) is multivariate normal with conditional mean \( r I - \delta_t - \frac{1}{2} h_t \) and conditional covariance matrix \( H_t \).

Summing up, the dynamics of \( y_t \) under the risk-neutral Esscher measure is given by

\[
y_t = r I - \delta_t - \frac{1}{2} h_t + \tilde{a}_t,
\]

where \( \tilde{a}_t|\mathcal{F}_{t-1} \sim \text{MVN}(0, H_t) \).

By specifying \( H_t \) accordingly, the result above can be applied to each of the three multivariate GARCH models we consider.

### 3.4. Regime-Switching Lognormal Models

Recall that if \( y_t \) follows a multivariate regime-switching lognormal model, the dynamics of \( y_t \) is given by

\[
y_t = \mu_j t + a_t,
\]

\[
a_t|\mathcal{F}_t \sim \text{MVN}(0, H_j),
\]

where \( a_t \) is the innovation vector at time \( t \), and \( \mu_j \) and \( H_j \) are the conditional mean and covariance matrices of the volatility of the log total return in regime \( J_t \), respectively. If we let \( \alpha_i,j \) and \( \delta_i,j \) be the mean total return and the dividend yield of the \( i \)th risky asset in the \( j \)th regime, then

\[
\mu_j = \begin{bmatrix}
\alpha_{1,j} \\
\alpha_{2,j} \\
\vdots \\
\alpha_{N,j}
\end{bmatrix} - \begin{bmatrix}
\delta_{1,j} \\
\delta_{2,j} \\
\vdots \\
\delta_{N,j}
\end{bmatrix} - \frac{1}{2} \begin{bmatrix}
\bar{h}_{1,j} \\
\bar{h}_{2,j} \\
\vdots \\
\bar{h}_{N,j}
\end{bmatrix}.
\]

For simplicity, we write the above as \( \mu_j = \alpha_j - \delta_j - \frac{1}{2} h_j \). Conditioning on the regime, \( a_t \) are independent random vectors. Such a sequence of normal random vectors can be constructed by using a sequence of standard normal random vectors \( B \sim \text{MVN}(0, I_{N \times N}) \) and setting \( a_t = \text{Chol}(H_j)B_t \), where \( \text{Chol} \) is the Choleski decomposition of a matrix. The natural filtration generated by \( B_t \) is denoted by \( \mathcal{F}_t^B \). We also denote the filtration \( \mathcal{F}_t^J \vee \mathcal{F}_t^B \) by \( \mathcal{F}_t \). Under the real-world measure \( \mathbb{P} \),

\[
y_t|\mathcal{F}_t^J \sim \text{MVN}(\mu_j, H_j).
\]

Let \( \lambda_j = (\lambda_{1,j}, \lambda_{2,j}, \ldots, \lambda_{k,j})' \) be a sequence of Esscher parameter vectors. Generalizing the work of Elliott et al. (2005), the regime-switching conditional Esscher transform is obtained by using the Radon-Nikodym derivative

\[
Z_t = \prod_{j=1}^{t} e^{\lambda_j y_j} \mathbb{E} \left[ e^{\lambda_j y_j} | \mathcal{F}_j^J \right].
\]

It is not difficult to verify that \( \{Z_t\} \) is a \( (\mathbb{P}, \mathcal{F}) \)-martingale. As for the multivariate GARCH models, there exists a probability measure \( \widehat{\mathbb{P}} \) such that \( \widehat{\mathbb{E}}(X) = \mathbb{E}(XZ_t) \) for \( X \in \mathcal{F}_t \).

We now consider the distribution of \( y_t|J_t \) under \( \widehat{\mathbb{P}} \). Let \( \nu_t = (v_{1,t}, v_{2,t}, \ldots, v_{N,t})' \).

As before, it is easy to show that the joint moment generating function of \( y_t \) under \( \widehat{\mathbb{P}} \), conditioning on \( \mathcal{F}_t^J \), is

\[
\mathbb{E}(e^{\nu_t | \mathcal{F}_t^J}) = e^{\nu_t' (\mu_j + H_j \lambda_t) + \nu_t' H_t \nu_t},
\]

which means

\[
y_t|\mathcal{F}_t^J \sim N(\mu_j + H_j \lambda_t, H_j)
\]
under $\hat{\mathbb{P}}$. As a result, conditioning on the regime, the risk-neutral Esscher measure is obtained by using
\[
\lambda_t = H_{jt}^{-1} \left( r_1 - \delta_{jt} - \mu_{jt} - \frac{1}{2} h_{jt} \right),
\]
where $r$ is the continuously compounded risk-free interest rate per period. Under the risk-neutral Esscher measure, we have
\[
y_{t|F_j} \sim \text{MVN} \left( r_1 - \delta_{jt} - \frac{1}{2} h_{jt}, H_{jt} \right).
\]

For simplicity, it is assumed in this article that the risk-free interest rates in all regimes are equal. Nevertheless, one may easily generalize the setup to incorporate regime-dependent risk-free interest rates by replacing $r$ with $r_{jt}$, where $r_j$ is the continuously compounded risk-free interest rate per period in the $j$th regime.

3.5. Numerical Examples

In this section we present the process for $y_t$ under the risk-neutral Esscher measure for all models presented in Section 2.4, assuming a continuously compounded risk-free interest of 2% per year ($r = 1/6\%$ per month).

3.5.1. Bivariate Lognormal

The model is
\[
\begin{align*}
y_{1,t} &= 0.0000095718 + \tilde{a}_{1,t} \\
y_{2,t} &= 0.00058708 + \tilde{a}_{2,t},
\end{align*}
\]
where $(\tilde{a}_{1,t}, \tilde{a}_{2,t})'$ follows a bivariate normal distribution with zero means, standard deviations of 0.057569 and 0.046467, and a correlation of 0.81839.

3.5.2. GARCH Models

- BEKK-GARCH model with $P = Q = K = 1$. The model is
\[
\begin{align*}
y_{1,t} &= 0.0016667 - 0.5[H_{jt}]_{11} + \tilde{a}_{1,t} \\
y_{2,t} &= 0.0016667 - 0.5[H_{jt}]_{22} + \tilde{a}_{2,t},
\end{align*}
\]
with $\tilde{a}_t \sim \text{MVN}(0, H_t)$, where
\[
H_t = CC' + A\tilde{a}_{t-1}A' + BH_{t-1}B'.
\]
The values of $A$, $B$, and $C$ are given in Section 2.4.

- CCC-GARCH model with $P = Q = 1$. The model is
\[
\begin{align*}
y_{1,t} &= 0.0014732 - 0.052545\tilde{a}_{1,t-1} - 0.39332h_{1,t-1} + \tilde{a}_{1,t} \\
y_{2,t} &= 0.0016337 - 0.095765\tilde{a}_{2,t-1} - 0.40165h_{2,t-1} + \tilde{a}_{2,t},
\end{align*}
\]
with $\tilde{a}_t \sim \text{MVN}(0, H_t)$, where
\[
H_t = \begin{bmatrix}
h_{1,t} & 0.80103\sqrt{h_{1,t}h_{2,t}} \\
0.80103\sqrt{h_{1,t}h_{2,t}} & h_{2,t}
\end{bmatrix},
\]
and
\[
\begin{align*}
h_{1,t} &= 0.00038685 + 0.10509\tilde{a}_{1,t-1} + 0.78663h_{1,t-1} \\
h_{2,t} &= 0.000065917 + 0.19153\tilde{a}_{2,t-1} + 0.80329h_{2,t-1}.
\end{align*}
\]
**DCC-GARCH model with** $P = Q = p = q = 1$. The model is

$$H_t = \begin{bmatrix}
    h_{1,t} & \rho_t \sqrt{h_{1,t} h_{2,t}} \\
    \rho_t \sqrt{h_{1,t} h_{2,t}} & h_{2,t}
\end{bmatrix},$$

where $\rho_t = \frac{Q_{12}}{\sqrt{Q_{11} Q_{22}}}$; the process for $Q_t$ is given by

$$Q_t = \begin{bmatrix}
    0.000002 & 0.00000164 \\
    0.00000164 & 0.000002
\end{bmatrix} + 0.10584 \tilde{a}_{t-1} \odot \tilde{a}_{t-1} + 0.89416 Q_{t-1}$$

and the process for $h_t$ is exactly the same as that for the risk-neutral CCC(1,1)-GARCH model.

### 3.5.3. Regime-Switching Lognormal Models

Under the risk-neutral Esscher measure, the probability of switching from regime 1 to regime 2 is 0.035248, while the probability of switching from regime 2 to regime 1 is 0.029042. In regime 1, the model is

$$\begin{cases}
    y_{1,t} = 0.00099769 + \tilde{a}_{1,t} \\
    y_{2,t} = 0.0013692 + \tilde{a}_{2,t}
\end{cases},$$

where $(\tilde{a}_{1,t}, \tilde{a}_{2,t})'$ follows a bivariate normal distribution with zero means, standard deviations of 0.036578 and 0.024391, and a correlation of 0.81150. In regime 2, the model is

$$\begin{cases}
    y_{1,t} = -0.00063772 + \tilde{a}_{1,t} \\
    y_{2,t} = 0.000074946 + \tilde{a}_{2,t}
\end{cases},$$

where $(\tilde{a}_{1,t}, \tilde{a}_{2,t})'$ follows a bivariate normal distribution with zero means, standard deviations of 0.067888 and 0.056422, and a correlation of 0.81150.

### 4. PRICING FORMULAS

Having identified a risk-neutral probability measure, we can calculate prices of various investment guarantees by using Monte Carlo simulations. We illustrate our pricing methodology with a guaranteed minimum maturity benefit (GMMB) and a guaranteed minimum death benefit (GMDB).

A variable annuity has two phases: an accumulation phase and a payout phase. During the accumulation phase, the policyholder makes premium payments, which he or she can allocate to a number of investment options. At the beginning of the payout phase, the policyholder may receive his or her purchase payments plus investment income and gains (if any) as a lump-sum payment, or he or she may choose to receive them as a stream of payments at regular intervals. A GMMB guarantees the policyholder a specific monetary amount at the end of the accumulation phase (i.e., at maturity). We denote the maturity (in months) by $T$.

Let us suppose that at time 0, a policyholder makes a single premium payment $P$, which is invested in $N$ different subaccount funds. In particular, the policyholder’s portfolio contains $w_i$ units of the $i$th subaccount fund, which has a time-$t$ value of $S_{i,t}$ per unit. Note that $P = \sum_{i=1}^{N} w_i S_{i,0}$. For simplicity, we assume that there is no dividend income and that $w_i$ are constant over time.

Denote by $m$ the management charge rate deducted from the policyholder’s account per month. Upon maturity, the value of the policyholder’s portfolio grows to $F_T = (\sum_{i=1}^{N} w_i S_{i,t})(1 - m)^T$ after the deduction of monthly management charges. If the policyholder still survives and has not withdrawn, the payoff from the GMMB is $(P g - F_T)^+$, where $g$ is the guarantee level, expressed as a percentage of the initial premium. The time-0 price of this payoff is

$$e^{-rT} \mathbb{E}[(P g - F_T)^+ | \mathcal{F}_0],$$

where $r$ is the continuously compounded risk-free interest rate per month, and $\mathbb{E}$ is the expectation under the risk-neutral Esscher measure $\tilde{P}$. We assume that financial and demographic risks are independent. Taking into account the effects of death and
withdrawal, the value of the GMMB is

\[ V(0) = r p^{(x)}_s e^{-rT} \mathbb{E}[(P g - F g)|F_0], \]  

(5)

where \( p^{(x)}_s \) is the probability that a policyholder aged \( x \) (in years) at time 0 survives and does not withdraw after \( t \) months.

Then we consider the valuation of a GMDB, which guarantees a specific monetary sum upon death during the accumulation phase of a variable annuity contract. That is, if the policyholder dies, a person he or she selects as a beneficiary (such as his or her spouse or child) will receive the greater of (1) all the money in the policyholder’s account or (2) some guaranteed minimum.

For simplicity we assume that the GMDB liability is paid at the end of the month of death if death happens before maturity. Let \( q^{(d)}_x(t) \) be the probability that a policyholder who is currently aged \( x \) dies during the time interval \( t \) to \( t + 1 \), given that the policy is still in force at time \( t \). Then the value of the GMDB with a maturity of \( T \) months is given by

\[ V(0) = \sum_{t=1}^{T} q^{(d)}_x(t) e^{-rt} \mathbb{E}[(P g - F g)|F_0]. \]  

(6)

We can evaluate \( \mathbb{E}[(P g - F g)|F_0] \) in equations (5) and (6) numerically with the following procedure:

1. Simulate 10,000 sample paths of \( F_t \) from a multivariate stochastic investment model under the risk-neutral Esscher measure\(^5\)
2. For each sample path, calculate the value of \( (P g - F g)^+ \) for \( t = 1, 2, \ldots, T \)
3. Average the simulated values of \( (P g - F g)^+ \) to obtain \( \mathbb{E}[(P g - F g)|F_0] \) for \( t = 1, 2, \ldots, T \).

We now apply the multivariate pricing method to the hypothetical GMDB and GMMB liabilities. In our illustrations, it is assumed that the policyholder’s portfolio contains two subaccount funds, both of which are nondividend paying. The returns on these two subaccount funds follow the multivariate processes estimated in Section 2 (in the real-world probability measure) and in Section 3 (in the risk-neutral probability measure). In addition, the following pricing assumptions are used:

| Age at inception: | 45, 50, and 60 |
| Premium: | A single premium of $10,000, paid at inception |
| Portfolio: | Two subaccount funds, one unit each |
| Fund values at inception: | $5,000 per unit for both funds |
| Mortality: | Independent mortality rates follow the Makeham law with \( \mu(x) = A + Bx^c \), where \( \mu(x) \) is the force of mortality at age \( x \). The Makeham parameters \( A = 0.00054, B = 9.5929 \times 10^{-6} \), and \( c = 1.1085 \) are estimated from crude central death rates for Canadian females in 2007. |
| Withdrawal: | 0% (for simplicity) |
| Guarantee: | 60%, 80%, 100%, and 120% of the premium paid |
| Management charge: | A nominal rate of 3% of the account value per year, implying \( m = 0.0025 \) |
| Maturity: | 5, 10, and 20 years |

The resulting guarantee values under different assumed multivariate investment models are displayed in Tables 1 to 5. The values shown are expressed as a percentage of the initial premium. We have the following comments on the pricing results.

First, let us review the volatility structures of GARCH-type models and lognormal models. For the former, each volatility can take a wide range of values, but for the latter, each volatility can take at most only two values (depending if one or two regime is used). The results indicate that this difference may matter. In particular, we observe in this example that the three GARCH-type models yield higher guarantee values than the two lognormal models do.

Second, the three GARCH-type models yield similar guarantee values. This observation suggests that as a practical shortcut, practitioners may use a simpler GARCH-type model, which may be easier to estimate and simulate.

\(^5\)We use crude Monte Carlo in our calculations. However, variance reduction methods described in Glasserman (2003) may be used.
### Table 1
Simulated Guarantee Values Expressed as Percentage of Initial Premium

#### Age at inception: 45

<table>
<thead>
<tr>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
</tr>
<tr>
<td>60%</td>
<td>1.49</td>
</tr>
<tr>
<td>80</td>
<td>6.46</td>
</tr>
<tr>
<td>100</td>
<td>15.66</td>
</tr>
<tr>
<td>120</td>
<td>28.34</td>
</tr>
</tbody>
</table>

#### Age at inception: 50

<table>
<thead>
<tr>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
</tr>
<tr>
<td>60%</td>
<td>1.43</td>
</tr>
<tr>
<td>80</td>
<td>6.28</td>
</tr>
<tr>
<td>100</td>
<td>15.59</td>
</tr>
<tr>
<td>120</td>
<td>28.41</td>
</tr>
</tbody>
</table>

#### Age at inception: 60

<table>
<thead>
<tr>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
</tr>
<tr>
<td>60%</td>
<td>1.44</td>
</tr>
<tr>
<td>80</td>
<td>6.43</td>
</tr>
<tr>
<td>100</td>
<td>15.88</td>
</tr>
<tr>
<td>120</td>
<td>28.67</td>
</tr>
</tbody>
</table>

*Note: Underlying model: bivariate lognormal model with no regime switching.*

Third, let us compare the volatility structures of the RSLN model and the simple lognormal model. Generally speaking, the volatilities in a simple lognormal model are lower than those in the high-volatility regime of an RSLN model but higher than those in the low-volatility regime of a RSLN model. It follows that a guarantee value implied by an RSLN model may or may not be higher than that implied by the corresponding simple lognormal model. In our example, the RSLN model yields lower guarantee values.

### 5. HEDGING IN A MULTIVARIATE FRAMEWORK

#### 5.1. Greek Approximations

Hedging with Greeks can be performed in a multivariate framework. Let us consider a delta hedge for a guarantee liability with a time-$t$ value of $V(t)$. The delta of the guarantee liability with respect to the $i$th subaccount fund at time $t$ is given by

$$
\Delta_{i,t} = \frac{\partial V(t)}{\partial S_{i,t}}.
$$

In practice, $\Delta_{i,t}$ may be estimated by a first difference approximation.\(^6\)

\(^6\)A symmetric first difference approximation is used in our calculations.
### TABLE 2
Simulated Guarantee Values Expressed as Percentage of Initial Premium

<table>
<thead>
<tr>
<th>Age at inception: 45</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.83</td>
<td>5.82</td>
</tr>
<tr>
<td>80</td>
<td>8.86</td>
<td>13.49</td>
</tr>
<tr>
<td>100</td>
<td>18.74</td>
<td>23.59</td>
</tr>
<tr>
<td>120</td>
<td>31.58</td>
<td>35.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 50</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.74</td>
<td>5.74</td>
</tr>
<tr>
<td>80</td>
<td>8.69</td>
<td>13.06</td>
</tr>
<tr>
<td>100</td>
<td>18.50</td>
<td>22.96</td>
</tr>
<tr>
<td>120</td>
<td>31.23</td>
<td>34.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 60</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.74</td>
<td>5.58</td>
</tr>
<tr>
<td>80</td>
<td>8.62</td>
<td>12.64</td>
</tr>
<tr>
<td>100</td>
<td>18.15</td>
<td>22.06</td>
</tr>
<tr>
<td>120</td>
<td>30.59</td>
<td>33.18</td>
</tr>
</tbody>
</table>

*Note: Underlying model: BEKK-GARCH model.*

To hedge the guarantee liability, at time $t$ we should hold a short position of $|\Delta_{i,t}|$ units of fund $i$ and invest

$$V(t) + \sum_{i=1}^{N} |\Delta_{i,t}|S_{i,t}$$

dollar amount of risk-free bonds.

Because the values of $\Delta_{i,t}, i = 1, \ldots, N$, change with time, the hedge portfolio has to be rebalanced periodically. Rebalancing the hedge portfolio incurs a transaction cost. Also, because the hedge portfolio may not be self-financing, additional funds may be required when the hedge portfolio is rebalanced. Transaction costs and hedging errors may give rise to additional costs on top of the value of the hedge portfolio itself.

According to Hardy (2003), the capital requirement under a dynamic hedging strategy is the capital allocated to the hedge itself (the value of the guarantee liability), plus an allowance for the additional costs that may be required to cover transaction costs and hedging error. These additional costs may be quantified by an actuarial approach. In more detail, we simulate (under the real-world probability measure) an empirical distribution of the additional costs; then the distribution’s upper 95th percentile, for example, indicates the amount of risk capital required for cushioning against the additional costs.

Hardy (2003) presented this risk management approach in a univariate framework. We now generalize it to a multivariate setup and apply it to a GMDB.
TABLE 3
Simulated Guarantee Values Expressed as Percentage of Initial Premium

<table>
<thead>
<tr>
<th>Age at inception: 45</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.64</td>
<td>5.92</td>
</tr>
<tr>
<td>80</td>
<td>8.35</td>
<td>13.38</td>
</tr>
<tr>
<td>100</td>
<td>18.02</td>
<td>23.33</td>
</tr>
<tr>
<td>120</td>
<td>30.77</td>
<td>35.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 50</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.82</td>
<td>6.16</td>
</tr>
<tr>
<td>80</td>
<td>8.60</td>
<td>13.64</td>
</tr>
<tr>
<td>100</td>
<td>18.26</td>
<td>23.55</td>
</tr>
<tr>
<td>120</td>
<td>31.11</td>
<td>35.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 60</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>60%</td>
<td>2.65</td>
<td>5.65</td>
</tr>
<tr>
<td>80</td>
<td>8.35</td>
<td>12.71</td>
</tr>
<tr>
<td>100</td>
<td>17.90</td>
<td>22.21</td>
</tr>
<tr>
<td>120</td>
<td>30.44</td>
<td>33.46</td>
</tr>
</tbody>
</table>

Note: Underlying model: CCC-GARCH model.

5.2. Hedging Errors

We let $\mathcal{H}(t)$ be the value of the hedge portfolio $t$ months from inception, given that the contract is still in force. Generalizing equation (6), we have

$$
\mathcal{H}(t) = \sum_{s=t+1}^{T} s^{-1} q_{s/(s+12)}^{(d)} e^{-rs} \mathbb{E}[ (P g - F_{s})^{+} | \mathcal{F}_{t} ].
$$

Note that $V(0) = \mathcal{H}(0)$.

Let us assume that the hedge portfolio is rebalanced monthly. Immediately before rebalancing at time $t + 1$, the hedge portfolio from $t$ will have accumulated to

$$
\mathcal{H}(t + 1^{-}) = - \sum_{i=1}^{N} \Delta_{i,t+1} |S_{i,t+1} + \bigg( \mathcal{H}(t) + \sum_{i=1}^{N} \Delta_{i,t} |S_{i,t} \bigg) e^{r}.
$$

Let $q_{x}^{(d)}$ be the probability that a policyholder who has attained age $x$ dies before age $x + 1/12$, and let $p_{x}^{(r)}$ be the probability that a policyholder aged $x$ survives and does not withdraw after 1 month. One month from inception, we require an amount of

$$
q_{x}^{(d)} (P g - F_{1})^{+} + p_{x}^{(r)} \mathcal{H}(1),
$$
TABLE 4
Simulated Guarantee Values Expressed as Percentage of Initial Premium

<table>
<thead>
<tr>
<th>Age at inception: 45</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g 5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>60%</td>
<td>2.90</td>
<td>6.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>8.87</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>18.69</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>31.61</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 50</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g 5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>60%</td>
<td>2.66</td>
<td>6.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>8.40</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>17.98</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>30.64</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 60</th>
<th>GMMB Maturity (years)</th>
<th>GMDB Maturity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g 5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>60%</td>
<td>2.85</td>
<td>5.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>8.57</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>18.24</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>30.83</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Note: Underlying model: DCC-GARCH model.

de means the hedging error at time 1 can be written as
\[
E_1 = q^{(d)}_x(Pg - F_1) + p^{(r)}_xH(1) - H(1^-).
\]

After rebalancing at time 1, we hold a hedge portfolio of \(p^{(r)}_xH(1)\). This means that the hedging error at time 2 is given by
\[
E_2 = p^{(r)}_x \left(q^{(d)}_{x+1/12}(Pg - F_2) + p^{(r)}_{x+1/12}H(2) - H(2^-)\right).
\]

Inductively, the hedging error at time \(t, t = 2, \ldots, T\), can be expressed as
\[
E_t = t^{-1}p^{(r)}_x \left(q^{(d)}_{x+(t-1)/12}(Pg - F_t) + p^{(r)}_{x+(t-1)/12}H(t) - H(t^-)\right).
\]

We are interested in the empirical distribution of the present value of the aggregate hedging error. We can generate this distribution with the following procedure:

1. Simulate a large number of sample paths of log returns under the real-world probability measure.
2. For each sample path, calculate the hedging error at time points when the hedge portfolio is rebalanced and then compute the present value of all hedging errors.

Note that this procedure requires stochastic-on-stochastic simulations, which may be computationally demanding.
TABLE 5
Simulated Guarantee Values Expressed as Percentage of Initial Premium

<table>
<thead>
<tr>
<th>Age at inception: 45</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>GMMB Maturity</td>
<td>GMDB Maturity</td>
<td>g</td>
</tr>
<tr>
<td>51</td>
<td>5 years</td>
<td>0 years</td>
<td>20 years</td>
</tr>
<tr>
<td>60%</td>
<td>1.31</td>
<td>3.73</td>
<td>6.92</td>
</tr>
<tr>
<td>80</td>
<td>5.78</td>
<td>10.12</td>
<td>13.93</td>
</tr>
<tr>
<td>100</td>
<td>14.82</td>
<td>19.66</td>
<td>22.57</td>
</tr>
<tr>
<td>120</td>
<td>27.78</td>
<td>31.57</td>
<td>32.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 50</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>GMMB Maturity</td>
<td>GMDB Maturity</td>
<td>g</td>
</tr>
<tr>
<td>51</td>
<td>5 years</td>
<td>0 years</td>
<td>20 years</td>
</tr>
<tr>
<td>60%</td>
<td>1.36</td>
<td>3.60</td>
<td>6.61</td>
</tr>
<tr>
<td>80</td>
<td>5.83</td>
<td>9.84</td>
<td>13.25</td>
</tr>
<tr>
<td>100</td>
<td>14.87</td>
<td>19.18</td>
<td>21.47</td>
</tr>
<tr>
<td>120</td>
<td>27.75</td>
<td>30.88</td>
<td>30.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at inception: 60</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>GMMB Maturity</td>
<td>GMDB Maturity</td>
<td>g</td>
</tr>
<tr>
<td>51</td>
<td>5 years</td>
<td>0 years</td>
<td>20 years</td>
</tr>
<tr>
<td>60%</td>
<td>1.29</td>
<td>3.39</td>
<td>5.34</td>
</tr>
<tr>
<td>80</td>
<td>5.70</td>
<td>9.33</td>
<td>10.78</td>
</tr>
<tr>
<td>100</td>
<td>14.65</td>
<td>18.28</td>
<td>17.50</td>
</tr>
<tr>
<td>120</td>
<td>27.38</td>
<td>29.36</td>
<td>25.11</td>
</tr>
</tbody>
</table>

Note: Underlying model: bivariate regime-switching log-normal model with two regimes.

5.3. Transaction Costs

A transaction cost is incurred whenever the hedge portfolio is rebalanced. We assume transaction costs of \( \gamma \) times the total value of the risky assets that have to be traded when the hedge portfolio is rebalanced.

Let us assume again that the hedge portfolio is rebalanced monthly. At time 1, immediately before the hedge is rebalanced, we hold a short position of \( |\Delta_{i,0}| \) units of fund \( i, i = 1, \ldots, N \). When the hedge is balanced, the required position of fund \( i \) depends on the status of the contract. If the policy is still in force, we need to hold a short position of \( |\Delta_{i,1}| \) units of fund \( i \). On the other hand, if the policyholder has died during the first month, then the required short position is \( w_i(1 - m) \) if the guarantee is in the money and zero otherwise. Therefore, the transaction cost at time 1 is given by

\[
C_1 = \gamma \sum_{i=1}^{N} S_{i,1} |p_x^{(r)}|\Delta_{i,1}| - q_x^{(d)} w_i(1 - m) I_{\{P_g - F_i > 0\}} + |\Delta_{i,0}|,
\]

where \( I_{\{A\}} \) denotes the indicator function for event \( A \).

Inductively, we have the following expression for the transaction cost at time \( t, t = 2, \ldots, T \):

\[
C_t = (t-1)p_x^{(r)} \gamma \sum_{i=1}^{N} S_{i,t} |p_x^{(r)(t-1)/12}|\Delta_{i,t}| - q_x^{(d)(t-1)/12} w_i(1 - m) I_{\{P_g - F_i > 0\}} + |\Delta_{i,t-1}|.
\]

Given the above equation, we can easily simulate a distribution of the aggregate transaction cost with a procedure similar to that provided in Section 5.2.
5.4. An Example

We illustrate the calculation of hedging errors and transaction costs with a GMDB written to a policyholder who is aged 50 at inception. The maturity of the contract is five years, and the guarantee level is $g = 1.0$. All other pricing assumptions remain the same as those in Section 4. The CCC-GARCH model is used.\(^7\) We assume that S&P500 and S&P600 indices are traded in the market.

The top panel of Figure 1 depicts the simulated distribution of the aggregate transaction cost. The simulations are based on the assumption that $\tau = 0.002$. From the diagram we observe that the average aggregate transaction cost is approximately $0.41, which is quite small relative to the value of the guarantee ($18; see Table 3).

In the middle panel of Figure 1, we show the simulated distribution of the aggregate hedging error. Note that a negative value means a hedge profit, while a positive value means a hedge loss.

---

\(^7\)The value of the hedge is calculated from the risk-neutral CCC-GARCH model, while the sample paths of the log returns are simulated from the CCC-GARCH model under the real-world probability measure.
The lower panel of Figure 1 displays the simulated distribution of the total unhedged liability (the sum of transaction costs and hedging errors). From the simulated distribution, we find that, at a 95% level of significance, the Value-at-Risk (VaR) of the total unhedged liability is $8.86. As mentioned earlier, this amount plus the guarantee’s value ($18) may be viewed as the capital requirement for the guarantee.

6. DISCUSSION AND CONCLUSION

In this article, we introduced a multivariate framework for pricing and hedging variable annuity guarantees. We believe that the use of a multivariate framework is more efficient, because a single multivariate model can accommodate different asset allocations and generate prices that are internally consistent. We demonstrated that our proposed framework can work with common multivariate stochastic investment models, including RSLN and GARCH. One can implement our proposed framework on the basis of the model that best fits to his or her data.

Unless the simplest lognormal model is assumed, there is a need to identify an appropriate risk-neutral probability measure for pricing purposes. In our work, we accomplished this important step by using the multivariate conditional Esscher transform. According to Gerber and Shiu (1994), the risk-neutral Esscher measure can be justified within an expected utility framework. The justification is presented in the Appendix of this article.

We derived in Section 3 the explicit formulas for multivariate GARCH and RSLN models under the risk-neutral Esscher measure. For both types of models, the volatility and correlation structures under the real-world and the risk-neutral Esscher measures are exactly the same. This relationship permits practitioners to straightforwardly simulate sample paths of asset returns in the risk-neutral Esscher measure and subsequently obtain Monte Carlo estimates of investment guarantee prices.

We demonstrated in Section 5 how dynamic delta hedging can be performed in a multivariate setup. Because continuous trading is impossible in practice, the hedge portfolio may not be self-financing, which means additional funds may be required when the hedge is rebalanced. We illustrated in Section 5 how the potential hedging error may be quantified by using stochastic simulations under the real-world probability measure.

For simplicity, we assumed in our illustrations that the quantity of each subaccount fund remains constant over time. In practice, however, policyholders can alter their asset allocations at different time points. The assumption of a fixed asset allocation can be relaxed if a model for dynamic policyholder behavior is available. Developing such a model is challenging, especially when there is a paucity of the relevant data. The contributions by Kim (2005) and Ulm (2010) can be a starting point.

Sometimes, policyholders may be allowed to choose subaccount funds that are linked to equity indices measured in foreign currencies. For instance, an U.S. insurer may offer (§-denominated) subaccount funds that are linked to the Nikkei 225 index (measured in yen) in Japan. When there is one or more foreign subaccount funds in a portfolio, a guarantee written on the portfolio would become a quanto (or cross-currency) derivative. To price the guarantee, we require a probability measure that is risk-neutral as seen from the perspective of a domestic investor. It would be interesting to explore in future research how such a measure can be constructed.

REFERENCES


APPENDIX
A Justification for Using the Esscher Transform

According to Gerber and Shiu (1994), the risk-neutral Esscher measure can be justified within an expected utility framework. In more detail, consider an economic agent with a power utility function

$$u(x) = \frac{x^{1-c}}{1-c},$$

where $c \neq 1$. Suppose that in the market, there is a derivative with a payoff $V$, which depends on the value of the risky asset at time $1$. The time-0 price of this derivative is denoted by $\pi_0$. We let

$$\phi(\eta) = E \left[ u \left( e^{\delta S_1} + \eta (V - e^{\delta \pi_0}) \right) \right]$$

$$= \frac{1}{1-c} E \left[ (e^{\delta S_1} + \eta (V - e^{\delta \pi_0}))^{1-c} \right],$$

where $\eta$ is a constant. It is optimal for the agent not to buy or sell any fraction or multiple of the derivative when $\eta = 0$. Hence, we require

$$\phi'(\eta) = E \left[ (e^{\delta S_1} + \eta (V - e^{\delta \pi_0}))^{-(c-1)} (V - e^{\delta \pi_0}) \right] = 0$$

when $\eta = 0$. It follows that

$$\pi_0 = e^{-\gamma} \frac{E \left( S_1^{1-c} V \right)}{E \left( S_1^{1-c} \right)} = e^{-\gamma} E(Z_1 V),$$

(7)
where

\[ Z_1 = \frac{e^{-y_1}}{\mathbb{E}(e^{-y_1})}. \] (8)

Note that equation (8) is identical to equation (2). Since equation (7) must hold for all derivative securities, we have the following condition:

\[ S_0 = e^{-r} \frac{\mathbb{E}(e^{-y_1} S_1 e^\delta)}{\mathbb{E}(e^{-y_1})} = e^{-r} \frac{\mathbb{E}(S_0 e^{(1-c) y_1 + \delta})}{\mathbb{E}(e^{-y_1})}. \] (9)

When \( y_1 \) is normally distributed with mean \( \mu \) and variance \( \sigma^2 \), condition (9) becomes

\[ c = \frac{\mu - r + \delta + \sigma^2/2}{\sigma^2} = \frac{\alpha - r}{\sigma^2}. \]

which means that the Esscher parameter is \( \lambda = -c \). Summing up, prices resulting from the Esscher transform are the same as those resulting from the equilibrium when the underlying economic agent has a power utility function.