Hedging against volatility, jumps and longevity risk in participating life insurance contracts – a Bayesian analysis

Arto Luoma

Department of Mathematics and Statistics,
Statistics Unit, FIN-33014 University of Tampere, Finland.

Tel: +358 3 3551 6428

E-mail: arto.luoma@uta.fi

and

Anne Puustelli
(presenting and corresponding author)

Department of Mathematics and Statistics,
Statistics Unit, FIN-33014 University of Tampere, Finland.

E-mail: anne.puustelli@uta.fi

This paper introduces a Bayesian approach to market consistent valuation and hedging of a participating life insurance contract. The contract is valued in a general and realistic framework allowing interest rate, volatility and jumps in the asset dynamics to be stochastic. In our set-up we also incorporate stochastic mortality and study its effect on pricing and hedging. All underlying models are estimated using the Markov Chain Monte Carlo method, and their simulation is based on their posterior predictive distribution. In our case the contract is an American-style path-dependent derivative, and we value it using the regression method. As a hedging strategy we employ minimum variance hedging which relies on the underlying asset as a single hedging instrument. We compare its hedging effectiveness with a conventional delta-neutral hedge which uses a simpler model for asset dynamics.

Key Words: American-style option, jump-diffusion, risk-neutral valuation, single-instrument hedging, stochastic interest rate, stochastic mortality, stochastic volatility

1. INTRODUCTION

Participating life insurance policies are characterized by a vast number of features including, for example, interest rate guarantees, equity-linked policies and bonus and surrender options. All these features have values which need to be priced. Pricing these policies in a market consistent framework was first studied by Briys and de Varenne (1997a,b). Since then a number of articles have appeared on the topic; see, for example, Grosen and Jorgensen (2000), Tanskanen and Lukkarinen (2003), Bernard et al. (2005), Ballotta et al. (2006), Bauer et al. (2006) and Zaglauer and Bauer (2008). However, most valuation models assume a simplified set-up. Our objective is to present a realistic valuation framework in which the guarantee and the bonus are priced in a stochastic framework, and a surrender option is included in the contract. More specifically, we allow the interest rate, volatility and jumps to be stochastic in the asset dynamics, and value the contract as an American-style path-dependent derivative. Besides this, we also incorporate stochastic mortality into the model.

The price of an option depends on the assumption of the model describing the behavior of the underlying instrument. Most approaches specify a particular stochastic process to represent the price dynamics of the underlying asset and then derive an explicit pricing formula. A traditional approach involves solving a partial differential equation. However, when the asset dynamics are assumed to follow a fairly complex model, a closed form solution of the partial differential equation may not exist or its numerical solution may become intractable. When the payoff of an option depends on the path of the underlying asset, the price cannot be evaluated in this manner. Instead, Monte Carlo simulation methods may be used (Glasserman, 2003). For example, Bacinello et al. (2008) apply the least squares Monte Carlo (or regression) method in pricing a participating life insurance with early exercise.

Most papers on pricing participating life insurance contracts lack paying attention to parameter and model errors. Neither the true underlying model nor its parameter values are known. Typically, a relatively simple model is assumed and the point estimates of the parameters are used. This might lead to a crucial valuation error. In the Bayesian approach, parameter and model uncertainty plays a major role. While frequentist methods typically rely on large sample approximations, Bayesian inference is exact in finite samples. In derivative pricing an exact characterization of finite sample uncertainty is critical from the insurance company's risk management point of view. The Bayesian approach is particularly attractive, since it can link the uncertainty of parameters and latent variables to the predictive uncertainty of the process. Another advantage of Bayesian inference is its ability to incorporate prior information into the model.

In estimating the equity index process with stochastic volatility and jumps we will follow the guidelines provided by Jones (1998), but make some generalizations. In our modelling framework we also allow the interest rate process to be stochastic, and we allow it to be correlated with the index and volatility processes. In order to value an American-style option we use the regression approach (see, for example, Longstaff and Schwartz, 2001).

Participating life insurance policies involve not only risks arising from financial factors, but also a risk related to mortality. Bacinello (2003) and Shen and Chu (2005) introduce mortality risk, but only in a simple set-up with deterministic or constant mortality rates. Biffis (2005) and Bacinello et al. (2008) incorporate stochastic mortality to the pricing framework. With a stochastic mortality model we do not need to make an assumption of a large insurance portfolio, and we avoid invoking to the law of large numbers. This again is significant from the risk management point of view.

Here we study dynamic hedging strategies to control for various risks by utilizing a replicating portfolio. As a hedging strategy we employ minimum-variance hedging which relies on the underlying asset as a single hedging instrument. We follow the work by Bakshi et al. (1997) when deriving a minimum-variance hedge. This type of hedge is needed, since a perfect delta-neutral hedge is not feasible due to untraded risks. However, a single-instrument hedge can only be partial, since in our set-up there is more than one source of risk. We also construct a conventional delta-neutral hedge which uses a simpler model for asset dynamics and compare the hedging performances.

The paper is organized as follows. Section 2 introduces the framework and models for the asset dynamics and mortality, Section 3 presents the estimation and evaluation procedures and Section 4 the empirical results. The final Section 5 concludes. The full conditional distributions of the option pricing and mortality models as well as the estimation results are provided in the appendices.

2. THE FRAMEWORK

2.1. The participating life insurance contract

We define the participating life insurance contract as in Luoma et al. (2008). The contract consists of two parts, the first being a guaranteed interest and the second a bonus depending on the yield of some total return equity index. We denote the amount of savings in the insurance contract at time t_i by $A(t_i)$. Then its growth during a time interval of length $\delta = t_{i+1} - t_i$ is given by

$$\log \frac{A(t_{i+1})}{A(t_i)} = g\,\delta + b \max\left(0, \log \frac{X(t_{i+1})}{X(t_i)} - g\,\delta\right), \tag{1}$$

where $X(t_i) = \sum_{j=0}^{q} S(t_{i-j})/(q+1)$ is a moving average of the total return equity index $S(t_i)$. The guarantee rate g is set to be less than the riskless interest rate, and it is fixed for one year at a time. It is set annually at kr_t, where r_t is the riskless short-term interest rate at time t and k is a positive constant less than 1. The bonus rate b is the proportion of the excessive equity index yield that is returned to the customer.

In this study we use the time interval $\delta = 1/255$, where 255 is approximately the number of the days in a year on which the index is quoted. The model also incorporates a surrender (early exercise) option. A further condition is that there will be a 1 % penalty if the contract is reclaimed during the first 10 working days. The penalty is not applied if the contract is reclaimed due to mortality.

In our framework the parameters k, g and b are predefined by the insurance company. Luoma et al. (2008) introduce a method to evaluate a fair bonus rate b so that the risk-neutral price of the contract is equal to initial savings. This gives the contract a simple structure and makes its costs and returns visible and predictable for the insurer and the customer.

2.2. Option pricing models

We assume that the dynamics of stock index S_t, variance V_t and riskless short-term rate r_t are described by the following system of SDEs:

$$\mathrm{d}\log S_t = \mu \mathrm{d}t + \sqrt{V_t}\mathrm{d}B_t^{(1)} + U_t \mathrm{d}q_t \tag{2a}$$

$$\mathrm{d}V_t = (\alpha_1 + \beta_1 V_t)\mathrm{d}t + \sigma_V \sqrt{V_t}\mathrm{d}B_t^{(2)} \tag{2b}$$

$$\mathrm{d}r_t = (\alpha_2 + \beta_2 r_t)\mathrm{d}t + \sigma_r \sqrt{r_t}\mathrm{d}B_t^{(3)} \tag{2c}$$

where $B_t^{(1)}$, $B_t^{(2)}$ and $B_t^{(3)}$ are standard Brownian motions, and q_t is a jump process with jump size U_t. We further assume that these Brownian motions have the correlation structure

$$\mathrm{Cor}\left(B_t^{(1)}, B_t^{(2)}, B_t^{(3)}\right) = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}, \tag{3}$$

and q_t is a Poisson process with intensity λ, that is, $\Pr(\mathrm{d}q_t = 1) = \lambda \mathrm{d}t$ and $\Pr(\mathrm{d}q_t = 0) = 1 - \lambda \mathrm{d}t$. Conditional on a jump occurring, we assume that $U_t \sim \mathrm{N}(a, b^2)$. In addition, we assume that q_t is uncorrelated with U_t or with any other process. We abbreviate this model as SVJ-SI.

In order to facilitate estimation, we reparameterize models (2a) -(2c) as follows:

$$\mathrm{d}\log S_t = \mu \mathrm{d}t + \sigma_1 \sqrt{Y_t}\mathrm{d}B_t^{(1)} + U_t \mathrm{d}q_t \tag{4a}$$

$$\mathrm{d}Y_t = (\alpha_1^* + \beta_1 Y_t)\mathrm{d}t + \sigma_2 \sqrt{Y_t}\mathrm{d}B_t^{(2)} \tag{4b}$$

$$\mathrm{d}R_t = (\alpha_2^* + \beta_2 R_t)\mathrm{d}t + \sigma_3 \sqrt{R_t}\mathrm{d}B_t^{(3)} \tag{4c}$$

where $Y_t = V_t/\sigma_1^2$ is rescaled variance and $R_t = 100\, r_t$ is the interest rate given in percentages. The new parameters are $\alpha_1^* = \alpha_1/\sigma_1^2$, $\sigma_2 = \sigma_V/\sigma_1$, $\alpha_2^* = 100\alpha_2$ and $\sigma_3 = 10\sigma_r$.

We introduce a risk-neutral probability measure Q under which the discounted index process $\tilde{S}_t = S_t \exp(-\int_0^t r_s ds)$ is a martingale. Specifically, we assume the risk neutral dynamics to be

$$\mathrm{d}\log S_t = \left(r_t - \frac{1}{2}V_t - \lambda\mu_J\right)\mathrm{d}t + \sqrt{V_t}\mathrm{d}Z_t^{(1)} + U_t \mathrm{d}q_t \tag{5a}$$

$$\mathrm{d}V_t = (\alpha_1 + \beta_1 V_t)\mathrm{d}t + \sigma_V \sqrt{V_t}\mathrm{d}Z_t^{(2)} \tag{5b}$$

$$\mathrm{d}r_t = (\alpha_2 + \beta_2 r_t)\mathrm{d}t + \sigma_r \sqrt{r_t}\mathrm{d}Z_t^{(3)} \tag{5c}$$

where $\mu_J = \exp(a+\frac{1}{2}b^2)-1$, and $Z_t^{(1)}$, $Z_t^{(2)}$ and $Z_t^{(3)}$ are three standard Brownian motions with correlation structure (3) under Q.

For the intensity of mortality, we will use a generalization of the Gompertz model. The Gompertz model describes the age dynamics of human mortality fairly accurately in the middle span of ages, approximately between 30 and 80 years, which is enough for our purposes (see, for example, Promislow, 2006). We use a stochastic generalization of the form

$$\log(\mu_{ku}) = \beta_{00} + \beta_{01}u + \beta_{10}k + \beta_{11}ku + \epsilon_{ku}, \tag{6}$$

where μ_{ku} is the death rate for age k and for cohort u set by the year of birth. We assume the error term ϵ_{ku} to follow an autoregressive process of order one: $\epsilon_{ku} = \phi\epsilon_{k-1,u} + a_{ku}$, where $a_{ku} \sim$ i.i.d. $\mathrm{N}(0, \sigma_m^2)$. Parameters $\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$, ϕ and σ_m^2 are unobservable and must be estimated.

3. ESTIMATION AND EVALUATION PROCEDURES

3.1. Finance model estimation

We use Bayesian methods to estimate the unknown parameters of the stock index, volatility and interest rate models as well as to estimate the latent volatility and jump processes. By doing so it is possible to take parameter uncertainty into account when the fair price of the contract is evaluated and a hedging strategy is employed. The major challenge in estimation is its high dimensionality, which results from the need to estimate latent processes.

We will use Euler discretization in the estimation of unknown parameters, since the transition density of the multivariate process described by (4a), (4b) and (4c) does not have a closed form solution. Accordingly, we will simulate the risk-neutral process using the Euler discretization of (5a), (5b) and (5c).

A discrete version of (4a), (4b) and (4c) is given by

$$\begin{aligned}\log S_{k+1} &= \log S_k + \mu\delta + \sqrt{Y_k\,\delta} e^{(1)}_{k+1} + U_{k+1} l_{k+1}\\ Y_{k+1} &= Y_k + (\alpha_1^* + \beta_1 Y_k)\delta + \sqrt{Y_k\,\delta} e^{(2)}_{k+1}\\ R_{k+1} &= R_k + (\alpha_2^* + \beta_2 R_k)\delta + \sqrt{R_k\,\delta} e^{(3)}_{k+1}\end{aligned}$$

where δ denotes discretization interval length, $e^{(1)}_k$, $e^{(2)}_k$ and $e^{(3)}_k$ are three normal variables with zero means, variances σ_1^2, σ_2^2 and σ_3^2, and correlation structure (3), $U_k \sim \mathrm{N}(a, b^2)$ is jump size and $l_k \sim \mathrm{Ber}(\lambda\delta)$ an indicator variable of a jump.

Our estimation procedure is a single-component (or cyclic) Metropolis-Hastings algorithm (see, for example, Gilks et al. (1996)). The Metropolis-Hastings (M-H) algorithm is a general term for Markov Chain Monte Carlo (MCMC) methods that are used to simulate posterior distributions. The algorithm was introduced by Hastings (1970) as a generalization of the Metropolis algorithm (Metropolis et al., 1953). Also the Gibbs sampler (Geman and Geman, 1984) can be viewed as its special case.

The single-component M-H algorithm differs from the basic algorithm in that the simulated random vector is divided into components which are updated one by one. The purpose is to simulate the conditional distribution of each block given the current values of the other blocks. In the case of the Gibbs sampler, random variates from these distributions are drawn directly. In the more general case, a proposal is first generated and it is accepted with certain probability, or otherwise the old value is retained.

In the case of our model, it is possible to divide the vector of all parameters to blocks which can be updated using Gibbs sampling, that is, the full conditionals of these blocks can be simulated directly. This is possible, since we have introduced a superfluous parameter σ_1 and we use the general correlation structure. Now posterior simulations of the dispersion matrix of the error vector $(e^{(1)}_k, e^{(2)}_k, e^{(3)}_k)$ can be drawn from the Inverse-Wishart density. Further details about the updating procedure are given in Appendix A.

Note that the data do not contain enough information to estimate σ_1 and the vector of scaled variances Y separately, but their joint posterior distribution determines the posterior of the variance vector V, which is of interest. Adding a new parameter is called parameter expansion, and it can be more generally used to improve the convergence of Markov chain simulation. This is discussed in Liu and Wu (1999), van Dyk and Meng (2001) and Liu (2003), and a simple example is provided by Gelman et al. (2004).

The volatility and jump processes cannot be updated using Gibbs sampling. Here we follow the guidelines provided by Jacquier et al. (1994) and Jones (1998). The scaled variances Y_k are updated one by one. Their full conditional distribution is $p(Y_k|Y_{-k}, H, \phi)$ where Y_{-k} comprises all of Y except Y_k, H comprises the index, interest rate and jump processes, and ϕ is a vector of all parameters. Since we are dealing with Markov processes,

$$p(Y_k|Y_{-k}, H, \phi) \propto p(Y_k|Y_{k-1}, H_{k-1}, H_k, \phi)p(Y_{k+1}, H_{k+1}|Y_k, H_k, \phi).$$

Now Y_k may be updated by first generating a proposal Y_k^* from $p(Y_k|Y_{k-1}, H_{k-1}, H_k, \phi)$ and accepting it with probability

$$\min\left(1, \frac{p(Y_{k+1}, H_{k+1}|Y_k^*, H_k, \phi)}{p(Y_{k+1}, H_{k+1}|Y_k, H_k, \phi)}\right).$$

A detailed description of this update can be found in Appendix A.

The jump process can be updated similarly. Let us denote the joint process of jumps and jump sizes as $I_k = (l_k, U_k)$ and the other processes as $L_k = (S_k, Y_k, R_k)$. Because the jumps are independent, their full conditional is given by $p(I_k|I_{-k}, L, \phi) = p(I_k|L_{k-1}, L_k, \phi)$, which is proportional to

$$p(I_k|\phi)p(L_k|L_{k-1}, I_k, \phi).$$

Now I_k is updated by first generating I_k^* from its marginal distribution $p(I_k|\phi)$ and accepting it with probability

$$\min\left(1, \frac{p(L_k|L_{k-1}, I_k^*, \phi)}{p(L_k|L_{k-1}, I_k, \phi)}\right).$$

The jumps l_k and their sizes U_k could also be updated separately. A detailed description of this update can be found in Appendix A.

3.2. Mortality estimation and prediction

To estimate the mortality model (6) we use Gibbs sampling except that the correlation parameter ϕ is updated with a Metropolis step. The needed conditional posterior distributions can be found in Appendix B. The data is imbalanced, since later cohorts have less observations. The unobserved future death rates are considered as missing observations and they are estimated similarly to the unknown parameters using Gibbs sampling. Each missing value is initially given the corresponding death rate from the most recent cohort where it is available. In the Gibbs sampler the missing values of each cohort are updated by generating them from their multivariate normal conditional distribution.

When the mortality model is used to study hedging performance, we need to scale the estimated AR(1) model, which is based on yearly data, to daily observations. When the sampling frequency is changed from 1 to δ, the high frequency sampling parameters are given as $\phi_{hf} = \phi^\delta$ and $\sigma^2_{m.hf} = \sigma^2_m \frac{1-\phi^{2\delta}}{1-\phi^2}$ (see, for example, Gourieroux and Jasiak, 2001).

When pricing and hedging the contract, we use the worst-case scenario of mortality from the insurance company's viewpoint. In practice, we simulate 1000 paths of death rates and choose the minimum rate for each time point. These minimum death rates are then used to generate the date of death for each simulation path used in pricing and simulation of hedging performance.

3.3. Pricing American options with regression methods

Our participating life insurance contract is an American-style option with a path-dependent moving average feature. An American option gives the holder the right to exercise the option at any time up to the expiry date. In pricing we adopt the least squares method introduced by Longstaff and Schwartz (2001). It is a simple but powerful approximation method for American-style options. The pricing is based on an optimal exercising strategy in which the goal is to find a stopping time maximizing the expected discounted payoff of the option. The decision to continue is based on comparing the discounted immediate exercise value with the corresponding discounted continuation value. In regression methods it is assumed that the continuation value may be expressed as linear regression using some basis functions.

In our application, the continuation values of the option depend on the path of the underlying index value in a complicated way. However, we consider that the current value of the index, its moving average, and the first index value appearing in the moving average may be used to predict the continuation value reasonably well. The use of the moving average may be motivated by observing that the growth of savings in the insurance contract depends on the path of the moving average (see Equation 1). The current index value and the first value appearing in the moving average help predict the future evolution of the moving average. We also use the current values of interest rate and volatility to predict the continuation value. The current amount of savings also helps predict the continuation value, but it is not included in the regression variables. Instead, it is subtracted from the regressed value before fitting the regression and subsequently added to the fitted value.

To avoid under- and overflows in the computations, the regression variables related to the equity index are scaled by the first index value, and the current value of the interest rate is given in percentages. Thus, the following state variables are used: $X_1(t_i) = S(t_i)/S(0)$, $X_2(t_i) = \left[\sum_{j=0}^{q} S(t_{i-j})/(q+1)\right]/S(0)$, $X_3(t_i) = S(t_{i-q})/S(0)$, $X_4(t_i) = R(t_i)$ and $X_5(t_i) = V(t_i)$. However, multicollinearity problems would occur if the variables X_1, X_2 and X_3 were used at all time points. In fact, X_3 would be equal for all simulations paths for $i \leq q$ and the moving averages X_2 would be very close to each other for small values of i. Therefore, we apply the following rule: The variables X_1, X_4 and X_5 are used for $i < q/2$, variables X_1, X_2, X_4 and X_5 are used for $q/2 \leq i < 3q/2$, and all variables are used for $i \geq 3q/2$. In this study the lag length of the moving average is chosen to be $q = 125$ (that is, half a year).

We use Laguerre polynomials, suggested by Longstaff and Schwartz (2001), as basis functions. More specifically, we use the first two polynomials

$$\begin{aligned} L_0(X) &= \exp(-X/2) \\ L_1(X) &= \exp(-X/2)(1-X) \end{aligned}$$

for all variables. In addition, we use the cross-products $L_0(X_1)L_0(X_4)$, $L_0(X_1)L_0(X_5)$, $L_0(X_1)L_0(X_2)$, $L_0(X_1)L_1(X_2)$, $L_1(X_1)L_0(X_2)$, $L_0(X_1)L_0(X_3)$ and $L_0(X_2)L_0(X_3)$. Thus, we have altogether 17 explanatory variables in the regression.

3.4. Determining the fair bonus rate

We have presented a method to determine a fair bonus rate in Luoma et al. (2008). Using the regression method we can determine the option price (that is, the price of the insurance contract) when the bonus rate b and the guarantee rate g have been given. However, we are interested to determine the bonus rate so that the price of the contract is equal to initial savings. This gives the contract a transparent structure. Furthermore, it makes the different hedging strategies comparable, since the bonus rate affects the duration of the contract, which is the most significant factor to produce large hedging errors. If the bonus rate is set at a high level, the contract is almost never reclaimed before the final expiration date, and, on the other hand, if the bonus rate is too low, early surrender is highly probable.

The problem of determining b is a kind of inverse prediction problem, and we need to estimate the option value for various values of b. Since we also wish to estimate the variance of the Monte Carlo error related to the regression method, we repeat the estimation several times for fixed values of b. We end up estimating a regression model where option price estimates are regressed on bonus rates. (This regression model should not be confused with the regression method used in the estimation of the option value for a fixed b). We found the third degree polynomial curve to be flexible enough for this purpose. After fitting the curve, we solve the bonus rate b for which the option price is equal to 100, which we assume to be the initial amount of savings. In order to facilitate the estimation of the fair bonus rate we set the further condition that there is a 1% penalty for reclaiming the contract during the first ten days.

Prior to fitting the polynomial, it is, however, necessary to determine an initial interval for the solution. For this purpose we have developed a modified bisection method. In this method, one first specifies initial upper and lower limits for the bonus rate; we use the values $l = 0$ and $u = 1$. Then one estimates the option price at $(l + u)/2$. If the price is greater than 100, the upper limit of the bonus rate is set at $u - (u - l)/4$; if the price is smaller than 100, the lower limit of the bonus rate is set at $l + (u - l)/4$. This procedure is continued until $u - l = 0.25$. Note that the new limit is not set in the middle of the interval, as is done in the ordinary bisection method, since this might lead to missing the correct solution due to the randomness of the price estimates.

Figure 1 illustrates the estimation procedure. The option price is estimated for 10 different bonus rates, and the estimation is repeated 5 times for each bonus rate, which produces 50 points to the scatter plot. Each estimation is based on 1000 simulated paths. The initial limits of the bonus rate (0.14, 0.39) were determined using the modified bisection method described above. When producing this figure, the time to maturity was set at 3 years, the guarantee rate at 1/3, the starting level of interest rate at 0.07 and mortality was not included. We can see that the fair bonus rate is approximately 0.28.

As mentioned above, the bonus rate is solved from the equation $y = f(x)$, where y is the price of the contract and

$$f(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 = \mathbf{x}'\hat{\boldsymbol{\beta}},$$

where $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)'$ is the ordinary least squares (OLS) estimate of the cubic regression model and $\mathbf{x} = (1, x, x^2, x^3)'$ a regression vector. The purpose of the initial penalty rate is to ensure that there is exactly one solution in the relevant interval.

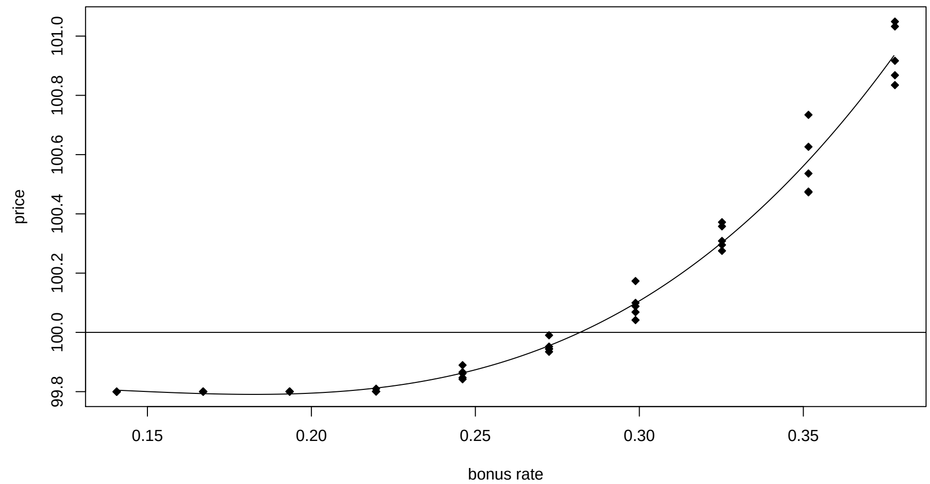

FIG. 1. Option price estimates vs. bonus rates.

Using the delta method, one also obtains an approximate variance for the estimate of x:

$$\mathrm{Var}(\hat{x}) \approx \frac{1}{[f'(x)]^2}\mathrm{Var}(f(x)) \approx \frac{1}{\left(\hat{\beta}_1 + 2\hat{\beta}_2\hat{x} + 3\hat{\beta}_3\hat{x}^2\right)^2}\hat{\mathbf{x}}'\mathrm{Cov}(\hat{\boldsymbol{\beta}})\hat{\mathbf{x}}.$$

4. HEDGING

4.1. Minimum variance hedging

We construct a single instrument hedge, which employs only the underlying stock. This hedge is only partial, since there are several sources of risks in our model. Uncontrolled risks are those which move the target option value but are uncorrelated with the underlying stock price. Such factors as model misspecification and transaction costs may render this type of hedge more practical to adopt than the conventional delta-neutral hedge. Besides a perfect delta-neutral hedge would be infeasible, since some of the risks are untraded.

Let N_t^S be the number of shares of the stock to be purchased and N_t^0 the residual cash position. Then the time t value of the replicating portfolio is $N_t^0 + N_t^S S_t$. Furthermore, the hedging error $H_{t+\delta}$ at time $t + \delta$ is given by

$$H_{t+\delta} = N_t^S S_{t+\delta} + N_t^0 e^{r_t \delta} - C_{t+\delta}, \tag{7}$$

where δ is the updating interval of the replicating portfolio and $C_{t+\delta}$ is the value of the contract at time $t + \delta$. In the limit when $\delta \to 0$, the mean squared hedging error is minimized by choosing

$$N_t^S = \frac{\mathrm{Cov}\,(\mathrm{d}S_t, \mathrm{d}C_t)}{\mathrm{Var}\,(\mathrm{d}S_t)}. \tag{8}$$

Let us denote the jump sizes of the index process S_t by $J_t \doteq e^{U_t} - 1$. The mean and variance of J_t are given by by $\mu_J = \exp(a + \frac{1}{2}b^2) - 1$ and $\sigma_J^2 = \exp(2a + b^2)(\exp(b^2) - 1)$, respectively. Under our framework, the total return variance can be decomposed into two components

$$\frac{1}{\mathrm{d}t}\mathrm{Var}\left(\frac{\mathrm{d}S_t}{S_t}\right) = V_t + V_t^J, \tag{9}$$

where the instantaneous variance of the jump component is given by

$$\begin{aligned} V_t^J &= (1/\mathrm{d}t)\mathrm{Var}\left(J_t \mathrm{d}q_t\right) = (1/\mathrm{d}t)\left(\mathrm{E}(J_t \mathrm{d}q_t)^2 - [\mathrm{E}(J_t)\mathrm{E}(\mathrm{d}q_t)]^2\right) \\ &= (1/\mathrm{d}t)\left[\sigma_J^2 + \left(\mu^J\right)^2\right]\left[\lambda \mathrm{d}t + (\lambda \mathrm{d}t)^2\right]. \end{aligned}$$

Now let $C_t(S_t, V_t, r_t)$ denote the value of the contract at time t with index value S_t, variance V_t and interest rate r_t. The differential of $C_t(S_t, V_t, r_t)$ may be written as

$$\begin{aligned} \mathrm{d}C_t(S_t, V_t, r_t) = {} & \frac{\partial C_t(S_t, V_t, r_t)}{\partial S_t}\sqrt{V_t}\mathrm{d}Z_t^{(1)} + \frac{\partial C_t(S_t, V_t, r_t)}{\partial V_t}\mathrm{d}V_t + \frac{\partial C_t(S_t, V_t, r_t)}{\partial r_t}\mathrm{d}r_t \\ & + \left[C_t(S_t + J_t S_t, V_t, r_t) - C_t(S_t, V_t, r_t)\right]\mathrm{d}q_t. \end{aligned}$$

Using this and equations (8) and (9) we obtain that

$$\begin{aligned} N_t^S = {} & \Delta_t^{(S)}\frac{V_t}{(V_t + V_t^J)} + \Delta_t^{(V)}\frac{\rho_{12}\sigma_V V_t}{S_t(V_t + V_t^J)} + \Delta_t^{(r)}\frac{\rho_{13}\sigma_r\sqrt{V_t r_t}}{S_t(V_t + V_t^J)} \\ & + \frac{\lambda\left[\mathrm{E}_t\left(J_t C_t(S_t + J_t S_t, V_t, r_t)\right) - C_t(S_t, V_t, r_t)\mu_J\right]}{S_t(V_t + V_t^J)} \end{aligned} \tag{10}$$

where we have denoted the deltas as $\Delta_t^{(S)} = \frac{\partial C_t(S_t, V_t, r_t)}{\partial S_t}$, $\Delta_t^{(V)} = \frac{\partial C_t(S_t, V_t, r_t)}{\partial V_t}$ and $\Delta_t^{(r)} = \frac{\partial C_t(S_t, V_t, r_t)}{\partial r_t}$.

Equation (10) shows that the position to be taken in the stock must control not only for the direct impact of stock price changes on the target option, but also for the indirect impacts of those parts of volatility and interest rate changes which are correlated with stock price fluctuations. We can see that the additional number of shares needed besides $\Delta^{(S)}$ is increasing both in ρ_{12} and ρ_{13}. Furthermore, since the jump risk is present as well, the position to be taken in the underlying stock must also hedge the impact of jump risk on the target option, which is reflected in the last term in (10). This term is increasing in λ and μ_J, meaning that the larger the random-jump risk, the more adjustment needs to be made in the hedging position.

In theory the constructed partial hedge requires continuous rebalancing to reflect the changing market conditions. In practice, only discrete rebalancing is possible. Suppose that portfolio rebalancing takes place at intervals of length δ. At time t, the replicating portfolio has N_t^S shares of the stock and the residual is invested in an instantaneously maturing riskfree bond. The combined position is a self-financed portfolio. At time $t + \delta$ the hedging error is as in (7).

Our contract is an American-style derivative, which we price using the regression method (see, for example, Longstaff and Schwartz, 2001). With the estimated regression model we may also compute the deltas $\Delta^{(S)}$, $\Delta^{(V)}$ and $\Delta^{(r)}$, and determine the optimal stopping times needed when simulating hedging performance in several cases. For each simulation path the used stopping time is the first time when the estimated continuation value is smaller than the immediate exercise value.

Comparison of the hedging schemes is done so that the difference of the replication portfolio and the balance is computed for each simulation path at its estimated optimal stopping time. Then the mean difference and mean squared error is computed over all the simulation paths. Moreover, this procedure is repeated 50 times using different regression estimates and the results of the repetitions are pooled.

4.2. Competing model and delta-neutral hedging

A similar approach is used when a delta-neutral hedge is constructed for a simpler model whose real-world asset dynamics of this model are described as

$$\mathrm{d}r_t = \kappa(\xi - r_t)\mathrm{d}t + \sigma r_t^{\gamma}\mathrm{d}W_t^{(1)} \tag{11a}$$

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \nu S_t^{1-\alpha}\mathrm{d}W_t^{(2)}. \tag{11b}$$

Here $W_t^{(1)}$ and $W_t^{(2)}$ are two standard Brownian motions, correlated through $W_t^{(2)} = \rho W_t^{(1)} + \sqrt{1-\rho^2}W_t^{(3)}$, where $W_t^{(1)}$ and $W_t^{(3)}$ are independent standard Brownian motions. The risk neutral dynamics are obtained by replacing the drift μ in (11b) with r_t. Details on estimation and pricing under this model may be found in Luoma et al. (2008). We abbreviate this model as CEV-SI.

In the delta-neutral hedge corresponding to this model, the number of shares in the replication portfolio is given by

$$N_t^S = \frac{\partial C_t(S_t, r_t)}{\partial S_t} = \Delta_t^{(S)} \geq 0.$$

Again we use the regression method to price the derivative and to compute $\Delta^{(S)}$.

5. EMPIRICAL RESULTS

5.1. Estimation of the parameters

In order to experiment with actual data and to estimate the unknown parameters of the models, we chose the following data sets: As an equity index we use the Total Return of Dow Jones EURO STOXX Total Market Index (TMI), which is a benchmark covering approximately 95 per cent of the free float market capitalization of Europe. The objective of the index is to provide a broad coverage of companies in the Euro zone including Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain. The index is constructed by aggregating the stocks traded on the major exchanges of Euro zone. Only common stocks and those with similar characteristics are included, and any stocks that have had more than 10 non-trading days during the past three months are removed. In estimation, we use daily quotes from March 4th, 2002 until December 6th, 2007.

As a proxy for riskless short-term interest rate, we use Eurepo, which is the benchmark rate of the large Euro repo market. Eurepo is the rate at which one prime bank offers funds in euro to another prime bank if in exchange the former receives from the latter Eurepo GC as collateral. It is a good benchmark for secured money market transactions in the Euro zone. In the estimation of the interest rate model we use the 3 month Eurepo rate, since it behaves more regularly than the rates with shorter maturities. Both the index and interest series are presented in Figure 2.

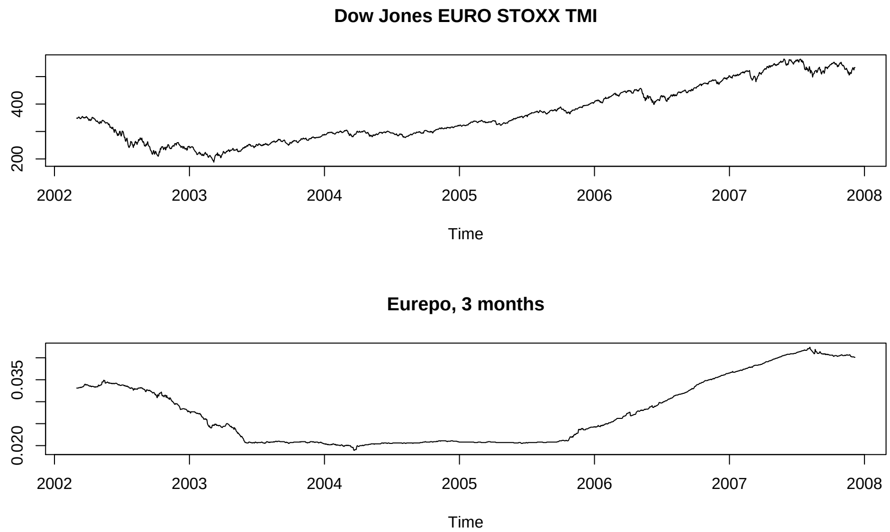

FIG. 2. The equity index and interest series.

In mortality modelling we use mortality data provided by the Human Mortality Database (see http://www.mortality.org). It was created to provide detailed mortality and population data to those interested in the history of human longevity. The project began as an outgrowth of earlier projects in the Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock, Germany. The project seeks to provide open, international access to the database which contains detailed population and mortality data for 37 countries or areas. In our work we use Finnish mortality data for females between ages 30 and 80. More specifically, we use cohort death rates for cohorts born between 1926 and 1961.

All computations were made and figures produced using the R computing environment (see http://www.r-project.org). To speed up computations we coded the most time consuming loops in C++. We had no remarkable convergence problems in the MCMC simulation used in estimation. Estimation of the finance model (2a)-(2c) was computationally most challenging, and we simulated three chains of length 200000 and picked every 10th simulation to obtain accurate results. In the estimation of the mortality model all chains converged rapidly to their stationary distributions. The summary of the estimation results, as well as Gelman and Rubin's diagnostics (see Gelman et al., 2004), are given in Appendix C. The values of the diagnostic are close to 1 and thus indicate good convergence.

5.2. Hedging results

There are several parameters which may be varied in the participating life insurance contract described by Equation (1). We set the lag length of the moving average at 125 days, the number of simulated paths in contract price estimation at 1000 and the number of estimation repetitions at 50. Furthermore, we set the duration of the contract to be 3 or 10 years and the starting level of interest rate

0.04 or 0.07. We do not fix the guarantee rate at a constant value throughout the entire contract period but set it at 0, 1/3 or 2/3 of the short-term rate at intervals of one year.

We calculate the results with and without mortality. When mortality is incorporated into the framework, we calculate the results for clients of ages 60 and 80. We assume that the contract period starts at the beginning of year 2008, which means that we use cohorts born in 1947 and 1927. There is only slight difference in the results when models without mortality are compared with models including mortality, and 60 years old clients are considered. Therefore, we present the results concerning 80 years old clients only. Moreover, although we set the update frequency of the replicating portfolio to be one day, one week and one month, we only present results for daily and monthly updates, since the results concerning daily and weekly updates do not differ considerably. Table 1 shows the fair bonus rates and hedging errors when minimum variance hedging with SVJ-SI model is used, while Table 2 shows the results when delta-neutral hedging with CEV-SI model is used. Table 3 shows the results of delta-neutral hedging with CEV-SI model when the real-world predictive simulations are however generated from the SVJ-SI model. The corresponding results with mortality and 80 year old clients may be found in Tables 5, 6 and 7.

TABLE 1.

Fair bonus rates and hedging errors with SVJ model (no mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.242	0.053	0.436	0.048	0.59
3	0.04	1/3	0.172	0.038	0.223	0.035	0.296
3	0.04	2/3	0.089	0.035	0.064	0.035	0.077
3	0.07	0	0.384	0.166	1.282	0.225	1.747
3	0.07	1/3	0.281	0.163	0.616	0.163	0.887
3	0.07	2/3	0.156	0.096	0.219	0.095	0.287
10	0.04	0	0.257	-0.043	9.081	0.081	8.423
10	0.04	1/3	0.188	-0.08	5.56	-0.023	5.206
10	0.04	2/3	0.092	0.024	0.9	0.031	0.887
10	0.07	0	0.389	0.076	28.569	0.107	33.261
10	0.07	1/3	0.291	-0.056	20.218	-0.091	22.916
10	0.07	2/3	0.158	-0.086	7.334	-0.065	7.624

From all the tables we may see that the estimated fair bonus rate increases as the guarantee rate decreases. This is logical but it is less obvious why the fair bonus rate also increases as the starting level of the interest rate increases. The probable explanation is as follows: When the interest rate is larger the level of the index grows more rapidly, since the ’percentage drift’ equals the riskless interest rate under risk-neutral probability. This makes negative returns in the moving average of the stock index less probable, and the feature of the contract which protects the accumulated capital against negative

TABLE 2.

Fair bonus rates and hedging errors with CEV model (pred. CEV)(no mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.325	0.125	0.35	0.147	0.474
3	0.04	1/3	0.235	0.095	0.175	0.098	0.243
3	0.04	2/3	0.137	0.032	0.074	0.038	0.094
3	0.07	0	0.50	0.207	0.74	0.226	0.985
3	0.07	1/3	0.383	0.152	0.426	0.156	0.549
3	0.07	2/3	0.223	0.071	0.141	0.079	0.181
10	0.04	0	0.333	0.19	5.235	0.175	4.903
10	0.04	1/3	0.254	0.068	1.69	0.063	1.753
10	0.04	2/3	0.145	0.026	0.25	0.043	0.305
10	0.07	0	0.496	-0.035	21.917	-0.021	19.832
10	0.07	1/3	0.384	0.043	6.91	0.029	7.957
10	0.07	2/3	0.241	-0.002	1.924	0.01	1.554

TABLE 3.

Fair bonus rates and hedging errors with CEV model (pred. SVJ)(no mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.325	-0.312	0.948	-0.287	1.118
3	0.04	1/3	0.235	-0.185	0.449	-0.211	0.521
3	0.04	2/3	0.137	-0.16	0.187	-0.15	0.2
3	0.07	0	0.50	-0.549	2.325	-0.509	2.436
3	0.07	1/3	0.383	-0.404	1.246	-0.362	1.246
3	0.07	2/3	0.223	-0.245	0.41	-0.205	0.407
10	0.04	0	0.333	-0.401	2.731	-0.36	2.664
10	0.04	1/3	0.254	-0.316	1.183	-0.342	1.293
10	0.04	2/3	0.145	-0.225	0.286	-0.194	0.291
10	0.07	0	0.496	-0.591	9.511	-0.672	12.614
10	0.07	1/3	0.384	-0.465	4.24	-0.491	5.42
10	0.07	2/3	0.241	-0.383	1.114	-0.408	1.37

returns becomes less important. This, in turn, decreases the contract price, which is compensated by the increase in the bonus rate.

From Table 1 we may see that the mean difference of the replicating portfolio value and the balance (MD) is positive with the 3 years contract, which is desirable from the insurance company's risk management point of view. The probable reason for this result is that the stopping rule based on the estimated regression model is slightly suboptimal. We also see that the MD decreases when the starting level of the interest rate increases or when the guarantee rate decreases. With the 10-year contracts the sign of the MD may be positive or negative, which may be due to uncertainty involved in long-term simulation. However, one should note that the MD is close to zero in all cases.

When comparing the mean square errors (MSEs) one can see that they increase as the initial interest rate increases, the duration of the contract increases or the guarantee rate decreases. The increase of the initial interest rate from 4% to 7% makes the MSE about 3 times larger with the 3-year contract and 4 to 7 times larger with the 10-year contract. The increase in the duration of the contract has the largest effect on the MSE. With 4% and 7% initial interest rates the MSE becomes about 15 and 30 times larger, respectively, when the duration of the contract changes from 3 to 10 years. The increase of the guarantee from 0 to $2/3$ reduces the error efficiently. This is understandable, since a larger guarantee reduces fluctuation in the value of the contract.

When the replicating portfolio is updated monthly instead of daily, the MSE increases slightly in the case of the 3-year contract. When the 10-year contract is considered, no systematic effect of the updating frequency is detected, probably because other types of errors become so large in a long-run simulation.

TABLE 4.

Hedging error of 2000 sample paths in a case of no guarantee, interest rate starting level 0.07, 10 year contract and no mortality.

hedge model	pred. model	Min	Q_1	Median	Mean	Q_3	Max	sd
SVJ	SVJ	-102.40	-0.186	0.075	-0.191	0.429	67.60	6.128
CEV	SVJ	-79.02	-0.764	-0.143	-0.660	0.039	6.31	2.945
CEV	CEV	-71.03	0.016	0.093	-0.049	0.297	24.45	3.241

As one can see from Table 2, the results are similar when the CEV-SI model and delta-neutral hedging are used. In almost all cases the hedging MSE is slightly lower than with SVJ-SI model and minimum variance hedging. This is natural, since there are less sources of error in a simpler model. When real-world predictions are simulated from the SVJ-SI model (see Table 3) and the hedging is based on the CEV-SI model, the MDs are substantially negative, which is not a desirable situation from the insurance company's viewpoint. As the initial interest rate increases the MD moves further away from the zero. The same effect takes place when the length of the contract increases or when the guarantee rate decreases. On the other hand, the MSEs of delta hedging are about double when compared with the MSEs of minimum variance hedging in the 3-year contracts, but in the case of 10-year contracts the MSEs of delta hedging are smaller as one can see from Table 3. This is somewhat surprising but may be understood by looking at the distribution of the hedging errors. Table 4 shows the basic statistics of the hedging errors when the two hedging strategies are applied in the cases of two predictive distributions. The hedging error distribution of the delta-neutral scheme combined with predictions from the SVJ-SI model is most skewed, since it does not have large positive errors like the other cases do. This makes the standard deviation smaller than in the other cases. On the other hand, the median of the distribution is negative and the upper quartile is only slightly positive whereas both of them are positive in the other cases.

TABLE 5.

Fair bonus rates and hedging errors with SVJ model (with mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.248	0.001	0.55	-0.024	0.708
3	0.04	1/3	0.176	-0.01	0.272	0.016	0.318
3	0.04	2/3	0.087	0.02	0.071	0.036	0.067
3	0.07	0	0.389	0.095	1.487	0.162	1.74
3	0.07	1/3	0.287	0.094	0.761	0.12	0.94
3	0.07	2/3	0.155	0.094	0.191	0.058	0.304
10	0.04	0	0.257	-0.403	7.077	-0.17	3.944
10	0.04	1/3	0.182	-0.172	2.294	-0.179	2.206
10	0.04	2/3	0.09	-0.004	0.265	-0.039	0.457
10	0.07	0	0.401	-0.377	21.226	-0.408	22.392
10	0.07	1/3	0.296	-0.259	10.137	-0.269	10.458
10	0.07	2/3	0.161	-0.107	2.808	-0.149	3.291

From Tables 5, 6 and 7 one can see the effect of mortality. The most dramatic problem seems to be the fact that the MD moves from the positive to negative side when 10-year contracts are considered. The most alarming situation is in the case when predictions come from the SVJ-SI model and the hedging is done with the delta-neural scheme and CEV-SI model. In this situation the MDs are farthest away from zero. With the 3-year contract the MDs are close to zero in all other cases except when SVJ-SI predictions and delta-neutral hedging with the CEV-SI model are used. The MSEs slightly increase in almost all cases. The only case where they become smaller is when the SVJ-SI model is used for the 10-year contracts. This may be explained by studying the error distributions summarized in Table 8. The maximum of the distribution is much smaller than in the case when mortality is not included.

These results indicate that model error might be crucial when hedging is applied to participating life insurance. In the worst scenarios the errors would mean large losses to the insurance company.

6. CONCLUSIONS

In this paper we present a full Bayesian analysis of valuation and hedging of a participating life insurance contract. The Bayesian approach enables us to exploit MCMC methods and to take parameter uncertainty into account in valuation and hedging. We value the contract with the regression method, since it embeds an American-style surrender option. In the valuation we take both financial and mor-

TABLE 6.

Fair bonus rates and hedging errors with CEV model (pred. CEV)(with mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.328	0.098	0.401	0.14	0.59
3	0.04	1/3	0.245	0.066	0.224	0.06	0.285
3	0.04	2/3	0.133	0.024	0.084	0.053	0.075
3	0.07	0	0.507	0.158	0.862	0.153	1.059
3	0.07	1/3	0.389	0.111	0.526	0.118	0.618
3	0.07	2/3	0.223	0.066	0.15	0.075	0.19
10	0.04	0	0.344	-0.038	3.712	-0.155	6.558
10	0.04	1/3	0.245	-0.039	2.486	-0.011	2.311
10	0.04	2/3	0.137	-0.011	0.675	-0.02	0.58
10	0.07	0	0.525	-0.376	29.431	-0.316	29.538
10	0.07	1/3	0.397	-0.19	12.943	-0.147	11.309
10	0.07	2/3	0.229	-0.033	2.011	-0.035	1.792

TABLE 7.

Fair bonus rates and hedging errors with CEV model (pred. SVJ)(with mortality).

contract length	interest rate starting level	guarantee rate	fair bonus rate	MD with daily update	MSE with daily update	MD with monthly update	MSE with monthly update
3	0.04	0	0.328	-0.318	0.963	-0.273	1.257
3	0.04	1/3	0.245	-0.245	0.58	-0.257	0.669
3	0.04	2/3	0.133	-0.161	0.21	-0.081	0.138
3	0.07	0	0.507	-0.577	2.818	-0.518	2.703
3	0.07	1/3	0.389	-0.436	1.447	-0.378	1.504
3	0.07	2/3	0.223	-0.223	0.411	-0.204	0.443
10	0.04	0	0.344	-0.387	2.246	-0.501	3.146
10	0.04	1/3	0.245	-0.283	1.189	-0.265	1.121
10	0.04	2/3	0.137	-0.17	0.395	-0.163	0.349
10	0.07	0	0.525	-0.783	10.474	-0.744	12.139
10	0.07	1/3	0.397	-0.548	5.496	-0.498	4.179
10	0.07	2/3	0.229	-0.261	1.051	-0.257	0.968

tality risks into account. The financial model allows the interest rate, volatility and jumps in the index process to be stochastic. As a stochastic mortality model we use a generalization of the Gompertz model.

TABLE 8.

Hedging error of 2000 sample paths in a case of no guarantee, interest rate starting level 0.07, 10 year contract and mortality.

hedge model	pred. model	Min	Q_1	Median	Mean	Q_3	Max	sd
SVJ	SVJ	-140.80	-0.138	0.156	-0.441	0.309	5.430	4.218
CEV	SVJ	-99.76	-0.772	-0.064	-0.799	0.091	8.248	3.626
CEV	CEV	-188.80	0.040	0.132	-1.118	0.233	18.650	10.017

The main steps in this paper are the estimation of the financial and mortality models, generation of the posterior predictive distributions, pricing the American-style contract, evaluation of the fair bonus rate, and hedging the contract with a single-instrument minimum variance hedge. We repeat all these steps using the CEV model with stochastic interest rate. With this simpler model we construct the conventional delta-neutral hedge, and compare its performance with minimum variance hedging and the more complicated model.

We find that the duration of the contract is the most significant factor to produce large hedging errors. Therefore, in order to make the different hedging strategies comparable, is important to determine a fair bonus rate for each case studied. If the bonus rate is set at a high level, the contract is almost never reclaimed before the final expiration date, and, on the other hand, if the bonus rate is too low, early surrender is highly probable.

One of the major findings of our simulation experiments is that the mean difference (MD) of the replicating portfolio value and the balance is positive in the 3-year contracts, which is desirable from the insurance company's risk management point of view. In the 10-year contracts the sign of the MD is sometimes positive and sometimes negative, which may be due to the uncertainty involved in long-term simulation. Moreover, the updating frequency of the replicating portfolio has no systematic effect on the hedging error, probably because other types of errors become so large in these longer contracts. The hedging MSEs are considerably larger in the 10-year contracts than in the 3-year contracts. The reason is that the error distributions are extremely heavy-tailed in the longer contracts.

We also find that there is only slight difference in the results between the case when mortality is not taken into account and the case when it is and the contract is started at the age of 60. When comparing the models without mortality with those with mortality and 80 years old clients, we note that for some reason the MD is clearly negative in the 10-year contracts.

Furthermore, we find that in almost all cases the hedging MSE with CEV-SI model and delta-neutral hedging is slightly lower than that with SVJ-SI model and minimum variance hedging. This is natural, since there are less sources of error in the simpler model. When the real-world predictions are simulated from the SVJ-SI model and the hedging is based on the CEV-SI model, the MDs are substantially negative, which is not a desirable situation from the insurance company's viewpoint.

The findings of this article indicate that in some cases the distribution of the hedging error has a negative mean. However, this is not a serious problem, since the location of the error distribution can be easily shifted by reducing the bonus rate slightly. The real problem are some rare paths of the financial series which cause a heavy left tail in the hedging error distribution in long-term contracts.

Although these paths are extremely rare, they might lead to crucial hedging errors and large losses to the insurance company, unless one finds a way to hedge against them.

REFERENCES

Ballotta, L., Haberman, S., Wang, N., 2006. Guarantees in with-profit and unitized with-profit life insurance contracts: fair valuation problem in presence of the default option. *The Journal of Risk and Insurance*, 73, 97–121.

Bacinello, A., 2003. Fair valuation of a guaranteed life insurance participating contract embedding a surrender option. *The Journal of Risk and Insurance*, 70, 461–487.

Bacinello, A.R., Biffis, E., Millossovich, P., 2008. Pricing life insurance contracts with early exercise features. *Journal of Computational and Applied Mathematics,* in press.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *Journal of Finance*, 52, 2003–2049.

Bauer, D., Kiesel, R., Kling, A., Ruß, J., 2006. Risk-neutral valuation of participating life insurance contracts. *Insurance: Mathematics and Economics*, 39, 171–183.

Bernard, C., Le Courtois, O., Quittard-Pinon, F., 2005. Market value of life insurance contracts under stochastic interest rates and default risk. *Insurance: Mathematics and Economics*, 36, 499–516.

Biffis, E., 2005. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37, 443–468.

Briys, E. and de Varenne, F., 1997. On the risk of previous termlife insurancenext term liabilities: debunking some common pitfalls. *Journal of Risk Insurance*, 64, 673–694.

Briys, E. and de Varenne, F., 1997. Valuing risky fixed rate debt: an extension, *Journal of Financial and Quantitative Analysis*, 32, 239–248.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC.

Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 6, 721–741.

Gilks, W. R., Richardson, S., and Spiegelhalter, D., eds., 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall.

Glasserman, P., 2003. *Monte Carlo Methods in Financial Engineering*. Springer.

Gourieroux, C. and Jasiak, J., 2001. *Financial Econometrics*. Princeton Series in Finance.

Grosen, A., Jorgensen, P., 2000. Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies. *Insurance: Mathematics and Economics*, 26, 37–57

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrica*, 57, 97–109.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12, 371–389.

Jones, C., 1998. Bayesian estimation of continuous-time finance models. Working paper, Simon School of Business, University of Rochester, New York.

Liu, C., 2003. Alternating Subspace-Spanning Resampling to Accelerate Markov Chain Monte Carlo simulation. *Journal of the American Statistical Association*, 98, 110–117.

Liu, J., and Wu, Y. N., 1999. Parameter expansion in for data augmentation. *Journal of the American Statistical Association*, 94, 1264–1274.

Longstaff, F.A., Schwartz, E.S., 2001. Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14, 113–148.

Luoma, A., Puustelli, A., Koskinen, L., 2008. Bayesian analysis of participating life insurance contracts with American-style options. In Proc. of AFIR Colloquium, Rome, Italy, 31 September – 3 October, 2008.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E., 1953. Equations of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21, 1087–1092.

Promislow, S.D., 2006. *Fundamentals of Actuarial Mathematics*. John Wiley & Sons, Ltd.

Shen, W. and Xu, H., 2005. The valuation of unit-linked policies with or without surrender options. *Insurance: Mathematics and Economics*, 36, 79–92.

Tanskanen, A. J., Lukkarinen, J., 2003. Fair valuation of path-dependent participating life insurance contracts. *Insurance: Mathematics and Economics*, 33, 595–609.

van Dyk, D. A., and Meng, X. L., 2001. The art of data augmentation (with discussion). *Journal of Computational and Graphical Statistics*, 10, 1–111.

Zaglauer, K., Bauer, D., 2008. Risk-neutral valuation of participating life insurance contracts in a stochastic interest rate environment. *Insurance: Mathematics and Economics*, in press.

APPENDICES

A. Full conditional distributions of the option pricing model

Let us denote $\gamma_1 = (\alpha_1^*, \beta_1)$, $\gamma_2 = (\alpha_2^*, \beta_2)$, $\phi = (\mu, \gamma_1, \gamma_2, \sigma_1, \sigma_2, \sigma_3, \rho_{12}, \rho_{13}, \rho_{23})$, $I_k = (l_k, U_k)$, $Y = (Y_1, \ldots Y_{K-1})$ and

$$e_k^{(1)} = \frac{\log S_k - \log S_{k-1} - \mu\delta - l_k U_k}{\sqrt{Y_{k-1}\delta}},$$

$$e_k^{(2)} = \frac{Y_k - Y_{k-1} - \left(\alpha_1^* + \beta_1 Y_{k-1}\right)\delta}{\sqrt{Y_k \delta}},$$

$$e_k^{(3)} = \frac{R_k - R_{k-1} - \left(\alpha_2^* + \beta_2 R_{k-1}\right)\delta}{\sqrt{R_{k-1}\delta}},$$

$$e_k = \begin{pmatrix} e_k^{(1)} \\ e_k^{(2)} \\ e_k^{(3)} \end{pmatrix},$$

and

$$\Sigma = \mathrm{Cov}\left(e_k^{(1)}, e_k^{(2)}, e_k^{(3)}\right) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}.$$

Update for μ with Gibbs sampler

Prior: $p(\mu) \propto 1$

Conditional posterior:

$$\{\mu | \mathrm{z}^{(1)}, Y, \sigma_{1.23}\} \sim \mathrm{N}\left(\frac{\sum_{k=1}^{K-1} \frac{z_{k+1}^{(1)}}{Y_k}}{\sum_{k=1}^{K-1} \frac{1}{Y_k}}, \frac{\sigma_{1.23}}{\delta \sum_{k=1}^{K-1} \frac{1}{Y_k}} \right),$$

where

$$z^{(1)} = \left(z_1^{(1)}, z_2^{(1)}, \ldots, z_K^{(1)}\right),$$

$$z_{k+1}^{(1)} = \frac{\log(S_{k+1}) - \log(S_k) - l_{k+1}U_{k+1}}{\delta} - \sqrt{\frac{Y_k}{\delta}}\mu_{k+1}^{(1.23)},$$

$$\mu_{k+1}^{(1.23)} = \begin{pmatrix}\sigma_{12} & \sigma_{13}\end{pmatrix}\begin{pmatrix}\sigma_2^2 & \sigma_{23}\\ \sigma_{23} & \sigma_3^2\end{pmatrix}^{-1}\begin{pmatrix}e_{k+1}^{(2)}\\ e_{k+1}^{(3)}\end{pmatrix},$$

and

$$\sigma_{1.23} = \sigma_1 - \begin{pmatrix}\sigma_{12} & \sigma_{13}\end{pmatrix}\begin{pmatrix}\sigma_2^2 & \sigma_{23}\\ \sigma_{23} & \sigma_3^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma_{12}\\ \sigma_{13}\end{pmatrix}.$$

Update for γ_1 with Gibbs sampler

Prior: $p(\gamma_1) \propto 1$
Conditional posterior:

$$\{\gamma_1 | z^{(2)}, Y, \sigma_{2.13}\} \sim \mathrm{N}\left(\left(X'\Delta^{-1}X\right)^{-1} X'\Delta^{-1}z^{(2)}, \frac{\sigma_{2.13}}{\delta}\left(X'\Delta^{-1}X\right)^{-1}\right),$$

where

$$z^{(2)} = \left(z_1^{(2)}, z_2^{(2)}, \ldots, z_K^{(2)}\right),$$

$$z_{k+1}^{(2)} = \frac{Y_{k+1} - Y_k}{\delta} - \sqrt{\frac{Y_k}{\delta}}\mu_{k+1}^{(2.13)},$$

$$X = \begin{pmatrix}1 & Y_1\\ 1 & Y_2\\ \vdots & \vdots\\ 1 & Y_{K-1}\end{pmatrix},$$

$$\Delta^{-1} = \begin{pmatrix}\frac{1}{Y_1} & 0 & \ldots & 0\\ 0 & \ddots & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \ldots & 0 & \frac{1}{Y_{K-1}}\end{pmatrix},$$

$$\mu_{k+1}^{(2.13)} = \begin{pmatrix}\sigma_{12} & \sigma_{23}\end{pmatrix}\begin{pmatrix}\sigma_1^2 & \sigma_{13}\\ \sigma_{13} & \sigma_3^2\end{pmatrix}^{-1}\begin{pmatrix}e_{k+1}^{(1)}\\ e_{k+1}^{(3)}\end{pmatrix},$$

and

$$\sigma_{2.13} = \sigma_2^2 - \begin{pmatrix}\sigma_{12} & \sigma_{23}\end{pmatrix}\begin{pmatrix}\sigma_1^2 & \sigma_{13} \\ \sigma_{13} & \sigma_3^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma_{12} \\ \sigma_{23}\end{pmatrix}.$$

Update for γ_2 with Gibbs sampler

Prior: $p(\gamma_2) \propto 1$
Conditional posterior:

$$\{\gamma_2 | z^{(3)}, X_*, \sigma_{3.12}\} \sim \mathrm{N}\left(\left(X_*'\Delta_*^{-1}X_*\right)^{-1} X_*'\Delta_*^{-1}z^{(3)}, \frac{\sigma_{3.12}}{\delta}\left(X_*'\Delta_*^{-1}X_*\right)^{-1}\right),$$

where

$$z^{(3)} = \left(z_1^{(3)}, z_2^{(3)}, \ldots, z_K^{(3)}\right),$$

$$z_{k+1}^{(3)} = \frac{R_{k+1} - R_k}{\delta} - \sqrt{\frac{R_k}{\delta}}\mu_{k+1}^{(3.12)},$$

$$X_* = \begin{pmatrix} 1 & R_1 \\ 1 & R_2 \\ \vdots & \vdots \\ 1 & R_{K-1} \end{pmatrix},$$

$$\Delta_*^{-1} = \begin{pmatrix} \frac{1}{R_1} & 0 & \ldots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \frac{1}{R_{K-1}} \end{pmatrix},$$

$$\mu_{k+1}^{(3.12)} = \begin{pmatrix}\sigma_{13} & \sigma_{23}\end{pmatrix}\begin{pmatrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2\end{pmatrix}^{-1}\begin{pmatrix}e_{k+1}^{(1)} \\ e_{k+1}^{(2)}\end{pmatrix},$$

and

$$\sigma_{3.12} = \sigma_3^2 - \begin{pmatrix}\sigma_{13} & \sigma_{23}\end{pmatrix}\begin{pmatrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma_{13} \\ \sigma_{23}\end{pmatrix}.$$

Update for Σ with Gibbs sampler

Prior: $p(\Sigma) \sim \text{Inv-Wishart}(\Psi, m)$
Posterior: $p(\Sigma | \ldots) \sim \text{Inv-Wishart}(\Psi + A, m + K - 1)$, where

$$A = \sum_{k=2}^{K} e_k e_k'.$$

Update for volatility with Metropolis-Hastings step

Let us denote $H_k = (\log S_k, I_k, R_k)$. The conditional posterior of Y_k:

$$p(Y_k|, Y_{k-1}, Y_{k+1}, H_{k-1}, H_k, H_{k+1}, \phi)$$
$$\propto p(Y_k|Y_{k-1}, H_{k-1}, H_k, \phi)\, p(Y_{k+1}, H_{k+1}|Y_k, H_k, \phi)$$

Proposal Y_k^* is generated from $p(Y_k|Y_{k-1}, H_{k-1}, H_k, \phi)$:

$$Y_k^* = Y_{k-1} + (\alpha_1^* + \beta_1 Y_{k-1})\delta + \sqrt{Y_{k-1}\delta} e_k^{(2)*},$$

where $e_k^{(2)*} \sim \mathrm{N}\left(\mu_k^{(2.13)}, \sigma_{2.13}\right)$. For $k = 1$ the proposal is generated from unconditional distribution $\Pr(Y_1|\phi)$. Since Y is a CIR process, its stationary distribution is $\mathrm{Gamma}\left(-\frac{2\alpha_1^*}{\sigma_2^2}, -\frac{\sigma_2^2}{2\beta_1}\right)$.
Acceptance probability:

$$\min\left(1, \frac{p(Y_{k+1}, H_{k+1}|Y_k^*, H_k, \phi)}{p(Y_{k+1}, H_{k+1}|Y_k, H_k, \phi)}\right)$$
$$= \min\left\{1,\ \exp\left[-\log(Y_k^*) + \log(Y_k) - \frac{1}{2}\left(e_{k+1}^{*\prime}\Sigma^{-1}e_{k+1}^* - e_{k+1}'\Sigma^{-1}e_{k+1}\right)\right]\right\},$$

where e_{k+1}^* is computed using Y_k^*. For $k = K$ the acceptance probability cannot be computed. The proposal is accepted with probability 1.

Update for the parameters of the jump process with Gibbs sampler

Prior: $\lambda_0 \sim \mathrm{Beta}(p_1, p_2)$
Posterior: $\Pr(\lambda_0|l) \propto \mathrm{Beta}(p_1 + \sum l_i, p_2 + K - \sum l_i)$
Priors: $b^2 \sim \text{Inv-}\chi^2(df_0, \sigma_0^2)$, $\quad a|b^2 \sim \mathrm{N}(a_0, b^2/b_0)$
Posteriors:

$$b^2|l, U \sim \text{Inv-}\chi^2\left(df_0 + n, \frac{1}{df_0 + n}\left(df_0\sigma_0^2 + (n-1)s^2 + \frac{b_0 n}{b_0 + n}(\bar{U} - a_0)^2\right)\right)$$

$$a|b^2, l, U \sim \mathrm{N}\left(\frac{b_0 a_0 + n\bar{U}}{b_0 + n}, \frac{b^2}{b_0 + n}\right)$$

where $n = \sum l_i$, $\bar{U} = \frac{1}{n}\sum l_i U_i$, $s^2 = \frac{1}{n-1}\sum l_i(U_i - \bar{U})^2$.

Update for the jump process with Metropolis-Hastings step

Let us denote and $L_k = (\log S_k, Y_k, R_k)$. Then the full conditional distribution of I_k is

$$p(I_k|I_{k-1}, I_{k+1}, L_k, L_{k-1}, L_{k+1}, \phi) = p(I_k|L_{k-1}, L_k, \phi) \propto$$
$$p(I_k|L_{k-1}, \phi)\, p(L_k|I_k, L_{k-1}, \phi) = p(I_k|\phi)\, p(L_k|I_k, L_{k-1}, \phi)$$

Proposal from distribution $p(I_k|\phi)$: $l_k^* \sim \mathrm{Ber}(\lambda_0)$, $U_k^* \sim \mathrm{N}(a, b^2)$.
Acceptance probability:

$$\min\left(1, \frac{p(L_k|L_{k-1}, I_k^*, \phi)}{p(L_k|L_{k-1}, I_k, \phi)}\right)$$
$$= \min\left\{1,\ \exp\left[-\frac{1}{2}\left(e_k^{*\prime}\Sigma^{-1}e_k^* - e_k'\Sigma^{-1}e_k\right)\right]\right\},$$

where e_k^* is computed using I_k^*.

B. Full conditional posterior distributions of the mortality model

Let us denote $y_{ku} = \log(\mu_{ku})$ for $k = 1, ..., K$ and $u = 1, ..., U$. Furthermore, $y_u = (y_{1u}, ..., y_{Ku})$, $y = (y_1, ..., y_U)$ and $X = (X_1', \ldots, X_U')'$, where

$$X_u = \begin{pmatrix} 1 & u & 1 & u \cdot 1 \\ 1 & u & 2 & u \cdot 2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & u & K & u \cdot K \end{pmatrix},$$

and $\beta = (\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11})$.

The inverse of $Cor(y|\beta, \phi)$ is

$$R_*^{-1} = \begin{pmatrix} R^{-1} & & 0 \\ & \ddots & \\ 0 & & R^{-1} \end{pmatrix} = I_U \otimes R^{-1}$$

where I_U is $U \times U$ identity matrix and

$$R = \begin{pmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{K-1} \\ \phi & 1 & \phi & \ldots & \phi^{K-2} \\ \vdots & & \ddots & & \vdots \\ \phi^{K-1} & \phi^{K-2} & \ldots & \phi & 1 \end{pmatrix}.$$

Update for σ_m^2 with Gibbs sampler

Prior: $p(\sigma_m^2) \propto \frac{1}{\sigma_m^2}$
Conditional posterior: $\sigma^2|y, \beta, \phi \sim \text{Inv-}\chi^2(KU, \text{SS}/KU)$, where $\text{SS} = (y - X\beta)' R_*^{-1} (y - X\beta)$.

Update for β with Gibbs sampler

Prior: $p(\beta) \propto 1$
Conditional posterior: $\beta|y, \phi, \sigma_m^2 \sim \text{N}(\mu_\beta, \sigma_m^2 V_\beta)$, where $\mu_\beta = (X' R_*^{-1} X)^{-1} X' R_*^{-1} y$ and $V_\beta = (X' R_*^{-1} X)^{-1}$.

Update for ϕ with Metropolis step

Prior: $p(\phi) = I_{(-1,1)}(\phi)$
Conditional posterior: $p(\phi|y, \beta, \sigma_m^2) \propto (1 - \phi^2)^{-\frac{1}{2}U(K-1)} \exp\left(-\frac{1}{2\sigma_m^2} SS\right) I_{(-1,1)}(\phi)$

C. Estimation results of finance and mortality models

The posterior simulations were performed using the R computing environment. The following outputs were obtained using the summary function of the add-on package MCMCpack:

TABLE 9.

Estimation results of finance model.

```
Number of chains = 3
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

               Mean        SD  Naive SE Time-series SE
mu        0.1157251 0.0365450 2.110e-04      7.554e-04
alpha1    0.1968175 0.0386987 2.234e-04      1.658e-03
beta1    -6.2996549 1.3530007 7.812e-03      5.483e-02
alpha2    0.2154899 0.1286814 7.429e-04      7.452e-04
beta2    -0.0495622 0.0429811 2.482e-04      2.270e-04
sigma22V  0.2218156 0.0444790 2.568e-04      2.137e-03
sigma33   0.0140068 0.0005198 3.001e-06      3.108e-06
rho12    -0.7681546 0.0523322 3.021e-04      2.070e-03
rho13     0.0792596 0.0264295 1.526e-04      2.204e-04
rho23    -0.1322315 0.0517667 2.989e-04      1.381e-03
a        -0.0060781 0.0068594 3.960e-05      1.366e-04
b2        0.0003484 0.0002063 1.191e-06      2.366e-06
lambda0   0.0090029 0.0027855 1.608e-05      6.318e-05
```

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
mu	0.0426332	0.0911853	0.1162357	0.1406814	0.1859535
alpha1	0.1271985	0.1697973	0.1944505	0.2216304	0.2790440
beta1	-9.1239610	-7.1682486	-6.2259077	-5.3563520	-3.8502831
alpha2	0.0281787	0.1218338	0.1949668	0.2853363	0.5272006
beta2	-0.1612529	-0.0704730	-0.0382768	-0.0170129	-0.0015817
sigma22V	0.1457985	0.1902363	0.2177842	0.2496907	0.3184338
sigma33	0.0130214	0.0136508	0.0139906	0.0143487	0.0150585
rho12	-0.8556995	-0.8054466	-0.7729459	-0.7370593	-0.6501383
rho13	0.0266995	0.0615367	0.0792856	0.0971150	0.1311709
rho23	-0.2335275	-0.1670558	-0.1321860	-0.0979035	-0.0299741
a	-0.0193522	-0.0104156	-0.0062760	-0.0019092	0.0080387
b2	0.0001421	0.0002256	0.0002996	0.0004109	0.0008392
lambda0	0.0051479	0.0065661	0.0085743	0.0111619	0.0145301

Gelman and Rubin's diagnostics
(Potential scale reduction factors):

	Point est.	97.5% quantile
mu	1.00	1.00
alpha1	1.00	1.00
beta1	1.00	1.01
alpha2	1.00	1.00
beta2	1.00	1.00
sigma22V	1.00	1.00
sigma33	1.00	1.00
rho12	1.01	1.02
rho13	1.00	1.00
rho23	1.00	1.00
a	1.00	1.01
b2	1.01	1.01
lambda0	1.01	1.02

TABLE 10.

Estimation results of mortality model.

```
Number of chains = 3
Sample size per chain = 2500

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

             Mean        SD  Naive SE Time-series SE
beta00 -7.218e+00 0.0230773 2.665e-04      3.524e-04
beta01 -1.391e-02 0.0011916 1.376e-05      6.582e-05
beta10  7.248e-02 0.0008204 9.473e-06      3.685e-05
beta11 -2.229e-05 0.0001162 1.341e-06      7.854e-06
sigma2m 2.941e-02 0.0015699 1.813e-05      4.937e-05
phi     2.486e-01 0.0366845 4.236e-04      1.603e-03

2. Quantiles for each variable:

             2.5%        25%        50%        75%      97.5%
beta00 -7.2635215 -7.2333503 -7.218e+00 -7.203e+00 -7.1731612
beta01 -0.0162149 -0.0146982 -1.392e-02 -1.315e-02 -0.0115150
beta10  0.0708985  0.0719250  7.247e-02  7.302e-02  0.0740813
beta11 -0.0002524 -0.0001005 -1.917e-05  5.772e-05  0.0002014
sigma2m 0.0266480  0.0283436  2.928e-02  3.038e-02  0.0327533
phi     0.1769527  0.2246326  2.472e-01  2.714e-01  0.3236979

Gelman and Rubin's diagnostics
(Potential scale reduction factors):

       Point est. 97.5% quantile
beta00       1.00           1.00
beta01       1.01           1.04
beta10       1.01           1.02
beta11       1.02           1.06
sigma2m      1.00           1.00
phi          1.00           1.01
```