Claims Reserving When There Are Negative Values in the Runoff Triangle: Bayesian Analysis Using the Three-Parameter Log-Normal Distribution

Enrique de Alba*

Abstract

This paper is concerned with the situation that occurs in claims reserving when there are negative values in the development triangle of incremental claim amounts. Typically these negative values will be the result of salvage recoveries, payments from third parties, total or partial cancellation of outstanding claims due to initial overestimation of the loss or to a possible favorable jury decision in favor of the insurer, rejection by the insurer, or just plain errors. Some of the traditional methods of claims reserving, such as the chain-ladder technique, may produce estimates of the reserves even when there are negative values. However, many methods can break down in the presence of enough (in number and/or size) negative incremental claims if certain constraints are not met. Historically the chain-ladder method has been used as a gold standard (benchmark) because of its generalized use and ease of application. A method that improves on the gold standard is one that can handle situations where there are many negative incremental claims and/or some of these are large. This paper presents a Bayesian model to consider negative incremental values, based on a three-parameter log-normal distribution. The model presented here allows the actuary to provide point estimates and measures of dispersion, as well as the complete distribution for outstanding claims from which the reserves can be derived. It is concluded that the method has a clear advantage over other existing methods. A Markov chain Monte Carlo simulation is applied using the package WinBUGS.

1. Introduction

The estimation of adequate reserves for outstanding claims is one of the main activities of actuaries in property/casualty insurance. The need to estimate future claims has led to the development of many loss-reserving techniques. Probably the oldest and most widely used of these techniques is the well-known chain ladder. It is frequently used as a benchmark because of its generalized use and ease of application (Hess and Schmidt 2002). In its original form the chain ladder is a nonstochastic algorithm for producing estimates of outstanding claims. There are many variations of the method; a description of one of them that will be useful in what follows will now be provided.

Assume that the time (number of periods) it takes for the claims to be completely paid is fixed and known, that payments are made annually, and that the development of partial payments follows a stable pay-off pattern. This is in agreement with many existing models for claims reserving in nonlife (general) insurance that assume, explicitly or implicitly, that the proportion of claim payments, payable in the j-th development period, is the same for all periods of origin (de Alba 2002b; Hess and Schmidt 2002).

* Enrique de Alba, ASA, PhD, is Dean and Professor of Actuarial Science in the Division of Actuarial Science, Statistics and Mathematics, at ITAM, Rio Hondo No. 1, Mexico City, 01080, Mexico, dealba@itam.mx. He is also an Adjunct Professor in the Department of Statistics and Actuarial Science at the University of Waterloo, Ontario, Canada.

The results are applicable to any frequency of claim payments (years, quarters, etc.) and length of pay-off period.

The term *claims reserving* is used here in its most general sense. Essentially the data would correspond to a typical run-off triangle used in loss reserving. In this paper the following notation is used: let Z_{it} = incremental amount of claims in the t-th development year corresponding to year of origin (or accident year) i. Thus $\{Z_{it};\ i = 1, \ldots, k,\ t = 1, \ldots, s\}$, where s = maximum number of years (subperiods) it takes to completely pay out the total number (or amount) of claims corresponding to a given exposure year. This paper does not assume $Z_{it} > 0$ for all $i = 1, \ldots, k$ and $t = 1, \ldots, s$. Most claims-reserving methods usually assume that $s = k$ and that we know the values Z_{it} for $i + t \le k + 1$. The known values are presented in the form of a run-off triangle in Table 1.

Although the models described in this paper can be applied to more general shapes of claims data, other than a triangle, we will assume that we have a conventional triangle of data. Thus, for ease of comparability with other methods, and without loss of generality, assume that the data consist of a triangle of incremental claims: $\{Z_{it}: t = 1, \ldots, k - i + 1;\ i = 1, \ldots, k\}$.

The cumulative claims are defined by $W_{ij} = \sum_{t=1}^{j} Z_{it}$, $i = 1, 2, \ldots, k$, and the development factors of the chain-ladder technique are denoted by $\{\lambda_j: j = 2, \ldots, k\}$. The estimates of the development factors from the standard chain-ladder technique are

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{k-j+1} W_{ij}}{\sum_{i=1}^{k-j+1} W_{i,j-1}} \tag{1.1}$$

(see Verrall 1989). These estimates are then applied to the latest cumulative claims available in each row ($W_{i,k-i+1}$) to produce forecasts of future values of cumulative claims:

$$\hat{W}_{i,k-i+2} = W_{i,k-i+1}\hat{\lambda}_{k-i+2}, \qquad i = 2, \ldots, k \tag{1.2}$$

and

$$\hat{W}_{i,j} = \hat{W}_{i,j-1}\hat{\lambda}_j, \qquad i = 3, \ldots, k, \qquad j = k - i + 3, \ldots, k. \tag{1.3}$$

The required reserve for the i-th accident year will then be $\hat{R}_i = \hat{W}_{ik} - W_{i,k-i+1}$, and the total reserve is $\hat{R} = \sum_{i=2}^{k} \hat{R}_i$.

Note that the estimators of the development factors can be written as

$$\hat{\lambda}_j = 1 + \frac{\sum_{i=1}^{k-j+1} Z_{i,j}}{\sum_{i=1}^{k-j+1} W_{i,j-1}} = 1 + r_j, \tag{1.4}$$

where

Table 1

Typical Development Triangle Used in General Insurance Claims Reserving

Year of Origin	Development Year						
	1	2	…	t	…	$k-1$	k
1	Z_{11}	Z_{12}	…	Z_{1t}		$Z_{1,k-1}$	Z_{1k}
2	Z_{21}	Z_{22}	…	Z_{2t}		$Z_{2,k-1}$	—
3	Z_{31}	Z_{32}	…	Z_{3t}		—	—
⋮						—	—
$k-1$	$Z_{k-1,1}$	$Z_{k-1,2}$				—	—
k	Z_{k1}	—					—

$$r_j = \frac{\sum_{i=1}^{k-j+1} Z_{i,j}}{\sum_{i=1}^{k-j+1} W_{i,j-1}} \tag{1.5}$$

is the rate of change of the cumulative claims between the development years $(j - 1)$ and j. Hence we can see that if the numerator is negative and the denominator positive, we will have $r_j < 0$, and so $\hat{\lambda}_j < 1$. Furthermore, in the literature on the chain-ladder method, in general it is implicitly assumed that $\hat{\lambda}_j > 0$. Clearly, formulas (1.1)–(1.5) can be applied no matter what values we get for r_j and $\hat{\lambda}_j$. However, in some cases the resulting estimated reserves may be meaningless. Now, the numerator in formula (1.5) will be negative if there are enough (in size and/or number) $Z_{i,j} < 0$ in the j-th column.

The estimators of the development factors in equation (1.1) can be rewritten as

$$\hat{\lambda}_j = \sum_{i=1}^{k-j+1} \frac{W_{i,j-1}}{\sum_{i=1}^{k-j+1} W_{i,j-1}} \left(\frac{W_{ij}}{W_{i,j-1}}\right);$$

that is, they are weighted sums of the ratios $W_{ij}/W_{i,j-1}$. For these estimators to be well defined it is necessary to assume $W_{i,j} > 0$, $i = 1, \ldots, k$; $j = 1, \ldots, k - i + 1$. Otherwise some of the weights will be negative. This assumption is implicit in most discussions of the chain ladder, since the W_{ij} used are positive and no mention is made as to what might happen if this were not the case. It is only sometimes made explicit (Hess and Schmidt 2002; Mack 2004). Under this assumption, the estimates of the development factors will be defined even if the sum of the known incremental claims of the j-th column is negative, that is, $\Sigma_{i=1}^{k-j+1} Z_{i,j} < 0$, for some j. Thus the chain-ladder technique can be applied in the presence of negative incremental claim values, in which case the consequence will be to have $\hat{\lambda}_j < 1$ for one or more j. Other methods are not so "robust," and the estimates may not be obtainable.

Negative incremental values can arise in the run-off triangle as a result of salvage recoveries, payments from third parties, total or partial cancellation of outstanding claims due to initial overestimation of the loss or to a possible favorable jury decision in favor of the insurer, rejection by the insurer, or just plain errors. England and Verrall (2002) argue that it is probably better to use paid claims rather than incurred claims (paid losses and aggregate case reserve estimates combined) since negative values are less likely to appear in the former. That is because case reserve estimates, the amount set aside by the claims handlers (see Chamberlin 1989; Brown and Gottlieb 2001), are set individually and often tend to be conservative, resulting in overestimation in the aggregate. Adjusting for this overestimation in the later stages of development may lead to negative incremental amounts. Whatever their cause, the presence of these negative incremental values in the data may cause problems when applying some claims-reserving methods. Thus, ideally, before applying claims-reserving methods, the actuary will revise and correct the data to eliminate negative incremental values. In this respect de Alba and Bonilla (2002) provide a list of potential adjustments frequently used in practice. However, even after correcting the data it is not always possible to eliminate all the negative values. Hence it is convenient to have available claims-reserving methods that will allow the actuary to compute the necessary reserves even in the presence of the negative values that may remain in the data.

In this paper a stochastic model is proposed that will yield optimal Bayesian estimates of the reserves for outstanding claims. In particular this paper is concerned with the situation in which there are many negative incremental claims in the development triangle and/or some of these are large.

The remainder of the paper is structured as follows. Section 2 presents stochastic claims-reserving models, Bayesian and non-Bayesian, that can be applied when there are negative claim values. Section 3 describes a Bayesian model for claim amounts in the presence of negative values. Section 4 describes how prior distributions are specified in the model. Section 5 describes how to use this model to estimate reserves. Its implementation for Markov chain Monte Carlo simulations, as well as an example, is given in Section 6. All models are presented only in discrete time.

2. STOCHASTIC MODELS

Stochastic models for claims reserving improve on the classical approach by allowing the actuary to obtain measures of uncertainty and sometimes the complete distribution of outstanding claims. For a comprehensive, although not exhaustive, review of existing stochastic methods for claims reserving see England and Verrall (2002) or Hess and Schmidt (2002). Most of the methods presented in these references use the point of view of frequentist or classical statistics. Hess and Schmidt concentrate on stochastic models for the chain ladder.

Mack (1993) presents one of the earliest attempts at formalizing a stochastic model for claims reserving. He proposes a nonparametric model that reproduces the chain ladder and obtains distribution-free expressions for the standard of reserve estimates. The use of the model is not limited by the existence of negative incremental claims. What may be considered a limitation is that it is directed at reproducing the chain-ladder reserves. England and Verrall (2002) have proposed the use of bootstrapping to compute the prediction errors.

England and Verrall (2002) emphasize the framework of generalized linear models (GLMs; Anderson et al. 2004). They provide predictions and prediction errors for the different methods discussed and show how the predictive distributions may be obtained by bootstrapping and Monte Carlo methods. Among the models considered by England and Verrall there are several that can handle negative values: an (overdispersed) Poisson, a negative binomial, and a normal approximation to the negative binomial. They also mention the log-normal model that was introduced by Kremer (1982) and analyzed in detail in Verrall (1991) when there are some negative incremental claims. Referring to the Poisson model they argue that "this does not imply that it is only suitable for data consisting exclusively of positive integers. That constraint can be overcome using a 'quasi-likelihood' approach, which can be applied to non-integer data, positive and negative." A similar argument is used for the negative binomial. It is not the purpose of this paper, but one could question whether they are really using those distributions or some continuous approximation. Also, since long-tailed distributions are used for modeling claim data, the normal distribution is not appropriate.

In the context of GLMs the first stochastic version of the chain-ladder method that can be applied in the presence of negative incremental claim values is defined as a generalized linear model with an overdispersed Poisson distribution (Renshaw and Verrall 1998). In the overdispersed Poisson model the mean and variance are not the same. Using our previous notation $E[Z_{it}] = m_{it}$, with variance function $V(Z_{it}) = \phi m_{it}$ and scale parameter $\phi > 0$, combined with the function $\log(m_{it}) = \mu + \alpha_i + \beta_t$, where α_i and β_t represent row and column effects in the triangle, respectively. Overdispersion is achieved if $\phi > 1$. The model reproduces the estimates of the classical chain-ladder method. Estimates of the parameters, $\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\beta}_t$, are obtained by using a "quasi-likelihood" approach (Anderson et al. 2004).

Renshaw and Verrall (1998) point out that their procedure "is not applicable to all sets of data, and can break down in the presence of a sufficient number of negative incremental claims." "Sufficient" means that there are enough incremental claims and their values are such that they make $\sum_{i=1}^{k-j+1} Z_{ij} < 0$, for some $j = 1, \ldots, k$. Note that this is the condition for $\hat{\lambda}_j < 1$ in the chain ladder; see equation (1.4). They then point out that to obtain estimates using this method it is necessary that $\sum_{i=1}^{k-j+1} Z_{ij} > 0$ for all $j = 1, \ldots, k$ (Verrall 2000). Using this constraint, the same reserve is produced by this model and the chain-ladder technique.

The negative binomial model is closely related to the Poisson model (Verrall 2000). The distribution in the GLM is now assumed to be a negative binomial with mean $(\lambda_j - 1)W_{i,j-1}$ and variance $\phi\lambda_j(\lambda_j - 1)W_{i,j-1}$, where $W_{ij} = \sum_{t=1}^{j} Z_{it}$. The parameters $\{\lambda_j : j = 2, \ldots, n\}$ are the typical chain-ladder development factors defined in Section 1. As in the Poisson model, ϕ is an overdispersion parameter. This model yields essentially the same estimates as the (overdispersed) Poisson. Again, with enough negative incremental claims, it is possible that some of the λ (one would be enough) become less than one, and so that clearly the variance would not exist and the model cannot be applied.

A third stochastic model for the chain ladder, mentioned in England and Verrall (2002), that can be applied in the presence of negative incremental claim values is a normal approximation to the negative binomial model, under which the chain-ladder results can still be reproduced. The approximation assumes the distribution of incremental claims is normal with mean $(\lambda_j - 1)W_{i,j-1}$, as in the negative binomial model, but the variance is assumed to be $\phi_j W_{i,j-1}$. This model is seen to be equivalent to one proposed by Mack (1993). England and Verrall recommend that bootstrapping be used to estimate the variance under the normal approximation.

An extensive review of the application of Bayesian methods in actuarial science will not be given here. Rather the approach will be described very briefly and some of their applications in loss reserving mentioned. For general discussion on Bayesian theory and methods see Berger (1985), Bernardo and Smith (1994), or Zellner (1971a).

Bayesian analysis of IBNR reserves has been considered by Jewell (1989, 1990), Verrall (1990), and Haastrup and Arjas (1996); see de Alba (2004) for a review. For a discussion of Bayesian methods in actuarial science see Klugman (1992), Makov (2001), Makov, Smith, and Liu (1996), Scollnik (2001, 2002), Ntzoufras and Dellaportas (2002), de Alba (2002b, 2004), and Verrall (2004). Kunkler (2004) provides a Bayesian claims-reserving method for the situation in which there are zeros in the development triangle, but not negative incremental claim values.

Next consider one existing Bayesian model, which can be applied to situations where $Z_{it} < 0$ for some $i, t = 1, \ldots, k$. Verrall (2004) presents a Bayesian formulation of the Bornhuetter-Ferguson (B-F) technique, which is a variant of the chain ladder that uses external information to obtain an initial estimate for the amount of expected ultimate claims, W_{ik}, for each $i = 2, \ldots, k$. This is then combined with the development factors of the chain-ladder technique to estimate outstanding claims (Brown and Gottlieb 2001; Chamberlin 1989). This is clearly well suited for the application of Bayesian methods if the initial information about ultimate claims is given in terms of a prior distribution (Verrall 2004). However, the method may break down in the presence of enough negative values (in number and/or size), certainly if any column sum of the incremental claims in the development triangle is negative, that is, again if $\sum_{i=1}^{k-j+1} Z_{it} < 0$.

We have seen that there are a number of stochastic claims-reserving methods that may work in the presence of negative incremental claims, but that they will break down if enough negative values are present. In the next section a Bayesian model is proposed that can be applied under very general conditions.

3. A BAYESIAN MODEL FOR AGGREGATE CLAIMS

This section presents a model for the unobserved aggregate claim amounts and hence the necessary reserves for outstanding claims. The approach followed is set out in de Alba (2002a), where use is made of the three-parameter log-normal distribution to estimate outstanding claims reserves in the presence of negative incremental claims. Let the random variable Z_{it} represent the value of incremental claims amounts in the t-th development year of accident year i, $i, t = 1, \ldots, k$. The Z_{it} are known for $i + t \le k + 1$, and we let

$$Y_{it} = \log(Z_{it} + \delta), \tag{3.1}$$

where δ is called the "threshold" parameter. If $Y_{it} \sim N(\mu, \sigma^2)$, then Z_{it} has a three-parameter log-normal distribution, and its density is

$$f(z_{it}|\mu, \sigma^2, \delta) = \begin{cases} \dfrac{1}{\sigma(z_{it} + \delta)\sqrt{2\pi}} \exp\left\{-\dfrac{1}{2\sigma^2}(\log(z_{it} + \delta) - \mu)^2\right\} & \text{for } z_{it} > -\delta \\ 0 & \text{for } z_{it} \le -\delta \end{cases}.$$

In general, the threshold parameter does not have restrictions (Crow and Schimidzu 1988) except that $(z_{it} + \delta) > 0$ must hold. In our claims-reserving problem the threshold parameter $\delta > 0$ adjusts the negative incremental claim values so as to ensure $(z_{it} + \delta) > 0$, for $i, t = 1, \ldots, k$, with $i + t \le k + 1$. Let us also assume that

$$Y_{it} = \log(Z_{it} + \delta) = \mu + \alpha_i + \beta_t + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \sigma^2), \tag{3.2}$$

$i = 1, \ldots, k$, $t = 1, \ldots, k$, and $i + t \leq k + 1$ so that Z_{it} follows a three-parameter log-normal distribution, denoted by $Z_{it} \sim LN(\mu_{it}, \sigma^2, \delta)$ with $\mu_{it} = \mu + \alpha_i + \beta_t$ and

$$f(z_{it}|\mu, \alpha_i, \beta_t, \sigma^2, \delta) = \frac{1}{\sigma(z_{it} + \delta)\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log(z_{it} + \delta) - \mu - \alpha_i - \beta_t)^2\right] \times I_{(-\delta,\infty)}(z_{it}), \tag{3.3}$$

where $I_A(x)$ is the usual indicator function.

Here α_i and the β_t, $i, t = 1, \ldots, k$, represent the accident year (row) and development year (column) effects, respectively. The model in equation (3.2) corresponds to an unbalanced two-way analysis of variance (ANOVA) model. It is well known in ANOVA that certain restrictions must be imposed on the parameters to attain estimability in equation (1.3). This paper uses the assumption that $\alpha_1 = \beta_1 = 0$ (Verrall 1990).

Let $z = \{z_{it}; i, t = 1, \ldots, k, i + t \leq k + 1\}$ be a T_U-dimension vector that contains all the observed values of Z_{it}, where $Y_{it} = \log(Z_{it} + \delta)$, and $\theta = (\mu, \alpha_2, \ldots, \alpha_k, \beta_2, \ldots, \beta_k)'$ is the $((2k - 1) \times 1)$ vector of parameters. The likelihood function will be $f(z|\theta, \sigma^2, \delta) = \Pi f(z_{it}|\theta, \sigma^2, \delta)$, where the product is over the T_U known z_{it} values, $i, t = 1, \ldots, k, i + t \leq k + 1$.

In de Alba (2002a) the threshold parameter δ is first estimated by maximum likelihood, substituted to define $y_{it} = \log(z_{it} + \hat{\delta})$, and then the "profile" likelihood, $L(\theta, \sigma^2|z, \hat{\delta}) = f(z|\theta, \sigma^2, \hat{\delta})$, is obtained. This is the likelihood function with δ replaced by its ML estimator, say, $\hat{\delta}$ (Crow and Shimizu 1988, p. 123). So the profile likelihood is used instead of the likelihood $L(\theta, \sigma^2, \delta|z) = f(z|\theta, \sigma^2, \delta)$ to carry out the Bayesian analysis, which then is done using results for a two-parameter lognormal distribution (Zellner 1971b). The approach followed in de Alba (2002a) has the disadvantage that the variability due to estimating δ is not taken into account in the inference process. This can be potentially troublesome since it is well known that the estimates of δ can be very unstable (see Cohen and Whitten 1980; Johnson, Kotz, and Balakrishnan 1994, chap. 14; Hill 1963). Here the full likelihood function $L(\theta, \sigma^2, \delta|z)$ is used.

4. PRIOR DISTRIBUTIONS

To carry out the Bayesian analysis we must now specify prior distributions for the parameters θ, σ^2, and δ. Assume that a priori the parameters are independent so that

$$f(\theta, \sigma^2, \delta) = f(\theta)f(\sigma^2)f(\delta) = f(\mu) \times \left[\prod_{i=2}^{k} f(\alpha_i)\right]\left[\prod_{t=2}^{k} f(\beta_t)\right] \times f(\sigma^2) \times f(\delta).$$

Any dependence among the parameters that is not modeled explicitly a priori will be reflected in the posterior distribution since it will be introduced by the sample data through the likelihood function. The posterior distribution will be obtained as

$$f(\theta, \sigma^2, \delta|z) \propto f(z|\theta, \sigma^2, \delta) \times f(\mu) \times \left[\prod_{i=2}^{k} f(\alpha_i)\right]\left[\prod_{t=2}^{k} f(\beta_t)\right] \times f(\sigma^2) \times f(\delta). \tag{4.1}$$

A hierarchical model will be used below. In hierarchical models, at a first stage the data are specified to come from a given distribution, $f(z_{it}|\mu, \alpha_i, \beta_t, \sigma^2, \delta)$ in our case. At the second stage, the parameters are assumed to follow their own (prior) distributions, here $f(\mu)$, $f(\alpha_i)$, $f(\beta_t)$, $f(\sigma^2)$, and $f(\delta)$, $i, t = 2, \ldots, k$. Markov chain Monte Carlo (MCMC) simulation then will be used to generate samples from the posterior distributions of the parameters as well as the predictive distribution of the reserves. This can be implemented with the package WinBUGS 1.4 (Spiegelhalter et al. 2001) or OpenBUGS (http://mathstat.helsinki.fi/openbugs/). In the next section the specification of the prior distributions is described.

A prior distribution can be specified using any distribution that is reasonable according to the characteristics of the parameter. When there is no agreement on the prior information, or even when there is a total lack of it, we can use what are known as noninformative or reference priors; that is, the prior distribution $\pi(\theta)$ will be chosen to reflect our state of ignorance. Inference under these circumstances is known as objective Bayesian inference. The values of the parameters in the prior distributions must in turn be specified, or else they must be assumed to follow a distribution. The latter will be done specifying a distribution that reflects little or no information (Zellner 1971a). This is easily done in WinBUGS (Scollnik 2001).

5. Estimating the Reserves

We want to estimate (or obtain the distribution of) aggregate claims for the i-th accident year, $i = 2, \ldots, k$, given the information available in the development triangle. Recall that the cumulative claims are given by for $W_{ij} = \Sigma_{t=1}^{j} Z_{it}$ for $1 \leq j \leq k$. Hence, in the run-off triangle setup, we are really interested in estimating W_{ik} $i = 2, \ldots, k$, given Z_{it}, $i = 1, \ldots, k$ $t = 1, \ldots, k$, with $i + t \leq k + 1$. Now let $R_i = W_{ik} - W_{i,k-i+1}$, for $i = 2, \ldots, k$, where $W_{i,k-i+1}$ is the accumulation of Z_{it} up to the latest development period and R_i equals the total aggregate outstanding claims corresponding to business year i, for $i = 2, \ldots, k$. From the distribution of R_i the required reserves corresponding to this business year can be obtained by choosing the measure of central tendency or quantile desired in the distribution of outstanding claims. Finally, from the distribution of total aggregate outstanding claims $R = \Sigma_{i=2}^{k} R_i$ the required total reserves can be computed.

As mentioned in Section 2, in the Bayesian approach, when interest is on prediction, as in loss reserving, the past (known) data in the upper triangle, z, are used to predict the observations in the lower triangle by means of the posterior predictive distribution for outstanding claims in each cell:

$$f(z_{it}|z) = \int f(z_{it}|\theta, \sigma^2, \delta) f(\theta, \sigma^2, \delta|z) d\theta d\sigma^2 d\delta,$$

$$i = 1, \ldots, k, \quad t = 1, \ldots, k, \quad \text{with } i + t > k + 1.$$

The reserves for the outstanding aggregate claims are estimated as the mean of the predictive distribution. Hence for each cell $E[Z_{it}|z]$ must be obtained. Then the Bayesian estimate of the total outstanding claims for year of business i is $\Sigma_{t>k-i+1} E[Z_{it}|z]$. The Bayesian "estimator" of the variance of outstanding claims (the predictive variance) for that same year is too cumbersome to derive analytically. This is where MCMC proves to be very useful: it yields random observations for aggregate claims in each cell of the (unobserved) lower right triangle $z_{it}^{(j)}$, $i = 2, \ldots, k$, $t > k - i + 1$, for $j = 1, \ldots, N$. The resulting (predictive) values will include both parameter variability and process variability. These predictive values can be used to compute the mean and variance of the reserves. First let us compute a random value of the total outstanding claims $R^{(j)} = \Sigma_{i,t} Z_{it}^{(j)}$, for $j = 1, \ldots, N$. Then the mean and variance of total outstanding claims can be computed as

$$\sigma_R^2 = \sum_{j=1}^{N} \frac{(R^{(j)} - \bar{R})^2}{N} \qquad \text{and} \qquad \bar{R} = \frac{1}{N} \sum_{j=1}^{N} R^{(j)}.$$

The standard deviation σ_R thus obtained is an "estimate" for the prediction error of the claims to be paid. The simulation process has the added advantage that it is not necessary to obtain explicitly the covariances that may exist between parameters; they are dealt with implicitly. It is also possible to obtain directly the median and any other quantile for the predictive distribution of the outstanding claims directly. This is illustrated in the next section with an application.

6. Application

This section presents a set of data that contains many negative values. Table 2 presents data provided by Prof. R. L. Brown and kindly made available by an (anonymous) American insurance company.

The results of applying the following methods are compared: the chain ladder, using Mack's model to obtain bootstrap variance estimates; the Bayesian method, using a "profile" likelihood with δ

Table 2

General Insurance Paid Claims from an Anonymous Insurance Company

	1	2	3	4	5	6	7	8	9
1	33,250.717	2,097.059	78.897	21.117	−18.65	−0.121	−5.072	−1.292	−0.78
2	36,717.578	2,583.632	−34.240	19.080	10.120	−3.699	−2.492	1.259	
3	38,155.786	2,705.212	38.503	−0.247	6.442	−6.669	−9.525		
4	36,180.233	2,601.743	21.501	−8.662	−6.250	12.87			
5	35,980.821	2,892.427	52.478	10.982	−3.496				
6	37,518.185	2,901.650	−23.61	−39.496					
7	40,213.152	3,006.438	−14.59						
8	39,105.807	3,080.126							
9	41,184.755								

replaced with its ML estimate, mentioned in Section 3; and the Bayesian model presented above with MCMC simulation.

Appendix 1 provides details of how the hierarchical Bayesian model was specified. The parameters in the third stage were chosen to reflect lack of information. Those for σ^2_μ, $\sigma^2_{\alpha_i}$, $\sigma^2_{\beta_t}$, and σ^2_δ are specified as in Spiegelhalter et al. (2003) and Verrall (2004). The priors for ν and λ are specified so that the prior expected value of σ^2 corresponds approximately to the value that results when applying OLS to the data after correcting by adding the MLE of δ, $\hat{\delta} = 300.2$, as in de Alba (2002a). Finally, the distribution for δ was set so that a priori δ has a large variance, and hence it allows the random values generated in the MCMC simulation to cover a broad range. Its prior expected value is chosen to be between zero and its MLE $\hat{\delta}$. The WinBUGS code is given in Appendix 2. An initial burn-in sample of 10,000 iterations was used. Two parallel chains were generated, each with a burn-in sample of 5,000. The results of these observations were discarded, to remove any effect from the initial conditions and allow the simulations to converge. The results were examined using a number of different initial conditions to ensure that these had no effect on the results. Then a further 50,000 simulations for each of the two chains were run and then ''thinned out'' to one out of 10 observations to reduce autocorrelations and the results shown below obtained. Various checks were made of the convergence of the Markov chain using BOA (Smith 2005). The specific criteria used were the Gelman-Rubin convergence statistic, the Raftery-Lewis convergence diagnostic, and the Geweke convergence diagnostic, as well as visual inspection of the sampled values.

Some characteristics of the posterior distribution of the parameters δ, μ, and the precision $\tau = 1/\sigma^2$ are shown in Table 3. The posterior mean of δ is 144.5, while its MLE is $\hat{\delta} = 300.2$. Notice that the MLE is out on the right tail of the posterior distribution and beyond the 97.5% quantile. Other options for the prior distribution of δ were also analyzed, and it was always the case that the MLE was out in the tail of the posterior distribution. Thus if we had this information before trying to apply the Bayesian method with the profile likelihood, this would be a very unlikely value to use to substitute in the likelihood.

Specifying a large prior variance for δ allows the parameter to ''adjust'' to a value that combines prior and sample information, so the posterior mean of δ is smaller than its MLE. In fact, it is inter-

Table 3

MCMC Estimates of Parameters for the Log-Normal Distribution

Parameter	Mean	Std Dev.	Percentiles		
			2.50%	Median	97.50%
δ	144.50	41.89	76.53	140.20	235.60
μ	10.53	0.06231	10.39	10.53	10.65
$\tau = 1/\sigma^2$	44.63	23.94	11.91	40.22	102.70

esting to note that the posterior mean of δ lies between its MLE and the first-order statistic of the sample, since Cohen and Whitten (1980) proposed a ''modified method of moments'' for estimating the threshold parameter. The modification to the method of moments uses a function of the first-order statistic of the sample to obtain better estimates. So it may be argued that the method presented here combines the MLE and the first-order statistics, although in some unknown way, and is consistent with other existing methods.

Table 4 shows the results of applying the methods that produce estimates when there are negative values present: the chain-ladder method with bootstrap estimates of the prediction error, the Bayesian method using a ''profile'' likelihood with δ replaced with its ML estimate, and the Bayesian model presented above with MCMC simulation.

The chain-ladder development factors can be computed without any problem and applied to obtain the reserves, since all the cumulative claim values are positive. The first column in Table 4 shows the chain-ladder reserve estimates. These values are the same as would be obtained using Mack's model because this model reproduces the values of the reserves from the chain ladder. The difference between the two methods is that one is stochastic and the other deterministic. These estimates will serve as basis for comparison since this technique has been used as a gold standard. Notice that the reserves are negative for all but the last two accident years, a result of the fact that three of the development factors turn out to be less than one.

The prediction error for the reserves under Mack's model were obtained with bootstrapping and computed using spreadsheets developed for this purpose by England and Verrall (2002) and kindly made available by the authors. These prediction errors can be compared to the standard deviation of the predictive distribution in the Bayesian models. Notice that the bootstrap errors are smaller than in both Bayesian models, probably because Bayesian models have more sources of variation, they involve parameters, and all of them are allowed to vary. Mack's model is ''distribution-free'' and thus involves no parameters. This is a restriction since it does not provide a distribution, and to produce prediction intervals, normality is assumed (England 2002).

In this example the reserves in both Bayesian methods are not very close, as can be seen in Table 4. In this case the estimates of the total reserves obtained with the Bayesian method presented in this paper are lower than the others. The variance of the MCMC estimates is larger than those from the direct simulation with the profile likelihood for accident years 2 and 9 and for the total. In the other accident years they are smaller. Apparently, allowing δ to vary does not have a systematic effect on the variances of the reserves. But in any case these estimates are more realistic than when fixing δ equal to its MLE, because its sampling variability is explicitly taken into account. In the Bayesian model this is done through the variance of the prior distribution of δ. In this particular model the noninformative priors used in the third stage of the hierarchical model seem to allow a better fit to the data and hence result in a smaller variance of the predictive distribution.

Table 4

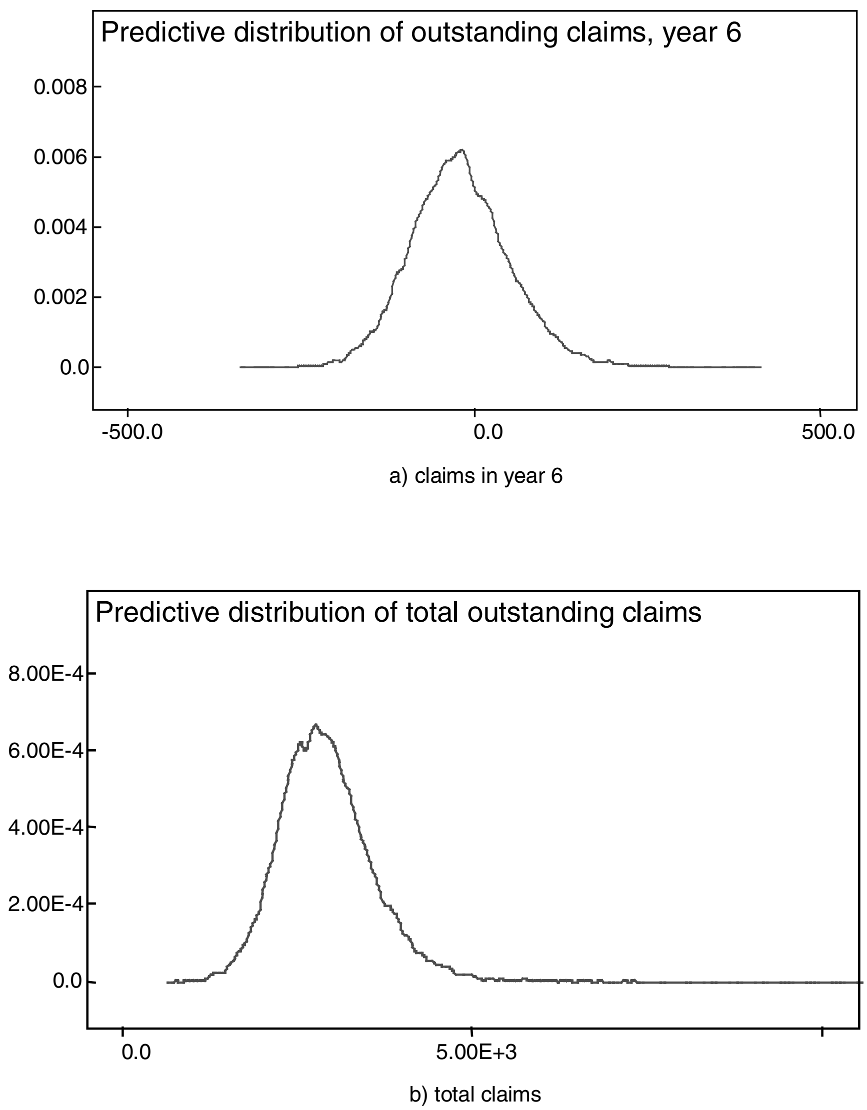
Estimation of Reserves by Accident Year Using Different Methods: Chain Ladder with Bootstrap Prediction Errors, Bayesian Method with Profile Likelihood, and Bayesian Model with MCMC

Accident Year	Chain Ladder with (Mack's) Bootstrap Prediction Errors		Bayesian Method with Profile Likelihood		Bayesian Method Using MCMC	
	Reserves	Pred. Error	Reserves	Std Dev.	Reserves	Std Dev.
2	−0.86	2.46	1.38	32.25	2.07	34.01
3	−0.91	3.55	21.35	47.19	11.10	46.97
4	−6.60	5.18	5.81	57.88	4.98	55.06
5	−6.02	10.69	64.14	73.00	22.24	65.06
6	−8.72	15.79	−54.89	81.64	−19.82	71.11
7	−8.82	30.14	77.69	107.55	9.43	80.04
8	9.51	54.31	244.08	148.15	44.07	93.80
9	3,041.18	235.66	3,363.20	514.50	2,851.00	616.20
Total	3,018.77	248.07	3,722.70	620.50	2,925.00	678.80

Notice that the reserves for accident year 6 are negative under all methods, so the complete distribution will be plotted to see the behavior of the outstanding claims for this year. Figure 1 shows the predictive distribution of the reserves for accident year 6 (panel a) and for the total (panel b). The distribution for reserves corresponding to accident year 6 has a long right tail and yet shows a large probability of negative values. In fact, the predictive mean is close to zero; this explains why the reserves are negative. The results for accident years 2 and 4 (not shown) also yield medians very close to zero, but even so the predictive mean is positive. The predictive distribution of total reserves is symmetric but is clearly not normal. The 95% predictive interval obtained from the MCMC results is given by (1767,4431). There is a large difference in total reserves between the Bayesian method using the profile likelihood substituting the MLE of δ and the predictive mean of total claims from the model presented in this paper. The means of total reserves are fairly close for the chain ladder and the Bayesian model presented here: 3,018.77 and 2,925.0, respectively. So in this respect the MCMC Bayesian results are close to the "gold standard." Nevertheless the reserves for each year of origin are very different between

Figure 1

Predictive Distribution for Outstanding Claims in Accident Year 6 and the Total, Obtained Using MCMC with 10,000 Simulations



the two methods. However, the complete posterior predictive distributions, and hence their quantiles, are available from the MCMC for the Bayesian method for each year, and further analysis can be carried out if desired. On the other hand, Mack's method would require some additional distributional assumptions.

7. CONCLUDING REMARKS

The Bayesian method presented here constitutes an appealing alternative to claims-reserving methods in the presence of negative values in incremental claims for some cells of the development triangle. It yields results comparable to those of other methods but does not have their limitations. Furthermore, the model is based on fairly standard and widely used assumptions. However, the main advantage is that this method will produce the predictive distribution of outstanding claims in the presence of enough negative values (in number and/or size), even if these make the total sum of one or more of the columns in the development triangle of incremental claims be negative, and so most other methods cannot produce estimates.

Concluding, if there are many negative incremental claims in the development triangle and one wants to obtain the distribution of outstanding claims, then the Bayesian model here is the solution, because Mack's method, which reproduces the chain ladder, yields the mean and variance, but not a distribution, and normality must be assumed to obtain predictive intervals (England 2002). Furthermore, the Bayesian method presented in this paper allows for the variability of all the parameters involved so it is more flexible, but it is also consistent with other methods. The method is not difficult to apply in practice, but it does require the use of an MCMC program and understanding how to specify the prior distributions.

APPENDIX 1

IMPLEMENTING MCMC IN WINBUGS

To implement the simulation using Markov chain Monte Carlo, the package WinBUGS 1.4 is used (Spiegelhalter et al. 2001). It is necessary to specify each of the stages in the hierarchical model. According to the specification of the model in Section 3, in the first stage of the hierarchical model the likelihood function is specified as

$$Y_{it} = \log(Z_{it} + \delta) \sim N(\mu_{it}, \sigma^2)$$
$$\mu_{it} = \mu + \alpha_i + \beta_t$$
$$\alpha_1 = \beta_1 = 0$$
$$i = 1, \ldots, k, \; t = 1, \ldots, k, \; i + t \leq k + 1$$
$$k = \text{number of accident years} = \text{number of development years} = 9.$$

In the second stage of the hierarchical model the prior distribution is specified for the unknown parameters. In the present model:

$$\mu \sim N(0, \sigma^2_{\mu})$$
$$\alpha_i \sim N(0, \sigma^2_{\alpha_i})$$
$$\beta_t \sim N(0, \sigma^2_{\beta_t})$$
$$\sigma^2 \sim IG(\nu, \lambda)$$
$$\delta \sim N(\mu_{\delta}, \sigma^2_{\delta}),$$

where $IG(\nu, \lambda)$ denotes an inverse Gamma distribution with parameters ν and λ. The prior distributions for μ, α_i, and β_t have been chosen as to allow the simulation process to converge to any possible value, positive or negative. This will be further enhanced by the choice of their variances or the distribution of their variances. If these are allowed to be large, the prior distributions will be of the noninformative kind. In the prior distribution for σ^2 we choose each of the parameters (ν, λ) to follow a distribution such that the prior one reflects ignorance about the possible value of σ^2 (Gelman et al. 2004). This is done at the third stage of the hierarchical model. In the second stage the prior distribution for δ is also specified. Notice that the prior distribution for δ will generate some values that will lead to $(z_{it} + \delta) < 0$, but these are excluded from the likelihood by definition, as implied by the indicator function in formula (3.3), and hence will have no effect on the posterior distribution.

In the third stage of the hierarchical model the distributions are specified for some of the (hyper) parameters in the distributions of stage two (Klugman 1992). These are usually given values that reflect lack of information (Scollnik 2001; Zellner 1971a). Thus

$$\sigma^2_\mu \sim IG(0.1, 0.1)$$
$$\sigma^2_{\alpha_i} \sim IG(0.001, 0.001)$$
$$\sigma^2_{\beta_j} \sim IG(0.001, 0.001)$$
$$\nu \sim G(2.5, 0.1)$$
$$\lambda \sim G(2, 0.1),$$

where $G(a, b)$ denotes a Gamma distribution with parameters a and b and mean a/b, and $IG(c, d)$ denotes an inverse Gamma distribution with parameters b and c. A distribution is also specified at a third stage for the parameters μ_δ and σ^2_δ in the prior distribution for δ: $\mu_\delta \sim N(200, 10000)$ and $\sigma^2_\delta \sim IG(0.0001, 0.1)$. Note that in WinBUGS the specification is in terms of precision $\tau = 1/\sigma^2$, and if $\sigma^2 \sim IG(a, b)$, then $\tau \sim G(a, b)$.

The "hyper-parameters" in the distributions of ν and λ were specified so that these parameters will have means 25 and 20, respectively, and both with large variances, so as to reflect lack of prior information. This, in turn, will result in a large variance on σ^2. The prior distribution on μ_δ is chosen to have a mean smaller than the MLE of δ and a variance of 10,000 (precision 0.0001).

APPENDIX 2

WINBUGS CODE

model

```
{
 # Model for the observations (upper left triangle)
 C <- 10000 # this constant must be large to guarantee that p[i]'s < 1
 for(i in 1 : r) {
  for(j in 1 : (r+1-i)) {
   ones[i , j] <- 1
   pp[i, j] <- ((sqrt(tau)/(x[i, j]+delta))*exp(-0.5 *pow(log(abs(x[i, j]+delta)) - mu[i, j], 2) * tau))/ C
   p[i, j] <- max(pp[i,j],0)
   cct[i,j] <- abs(p[i, j]/pp[i, j])
   ones[i , j] ~ dbern(p[i , j])
   mu[i , j] <- miu + alpha[i] + beta[j]
  }

  # Model for future claims (lower right triangle)
  for(j in (r+2-i) : r) {
  x[i, j] ~ dlnorm(mu[i, j], tau)
  mu[i, j] <- miu + alpha[i] + beta[j]
  }
 }

 alpha[1] <- 0
 beta[1] <- 0
 for(i in 2 : r) {
  alpha[i] ~ dnorm(0.0 , alpha.tau)
  beta[i] ~ dnorm(0.0 , beta.tau)
 }

 alpha.tau ~ dgamma(0.001 , 0.001)
 beta.tau ~ dgamma(0.001 , 0.001)
```

```
  miu ~ dnorm(0.0 , miu.tau)
  tau ~ dgamma(tau.a , tau.b)
  sigma2 <- 1 / tau

  miu.tau ~ dgamma(0.1 , 0.1)

  tau.a ~ dgamma(2.5 , 0.1)
  tau.b ~ dgamma(2 , 0.1)


  delta ~ dnorm(delta.a , delta.c)
  delta.a ~ dnorm(200.0, .0001)
  delta.c ~ dgamma(0.0001 , .1)
  # Estimation of the reserves.

  for(i in 2 : r) {
  outstand.row[i] <- sum(x[i, (r +2 - i) : r]) - (i - 1)*delta
  }
  outstand.row[10] <- sum(outstand.row[2 : 9])
 }
```

data

```
list(r = 9 ,
  x = structure(
   .Data = c(33250.72, 2097.06, 78.90, 21.12,-18.65,-0.12,-5.07,-1.29,-0.78,
  36717.58, 2583.63,-34.24, 19.08, 10.12,-3.70,-2.49, 1.26, NA,
  38155.79, 2705.21, 38.50,-0.25, 6.44,-6.67,-9.52, NA, NA,
  36180.23, 2601.74, 21.50,-8.66,-6.25, 12.87, NA, NA, NA,
  35980.82, 2892.43, 52.48, 10.98,-3.50, NA, NA, NA, NA,
  37518.19, 2901.65,-23.61,-39.50, NA , NA, NA, NA, NA,
  40213.15, 3006.44,-14.59, NA, NA, NA, NA, NA, NA,
  39105.81, 3080.13, NA, NA, NA, NA, NA, NA, NA,
  41184.76, NA, NA, NA , NA , NA , NA, NA , NA),
   .Dim = c(9, 9)))
```

initial values

```
list(
  alpha = c(NA, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1),
  alpha.tau = 1,
  beta = c(NA, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1),
  beta.tau = 1,
  delta = 65,
  delta.c =1,
  miu = 0,
  miu.tau = 1,
  tau = 201,
  tau.a =2001,
  tau.b = 9
)
list(
  alpha = c(NA, 1, 1, 1, 1, 1, 1, 1, 1),
  alpha.tau = 2,
  beta = c(NA, 1, 1, 1, 1, 1, 1, 1, 1),
  beta.tau = 2,
  delta = 60,
  delta.c =3,
  miu = 1,
  miu.tau = 2,
  tau = 100,
  tau.a =1001,
  tau.b = 20
)
```

ACKNOWLEDGEMENTS

This research was supported by Asociación Mexicana de Cultura, A.C. in Mexico City. The author wishes to acknowledge computational support from Marco A. Ramirez-Corzo, from ITAM.

REFERENCES

ANDERSON, DUNCAN, SHOLOM FELDBLUM, CLAUDINE MODLIN, DORIS SCHIRMACHER, ERNESTO SCHIRMACHER, AND NEEZA THANDI. 2004. A Practitioner's Guide to Generalized Linear Models. 2004 Discussion Paper Program, Casualty Actuarial Society, Arlington, VA.

BERGER, JAMES O. 1985. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer.

BERNARDO, JOSE M., AND ADRIAN F. M. SMITH. 1994. *Bayesian Theory*. New York: John Wiley & Sons.

BROWN, ROBERT L., AND LEON R. GOTTLIEB. 2001. *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*. 2nd ed. Providence, RI: ACTEX Publications.

CHAMBERLIN, GARY F., ED. 1989. *Claims Reserving Manual*, *Vol. 1*. London: Institute of Actuaries.

COHEN, A. CLIFFORD, AND BETTY JONES WHITTEN. 1980. Estimation in the Three Parameter Log-Normal Distribution. *Journal of the American Statistical Association* 75: 370, 399–404.

CROW, EDWIN L., AND KUNIO SHIMIZU. 1988. *Lognormal Distributions*: *Theory and Applications*. New York: Marcel Dekker.

DE ALBA, ENRIQUE. 2002a. Claims Reserving When There Are Negative Values in the Runoff Triangle. 37th Actuarial Research Conference, University of Waterloo, ARCH 2003 Proceedings. Schaumburg, IL: Society of Actuaries.

———. 2002b. Bayesian Estimation of Outstanding Claim Reserves. *North American Actuarial Journal* 6(4): 1–20.

———. 2004, Bayesian Claims Reserving. In *Encyclopedia of Actuarial Science*. London: John Wiley and Sons.

DE ALBA, ENRIQUE, AND ROBERTO BONILLA. 2002. Un Modelo para el Tratamiento de Valores Negativos en el Triángulo de Desarrollo Utilizado en la Estimación de Reservas para SONR. *In Transactions of the 27th International Congress of Actuaries,* Cancún, México.

ENGLAND, PETER. 2002. Addendum to Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving. *Insurance: Mathematics and Economics* 31: 461–66.

ENGLAND, PETER, AND RICHARD J. VERRALL. 2002. Stochastic Claims Reserving in General Insurance (with discussion). British Actuarial Journal 8: 443–544.

GELMAN, ANDREW, JOHN B. CARLIN, HAL S. STERN, AND DONALD B. RUBIN. 2004. *Bayesian Data Analysis*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

HAASTRUP, SVEND, AND ELJA ARJAS. 1996. Claims Reserving in Continuous Time: A Nonparametric Bayesian Approach. ASTIN Bulletin 26(2): 139–64.

HESS, KLAUS T., AND KLAUS D. SCHMIDT. 2002. A Comparison of Models for the Chain-Ladder Method. *Insurance: Mathematics and Economics* 31: 351–64.

HILL, BRUCE M. 1963. The Three-Parameter Lognormal Distribution and Bayesian Analysis of a Point-Source Epidemic. *Journal of the American Statistical Association* 58: 72–84.

JEWELL, WILLIAM S. 1989. Predicting IBNYR Events and Delays. I. Continuous Time. *ASTIN Bulletin* 19(1): 25–56.

———. 1990. Predicting IBNYR Events and Delays. II. Discrete Time. *ASTIN Bulletin* 20(1): 93–111.

JOHNSON, NORMAN L., SAMUEL KOTZ, AND N. BALAKRISHNAN. 1994. *Continuous Univariate Distributions in Statistics, Vol. I*. New York: John Wiley.

KLUGMAN, STUART A. 1992. *Bayesian Statistics in Actuarial Science*. Boston: Kluwer.

KREMER, ERHARD. 1982. IBNR Claims and the Two-Way Model of ANOVA. *Scandinavian Actuarial Journal*, 47–55.

KUNKLER, MICHAEL. 2004. Modelling Zeros in Stochastic Reserving Models. *Insurance: Mathematics and Economics* 34: 33–35.

MACK, THOMAS. 1993. Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Structure. *ASTIN Bulletin* 23: 213–25.

———. 2004. Chain-Ladder Method. In *Encyclopedia of Actuarial Science*. London: John Wiley and Sons.

MAKOV, UDI E. 2001. Principal Applications of Bayesian Methods in Actuarial Science: A Perspective. *North American Actuarial Journal* 5(4): 53–73.

MAKOV, UDI E., ADRIAN F. M. SMITH, AND Y.-H. LIU. 1996. Bayesian Methods in Actuarial Science. *The Statistician* 45(4): 503–15.

NTZOUFRAS, IOANNIS, AND PETROS DELLAPORTAS. 2002. Bayesian Modeling of Outstanding Liabilities Incorporating Claim Count Uncertainty. *North American Actuarial Journal* 6(1): 113–36.

RENSHAW, A. E., AND RICHARD J. VERRALL. 1998. A Stochastic Model Underlying the Chain-Ladder Technique. *British Actuarial Journal* 4(6): 905–23.

SCOLLNIK, DAVID P. M. 2001. Actuarial Modeling with MCMC and BUGS. *North American Actuarial Journal* 5(2): 96–124.

———. 2002. Implementation of Four Models for Outstanding Liabilities in WinBUGS: A Discussion of a Paper by Ntzoufras and Dellaportas. *North American Actuarial Journal* 6(1): 128–36.

SMITH, BRIAN. 2005. Bayesian Output Analysis Program (BOA). Version 1.1.5 for R. www.public-health.uiowa.edu/boa.

SPIEGELHALTER, DAVID J., ANDREW THOMAS, NICOLA G. BEST, AND WALTER R. GILKS. 2001. Win*BUGS* 1.4: *Bayesian Inference Using Gibbs Sampling*. Imperial College and MRC Biostatistics Unit, Cambridge, UK. www.mrc-bsu.cam.ac.uk/bugs.

SPIEGELHALTER, DAVID J., ANDREW THOMAS, NICOLA G. BEST, WALTER R. GILKS, AND DAVID J. LUNN. 2003. BUGS: Bayesian Inference Using Gibbs Sampling. MRC Biostatistics Unit, Cambridge, England. www.mrc-bsu.cam.ac.uk/bugs.

TAYLOR, GREGORY C. 2000. *Claims Reserving: An Actuarial Perspective.* New York: Elsevier.

THOMAS, ANDREW. 2005. OpenBUGS. http://mathstat.helsinki.fi/openbugs.

VERRALL, RICHARD J., ED. 1989. *Claims Reserving Manual*, *Vol. 2*. London: Institute of Actuaries.

———. 1990. Bayes and Empirical Bayes Estimation for the Chain Ladder Model. *ASTIN Bulletin* 20(2): 217–43.

———. 1991. Negative Incremental Claims: Chain Ladders and Their Models. *Journal of the Institute of Actuaries* 120(1): 171–83.

———. 2000. An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique. *Insurance: Mathematics and Economics* 26: 91–99.

———. 2004. A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving. *North American Actuarial Journal* 8(3): 67–89.

ZELLNER, ARNOLD. 1971a. *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley.

———. 1971b. Bayesian and Non-Bayesian Analysis of the Log-Normal Distribution and Log-Normal Regression. *Journal of the American Statistical Association* 66: 327–30.

Additional discussions on this paper can be submitted until January 1, 2007. The author reserves the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.