ROBUST BAYESIAN ANALYSIS OF LOSS RESERVES DATA USING THE GENERALIZED-$t$ DISTRIBUTION

BY

JENNIFER S.K. CHAN, S.T. BORIS CHoy AND UDI E. MAKOV

ABSTRACT

This paper presents a Bayesian approach using Markov chain Monte Carlo methods and the generalized-$t$ (GT) distribution to predict loss reserves for the insurance companies. Existing models and methods cannot cope with irregular and extreme claims and hence do not offer an accurate prediction of loss reserves. To develop a more robust model for irregular claims, this paper extends the conventional normal error distribution to the GT distribution which nests several heavy-tailed distributions including the Student-$t$ and exponential power distributions. It is shown that the GT distribution can be expressed as a scale mixture of uniforms (SMU) distribution which facilitates model implementation and detection of outliers by using mixing parameters. Different models for the mean function, including the log-ANOVA, log-ANCOVA, state space and threshold models, are adopted to analyze real loss reserves data. Finally, the best model is selected according to the deviance information criterion (DIC).

KEYWORDS

Bayesian approach, state space model, threshold model, scale mixtures of uniform distribution, deviance information criterion.

1. INTRODUCTION

An insurance policy is a promise of an insurance company to pay claims to the insureds if some defined events (death, accident, injury, etc.) occur. However in many cases, claims originating in a particular year are often not settled in that year, but with a time delay of years or perhaps decades. Therefore, the insurance company must have the necessary loss reserves to pay these outstanding claims and settlement costs incurred. With many uncertainties in the time lags inherently involved in the claims settlement process, the estimation procedure of the required loss reserves is extremely complicated. Since loss reserves generally represent by far the largest liability and the greatest source of financial uncertainty in an insurance company, accurate prediction of the
loss reserves is of great importance. Failure to estimate the loss reserves accurately may result in large profit losses and hence weaken the financial stability of the company which may ultimately drive it into insolvency.

Denote $Y_{i,j}$, $i = 1, \ldots, n$, the value of claims paid by an insurance company for accidents that occurred in policy year $i$ and were settled after $j - 1$ years (lag year $j$). The observed claim amounts $Y_{i,j}$, $i = 1, \ldots, n$, $j \leq n - i + 1$ over a period of $n$ policy years can be presented by a run-off triangle; see Table 1. The total number of observed claims in the upper triangle is $S_U = n(n + 1)/2$ and the number of unobserved claims to be predicted in the lower triangle is $S_L = n(n - 1)/2$. The aim is to predict the unobserved values in the lower triangle using an appropriate statistical model.

The most popular method for prediction is the chain-ladder method (Renshaw, 1989) which uses loss ratio estimates and loss development factors. However it is increasingly apparent that this method lacks some measures of variability. Over the years, stochastic models with hierarchical structure have been developed to overcome such shortcomings. For example, Verrall (1991, 1996) and Renshaw and Verrall (1998) adopt the ANOVA-type and ANCOVA-type models. The interaction between the policy year and lag year prompted the development of the state space models which allow dynamic evolution of parameters in a time-recursive way. See Verrall (1989, 1994). In addition, Hazan and Makov (2001) propose the threshold models to allow for structural changes in the trend of outstanding claims over time. Analyses of real data suggest that the dynamic nature of the state space model and threshold model improves the prediction.

It is well known that the normal error distribution falls short of allowing for irregular and extreme claims and hence contaminates the estimation procedure and leads to poor prediction. To allow for irregular claims, error distributions that possess flexible tails are recommended. The Student-$t$ distribution has been widely used by many researchers for this purpose while Choy and Chan (2003) use the exponential power (EP) family of distributions. In this

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>\ldots</th>
<th>n-1</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$Y_{1,1}$</td>
<td>$Y_{1,2}$</td>
<td>$Y_{1,3}$</td>
<td>\ldots</td>
<td>$Y_{1,n-1}$</td>
<td>$Y_{1,n}$</td>
</tr>
<tr>
<td>2</td>
<td>$Y_{2,1}$</td>
<td>$Y_{2,2}$</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>$Y_{2,n}$</td>
</tr>
<tr>
<td>3</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>n-1</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>n</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
</tbody>
</table>

Table 1: Run-off triangle for claim data.
paper, we shall adopt the generalized-\( t \) (GT) distribution (McDonald and Newey, 1988) for the errors. The GT distribution is symmetric and is governed by two shape parameters. By suitably choosing these shape parameter values, one can easily show that the normal, Cauchy, Student-\( t \), EP and Laplace distributions are members of the GT family.

For a long time the Bayesian approach had very limited applications in statistical inference because the posterior functionals that involve high-dimensional integration cannot be obtained analytically. Gelfand and Smith (1990) develop the simulated-based Markov chain Monte Carlo (MCMC) techniques that transform the integration problem into a sampling problem. The idea is to construct a Markov chain whose limiting distribution is the intractable joint posterior distribution. The simulated realizations from the Markov chain mimic a random sample from the joint posterior distribution and the posterior functionals can be estimated from these realizations. Amongst the MCMC algorithms, Gibbs sampling (Geman and Geman, 1984), Metropolis-Hastings (Metropolis et al., 1953 and Hastings, 1970) and Reversible Jump MCMC (Green, 1995) are very common. Bayesian statistical inferences can be easily done using WinBUGS (Spiegelhalter et al., 2004), a user-friendly software for MCMC algorithms. Readers who are interested in or unfamiliar to WinBUGS may find the educational materials on the official WinBUGS website very useful. See www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml. They are referred to Gilks et al. (1998) and Gelman et al. (2004) for MCMC applications in many statistical problems. For actuarial and insurance applications, see, for example, Makov (2001), De Alba (2002), Ntzoufras and Dellaportas (2002), Ntzoufras et al. (2005), Scollnik (1998, 2001, 2002) and Verrall (2004).

This paper makes the first attempt to adopt the GT error distribution via a scale mixture of uniforms (SMU) form for model implementation and for protecting inferences from the distorting effect of the possible outliers. The proposed data augmentation scheme makes the Gibbs sampler easier to set up and is able to detect possible outliers using the mixing parameters of the SMU representation.

The rest of this paper is structured as follows. Section 2 introduces the data that demonstrate the proposed loss reserves models. Section 3 reviews the models with normal error distribution. Section 4 outlines the GT distribution and its scale mixtures representations. Section 5 implements the models using the Bayesian approach and proposes the DIC for choosing the best model. Section 6 reports and compares the results. Finally, discussion of results and further research direction is outlined in Section 7.

2. THE DATA

To demonstrate the models proposed in Sections 2 and 3, a loss reserves data set as shown in Table 2 is analyzed. The data are the amount of claims paid to the insureds of an insurance company during the period of 1978 to 1995.
over \( n = 18 \) years. The upper triangle has \( N = 171 \) observations and the 153 observations in lower triangle are not yet observed. For mathematical convenience, two zero claim amounts are replaced by 0.01. Some general trends are obvious in the data. Given a policy year, the amount of claims paid follows an increasing trend in the first 4 to 6 lag years and then a decreasing trend thereafter. On the other hand, there are no obvious trends across policy years for each lag year. Although the data are not adjusted for inflation, our analysis is able to capture other sources of variability.

This data set contains some extreme outliers which are underlined in Table 2. For example, extremely large claims (in italic), amount to 11,920 and 15,546 dollars, were made in the 7-th lag year of policy year 1984 and in the 4-th lag year of policy year 1992, respectively. These outliers distort the general trends in the data and inflate the standard errors of the model parameters. Other outliers have much smaller amount of claims than their neighboring claims. These claims may lead to underestimates of loss reserves and hence lower the solvency and increase the risk of bankruptcy for an insurance company. Hence for robustness consideration, heavy-tailed error distributions are adopted to accommodate these irregular claims.

3. Modeling the Mean Function

The popular log-linear model was investigated by Renshaw (1989), Verrall (1991, 1996) and Renshaw and Verrall (1998) in actuarial context. Depending on the formulation of policy year and lag year effects, we consider the log-ANOVA, log-ANCOVA and state space models for the mean function. For the state space model, the interaction effects between the policy-year and the lag-year are assumed. Each of these models is then allowed to switch to the

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>3323</td>
<td>8332</td>
<td>9672</td>
<td>10172</td>
<td>7631</td>
<td>3855</td>
<td>3252</td>
<td>4433</td>
<td>2188</td>
<td>333</td>
<td>199</td>
<td>692</td>
<td>311</td>
<td>0.01</td>
<td>405</td>
<td>253</td>
<td>76</td>
<td>14</td>
</tr>
<tr>
<td>1979</td>
<td>3795</td>
<td>10342</td>
<td>8330</td>
<td>7849</td>
<td>2839</td>
<td>3577</td>
<td>1404</td>
<td>1721</td>
<td>1066</td>
<td>156</td>
<td>35</td>
<td>259</td>
<td>256</td>
<td>420</td>
<td>5</td>
<td>1</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>4577</td>
<td>9889</td>
<td>8746</td>
<td>10228</td>
<td>8572</td>
<td>5787</td>
<td>3855</td>
<td>1445</td>
<td>1612</td>
<td>626</td>
<td>589</td>
<td>438</td>
<td>473</td>
<td>370</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>5288</td>
<td>8089</td>
<td>12839</td>
<td>11829</td>
<td>7560</td>
<td>6383</td>
<td>4118</td>
<td>3016</td>
<td>1575</td>
<td>1985</td>
<td>2646</td>
<td>266</td>
<td>38</td>
<td>45</td>
<td>115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>2294</td>
<td>9869</td>
<td>10242</td>
<td>13908</td>
<td>8775</td>
<td>5419</td>
<td>2424</td>
<td>1597</td>
<td>4149</td>
<td>1296</td>
<td>917</td>
<td>295</td>
<td>428</td>
<td>359</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>3600</td>
<td>7514</td>
<td>8247</td>
<td>9327</td>
<td>8584</td>
<td>4245</td>
<td>4096</td>
<td>3216</td>
<td>2014</td>
<td>593</td>
<td>1888</td>
<td>691</td>
<td>368</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>3642</td>
<td>7394</td>
<td>9838</td>
<td>9733</td>
<td>6377</td>
<td>4884</td>
<td>17230</td>
<td>4189</td>
<td>4492</td>
<td>1760</td>
<td>944</td>
<td>921</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>2463</td>
<td>5033</td>
<td>6980</td>
<td>7722</td>
<td>6702</td>
<td>7834</td>
<td>5579</td>
<td>3622</td>
<td>1300</td>
<td>3069</td>
<td>1370</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>2257</td>
<td>5959</td>
<td>6175</td>
<td>7051</td>
<td>8102</td>
<td>6339</td>
<td>6978</td>
<td>4396</td>
<td>3107</td>
<td>903</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>2009</td>
<td>3700</td>
<td>5299</td>
<td>6985</td>
<td>6477</td>
<td>7570</td>
<td>6855</td>
<td>5742</td>
<td>3971</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>1960</td>
<td>5282</td>
<td>3640</td>
<td>7538</td>
<td>5157</td>
<td>5766</td>
<td>6892</td>
<td>2572</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>2331</td>
<td>3517</td>
<td>5310</td>
<td>6066</td>
<td>10149</td>
<td>9268</td>
<td>5262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>2314</td>
<td>4487</td>
<td>4112</td>
<td>7000</td>
<td>11163</td>
<td>10057</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>2077</td>
<td>3652</td>
<td>8228</td>
<td>7896</td>
<td>9317</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>2595</td>
<td>5403</td>
<td>6579</td>
<td>15546</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>3155</td>
<td>4974</td>
<td>7961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>2626</td>
<td>5704</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>2827</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The amount of claims paid to the insureds of an insurance during 1978 to 1995.
model with a new set of parameters after certain thresholds. All models are implemented using WinBUGS with vague and non-informative prior distributions assigned to the model parameters.

3.1. ANOVA models

Within the Bayesian hierarchical modeling framework, the log-adjusted claim amount $y_{ij}$, written in year $i$ and paid with a delay $j - 1$ years, adopts a two-way ANOVA model as below:

$$\log(\gamma_{ij}) = \mu_{ij} + \varepsilon_{ij}$$ (1)

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$ (2)

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$ (3)

for $i = 1, \ldots, n$, $j \leq n - i + 1$ subject to the constraints $\sum_{i=1}^{n} \alpha_i = \sum_{j=1}^{n} \beta_j = 0$. $\mu_{ij}$ is the mean function, $\alpha_i$ is the policy year effects and $\beta_j$ is the lag year effects. To complete the Bayesian framework, we adopt the following priors for the model parameters:

$$\mu \sim N(0, \sigma^2_\mu), \ \alpha_i \sim N(0, \sigma^2_\alpha), \ \beta_j \sim N(0, \sigma^2_\beta), \ \sigma^2 \sim IG(a, b)$$

where $IG(a, b)$ is the inverse gamma distribution with density

$$f(\sigma^2) = \frac{b^a}{\Gamma(a)} \left( \frac{1}{\sigma^2} \right)^{a+1} \exp\left( -\frac{b}{\sigma^2} \right).$$

Diffuse priors can be obtained by setting the hyperparameters $\sigma^2_\mu = \sigma^2_\alpha = \sigma^2_\beta = \infty$ and $a = b = 0$, i.e. inflating the variances of the prior distributions to reflect the lack of prior information (Ntzoufras and Dellaportas, 2002). In this case, the joint prior density is

$$p(\mu, \alpha, \beta, \sigma^2) \propto \frac{1}{\sigma^2}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n)$.

3.2. ANCOVA model

In the ANCOVA model, the effects of policy year $i$ and lag year $j$ are linear. The three different combinations of effects are given below.

1. Linear effect of policy year and categorical effect of lag year,
2. Categorical effect of policy year and linear effect of lag year and
3. Linear effect of both policy year and lag year.
Preliminary result of the analysis reveals that the first combination provides the best fit to the data and hence is chosen for all subsequent analyses of the ANCOVA model. Analysis of the model follows from the log-ANOVA model from (1) to (3) except that the mean function becomes

$$\mu_{ij} = \mu + \alpha \cdot i + \beta_j$$

(4)

for \(i = 1, \ldots, n, j \leq n - i + 1\) subject to the constraints \(\sum_{j=1}^{n} \beta_j = 0\). The diffuse priors are chosen to be

$$p(\mu, \alpha, \beta, \sigma^2) \propto \frac{1}{\sigma^2}$$

where \(\beta = (\beta_1, \ldots, \beta_n)\).

### 3.3. State space model

To account for the interaction between the policy year and lag year in the mean function, the state space model (Ntzoufras and Dellaportas, 2002, De Jong and Zehnwirth, 1983 and Verrall, 1991, 1994) is considered. In the model, parameters are allowed to evolve in a time-recursive pattern: \(\alpha_i\) depends on \(\alpha_{i-1}\) and an error term \(h_i\) while \(\beta_{ij}\) depends on \(\beta_{i-1,j}\) and an error term \(v_i\). The model again follows from log-ANOVA model from (1) to (3) except that the mean function becomes

$$\mu_{ij} = \mu + \alpha_i + \beta_{ij}$$

(5)

for \(i = 1, \ldots, n, j \leq n - i + 1\) where the recursive associations are

$$\alpha_i = \alpha_{i-1} + h_i, \quad h_i \sim N(0, \sigma_h^2), \quad i = 2, 3, \ldots, n$$

(6)

$$\beta_{ij} = \beta_{i-1,j} + v_i, \quad v_i \sim N(0, \sigma_v^2), \quad i, j = 2, 3, \ldots, n$$

(7)

subject to the constraints \(\alpha_i = 0\) and \(\beta_{i1} = 0\) for \(i = 1, \ldots, n\). The priors for \(\sigma_h^2\) and \(\sigma_v^2\) are assumed to be \(\sigma_h^2 \sim IG(a_h, b_h)\) and \(\sigma_v^2 \sim IG(a_v, b_v)\) respectively. Diffuse prior distributions are chosen to be

$$p(\mu, \beta_1, \sigma^2, \sigma_h^2, \sigma_v^2) \propto \frac{1}{\sigma^2 \sigma_h^2 \sigma_v^2}$$

where \(\beta_1 = (\beta_{12}, \ldots, \beta_{1n})\).

### 3.4. Threshold model

Hazan and Makov (2001) suggested a switching regression model in which the mean functions before and after a threshold \(T\) along the axis of policy year are assigned the same model but with a different set of parameters. The reasoning
behind is that some events such as a financial crisis or a change in the insurance regulation may take place on or before certain threshold year $T$. These events may change the effects of some factors on the claim amounts and hence a new set of parameters is adopted to reveal such changes. The threshold models based on different mean function structures are given below:

Threshold log-ANOVA model:

\[ m_{ij} = \mu_1 + \alpha_{1i} + \beta_{1j} \quad \text{for} \quad i < T \]
\[ m_{ij} = \mu_2 + \alpha_{2i} + \beta_{2j} \quad \text{for} \quad i \geq T, \; j \leq n - T + 1 \]

Threshold log-ANCOVA model:

\[ m_{ij} = \mu_1 + \alpha_1 i + \beta_{1j} \quad \text{for} \quad i < T \]
\[ m_{ij} = \mu_2 + \alpha_2 i + \beta_{2j} \quad \text{for} \quad i \geq T, \; j \leq n - T + 1 \]

Threshold state space model:

\[ m_{ij} = \mu_1 + \alpha_{1i} + \beta_{1ij} \quad \text{for} \quad i < T \]
\[ m_{ij} = \mu_2 + \alpha_{2i} + \beta_{2ij} \quad \text{for} \quad i \geq T, \; j \leq n - T + 1 \]

The main difference in these threshold models is related to the set of $\beta$ parameters which change after the threshold policy year $T$. On the contrary, the criterion that the $\alpha$ parameters are different before and after the threshold also holds for the simple models. Prior distributions for the parameters in the three models remain the same as those in Sections 3.1 to 3.3. The threshold $T$ can be selected based on some model selection measures such as the Akaike information criterion ($AIC$) (Akaike, 1974), Bayesian information criterion ($BIC$) (Schwarz, 1978) and deviance information criterion ($DIC$) (Spiergelhalter et al., 2002). We shall use the $DIC$ in this paper because it is particularly useful in Bayesian model selection problems where the posterior distributions of the model parameters have been obtained by MCMC simulation. Threshold models with threshold lag-years, for example,

Threshold log-ANOVA model:

\[ m_{ij} = \mu_1 + \alpha_{1i} + \beta_{1ij} \quad \text{for} \quad j < T, \]
\[ m_{ij} = \mu_2 + \alpha_{2i} + \beta_{2ij} \quad \text{for} \quad j \geq T, \; i \leq n - T + 1, \]

are also possible. However they are not considered further in this paper as the loss reserves data in the numerical illustration show no clear trends of claim amounts across policy years for each given lag-year.

4. Error distributions

Past experience shows that irregular claims which often present in loss reserves data may seriously distort the parameter estimates and hence affect the accuracy of the prediction. The traditional log-linear models with normal errors fall short of allowing for the irregularities in the claim amounts. Heavy-tailed distributions such as the Cauchy, Student-$t$, Laplace and EP distributions are adopted for the errors in order to robustify the statistical inferences. However searching across these error distributions for the most suitable distribution, though workable, is time-consuming. The GT distribution, nesting all these
competing heavy-tailed distributions as well as the lighter-tailed alternatives including the normal and uniform distributions, provides a favorable choice for the error distribution (Butler et al., 1990).

4.1. Generalized-$t$ distributions

Proposed by McDonald and Newey (1988) in econometric applications, the GT distribution is symmetric and unimodal and has a probability density function (PDF)

$$GT(x | \mu, \sigma, p, q) = \frac{p}{2q^{\frac{1}{2}} \sigma B\left(\frac{1}{p}, q\right) \left(1 + \left|\frac{x-\mu}{\sigma}\right|^p\right)^{q + \frac{1}{p}}}$$

(10)

where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, $p > 0$ and $q > 0$ are two shape parameters and $B(\cdot)$ is the beta function. When $\mu = 0$ and $\sigma = 1$, the $r^{th}$ moment of the GT distribution is given by

$$E[X^r] = \frac{q^r \Gamma\left(\frac{r+1}{p}\right) \Gamma\left(q - \frac{r}{p}\right)}{\Gamma\left(\frac{1}{p}\right) \Gamma(q)}$$

provided that $r < pq$.

In particular, the variance is given by

$$V[X] = \frac{q^2 \frac{3}{p} \Gamma\left(\frac{3}{p}\right) \Gamma\left(q - \frac{2}{p}\right)}{\Gamma\left(\frac{1}{p}\right) \Gamma(q)}$$

provided that $pq > 2$.

The tail behavior and other characteristics are controlled by $p$ and $q$. Large values of $p$ and $q$ signify distributions with thinner tails than the normal distribution whereas small values of $p$ and $q$ signify distributions with thicker tails. Hence the GT distribution can accommodate both leptokurtic and platykurtic distributions.

The GT family includes several important distributions: the normal ($p = 2$, $q \to \infty$), the Cauchy ($p = 2$, $q = 1/2$), the Student-$t$ with $2q$ degrees of freedom ($p = 2$), the Laplace ($p = 1$, $q \to \infty$), the EP with shape parameter $p(q \to \infty)$, and the uniform ($p \to \infty$) distributions. When $p = 2$, the PDF in (10) reduces to

$$t(x | \mu, \tau, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\tau \sqrt{v\pi} \Gamma\left(\frac{v}{2}\right) \left(1 + \left(\frac{x-\mu}{\tau}\right)^2\right)^{\frac{v+1}{2}}}$$
which is the PDF of the Student-\( t \) distribution with degrees of freedom \( \nu = 2q \) and scale parameter \( \tau = \sigma / \sqrt{2} \). Moreover, (10) converges to the PDF of the EP distribution with shape parameter \( p \)

\[
EP(x | \mu, \sigma, p) = \frac{\exp\left(-\frac{|x - \mu|}{\sigma}ight)}{2\sigma \Gamma\left(1 + \frac{1}{p}\right)}
\]

when \( q \to \infty \). Lastly, (10) also converges to the PDF of the uniform distribution

\[
U(x | \mu, \sigma) = \frac{1}{2\sigma} I(x \in (\mu - \sigma, \mu + \sigma))
\]

when \( p \to \infty \). In addition, when \( p \leq 1 \), the GT distribution is cuspidate. Figure 1 summarizes the relationship amongst the well-known distributions within the GT family. The graphs of the PDFs for different values of \( p \) and \( q \) when \( \mu = 0 \) and \( \sigma = 1 \) are available from the authors.

4.2. The generalized gamma distributions

The generalized gamma (GG) distribution with parameters \( \alpha, \beta, \gamma > 0 \), denoted by \( GG(\alpha, \beta, \gamma) \) has a PDF given by

\[
GG(x | \alpha, \beta, \gamma) = \frac{\gamma^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\beta^{-\gamma} x^\gamma\right)
\]

where \( \beta \) is a scale parameter and \( \alpha \) and \( \gamma \) are shape parameters. The \( s^{th} \) moment of the distribution is

\[
E[X^s] = \frac{\Gamma(\alpha + s\gamma^{-1})}{\beta^s \Gamma(\alpha)}
\]
The GG distribution includes the Weibull ($\alpha = 1$), gamma ($\gamma = 1$), exponential ($\alpha = \gamma = 1$) and lognormal ($\alpha = 0$) distributions as special cases. In particular, when $\gamma = 1$, the GG distribution reduces to the gamma $Ga(\alpha, \beta)$ distribution having the PDF

$$Ga(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x),$$

the mean $E[X] = \alpha/\beta$ and the variance $V[X] = \alpha/\beta^2$ For inferential procedures of the GG distribution, see Hager and Bain (1970). See also McDonald and Butler (1987) for the applications in finance.

4.3. The scale mixtures representation of the GT distribution

Recently, Arslan and Genc (2003) showed that the GT distribution can be expressed into a scale mixtures of Box and Tiao distribution (Box and Tiao, 1973), or the scale mixtures of exponential power (SMEP) distribution because the Box and Tiao distribution is another name for the EP distribution. Using the SMEP density representation, the PDF in (10) can be rewritten as

$$GT(x \mid \mu, \sigma, p, q) = \int_0^\infty EP(x \mid \mu, q^{1/p} s^{1/2} \sigma, p) GG(s \mid q, 1, p/2) ds. \quad (11)$$

The latent variable, $s$, arises in the above expression is called the mixing parameter of the SMEP density representation and has a $GG(q, 1, p/2)$ distribution. Similar to the mixing parameters of the scale mixtures of normal (SMN) distributions (Choy and Smith, 1997) and the scale mixtures of uniform (SMU) distributions (Choy and Chan, 2003), mixing parameter $s$ in (11) can be used as a global diagnostic of possible outliers. Small mean value of $s$ associates with possible outlier because the corresponding EP distribution in (11) must inflate its variance to accommodate the outlier. See Choy and Smith (1997) and Choy and Chan (2003) for outlier diagnostics using the mixing parameter of the SMN and SMU distributions, respectively.

Many symmetric and non-normal distributions are extremely difficult to handle in statistical inference because of the non-conjugate structure of the models. In Bayesian computation, expressing the PDF of a distribution into a certain kind of scale mixtures form enables efficient simulation-based algorithms. Without using a scale mixtures representation, sampling non-standard full conditionals in Bayesian Gibbs sampling procedure relies on the Metropolis-Hastings algorithm, the slice sampler (Damien et al., 1999) or other simulation algorithms. However, the Metropolis-Hastings algorithm and some other simulation algorithms require expertise in simulation techniques to provide reliable and efficient algorithms. The slice sampler is an easy-to-use algorithm that introduces an auxiliary variable to reduce a non-standard full conditional in a Gibbs sampler to standard full conditionals. However, there is no physical
meaning for the auxiliary variable. On the contrary, the use of a scale mixtures
density representation for a non-normal distribution results in adding extra
latent variables in the model and hopefully produces a set of standard full
conditionals for the Gibbs sampler. In addition, these latent variables are used
to identify possible outliers. Therefore, the slice sampler and the use of scale
mixtures density representation play different roles in Gibbs sampling even
though they both use latent variables. Although our proposed data augmentation
scheme, as well as the slice sampler, suffers from a slow convergence of
the Markov chain, and hence increases the computational time, it does make
the Gibbs sampler easier to set-up.

Although the GT distribution can be expressed as a SMEP distribution, the
EP distribution is difficult to handle analytically in statistical inference because
its density function involves an absolute term. Walker and Gutiérrez-Peña
(1999) showed that the EP distribution can be expressed as a SMU distribution
and Choy and Chan (2003) successfully adopt this SMU representation of the
EP distribution in an insurance application. In this paper, we combine the
SMEP for the GT distribution and the SMU for the EP distribution to obtain
a SMU representation for the GT distribution. The resulting Gibbs sampler
can be shown to have a set of easily sampled standard full conditional distri-
butions.

Let $U$ be a gamma $Ga(1 + 1/p, 1)$ random variable and $S$ be a generalized
gamma $GG(q, 1, p/2)$ random variable. Then, if $X$ is a GT random variable with
parameters $m, s, p$ and $q$ we have

$$X | U = u, S = s \sim Unif \left( \mu - q^{1/2} s^{-1/2} u^{1/2} \sigma, \mu + q^{1/2} s^{-1/2} u^{1/2} \sigma \right)$$

$$U \sim Ga \left(1 + \frac{1}{p}, 1\right)$$

and

$$S \sim GG \left(q, 1, \frac{p}{2}\right).$$

The PDF in (10) becomes

$$\int_0^\infty \int_0^\infty Unif \left( x | \mu - q^{1/2} s^{-1/2} u^{1/2} \sigma, \mu + q^{1/2} s^{-1/2} u^{1/2} \sigma \right) Ga \left( u \left| 1 + \frac{1}{p}, 1 \right. \right) GG \left( s \left| q, 1, \frac{p}{2} \right. \right) du \, ds$$

where

$$\phi = \frac{1}{q} \left| \frac{x - \mu}{\sigma} \right|^{p} \frac{q}{s^{q}}.$$

In other words, it can be expressed into an infinite mixture of uniform den-
sities. Therefore, conditioning on $u$ and $s$, the variable $X$ follows a uniform
distribution with a variance of \((qu)^{2/3} \sigma^2 / (3s)\) which inflates with a large \(u\) and a small \(s\).

Hence the Gibbs sampling is performed with the observed data \(x\) and the mixing parameters \(u\) and \(s\). Although the number of parameters to be sampled becomes larger with the inclusion of the mixing parameters, sampling can be conducted more efficiently from the standard full conditional distributions. Hence the SMU representation of the GT density function simplifies the Gibbs sampler for Bayesian computation.

Robust analysis using heavy-tailed distributions to accommodate outliers has been studied by many authors (Box and Tiao, 1973, Dawid, 1973). Choy and Smith (1997) and Choy and Chan (2003) highlight the use of latent variables of the normal and uniform scale mixtures distributions, respectively, to simplify the Bayesian computational algorithms and as a proxy for outlier diagnostics. For example, a normal scale mixtures distribution is an infinite mixture of the normal distributions with identical means and different variances. The weight and the variance of each normal component are determined by a latent variable, known as a mixing parameter of a scale mixture distribution, which follows a pre-deterministic distribution. A normal scale mixture distribution accommodates a possible outlier by inflating the variances of the normal components. Therefore, by comparing the posterior mean or median of the mixing parameter for each observation, we can identify possible outliers. This argument is similar for the uniform scale mixtures distributions.

For the GT distribution, we propose to use the ratio \(c = u^{1/3} s^{-1/2}\) as a proxy to identify outliers because it is proportional to the variances of the infinite mixture of uniform components. The observations with significantly large posterior mean values of \(c_i\) are possible outliers.

5. Bayesian Analysis

5.1. Model implementation

Bayesian approach is adopted for parameter estimation and is performed using WinBUGS. Command codes for the implementations are available from the authors. Parameters are estimated using the samples drawn from the intractable joint posterior distribution. Due to the complexity of the models, simulated values of certain parameters may be highly correlated and the Markov chains of the Gibbs sampler may suffer from a slow convergence. As a result, the number of iterations \(M\) should be large enough to ensure that the sample is uncorrelated, large and stationary. We set \(M = 105,000\) and the burn-in period is at least 5,000 iterations. After the burn-in period, parameters are taken from every 50th iteration to mimic a random sample of size at least \(K = 1,000\) from the intractable joint posterior distribution. Trajectory plots and autocorrelation plots of the simulated values are used to check for the convergence and independence of the sample. The posterior sample means, or medians where appropriate, are reported as parameter estimates.
5.2. Model Selection

To choose the most appropriate Bayesian model for the loss reserves data, the Deviance information criterion (DIC) is used. The DIC is proposed by Spiegelhalter, et al., (2002) as a model selection method for complex hierarchical models in which the number of parameters is not clearly defined. The DIC is defined as

\[
DIC = E[D(\theta)] + \Delta D(\hat{\theta})
\]

where \(E[D(\theta)]\) is a measure of the adequacy of model fitting and \(\Delta D(\hat{\theta}) = E[D(\theta)] - D(\hat{\theta})\) estimates the effective number of parameters in the model. Based on the posterior sample, we calculate the DIC as

\[
DIC = -2 \left( \frac{2}{K} \sum_{k=1}^{K} \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \ln f_{ij}(z_{ij} | \theta_k) - \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \ln f_{ij}(z_{ij} | \bar{\theta}) \right)
\]

where \(z_{ij} = \ln y_{ij}\) is the log claim size, \(f_{ij}(\cdot)\) is the GT density function given by (10) when the mean \(\mu\) is replaced by \(\mu_{ij}\) in (2), (4) and (5) for the log-ANOVA, log-ANCOVA and state space models, respectively, \(\theta_k\) is a vector of all parameters in the \(k^{th}\) posterior sample and \(\bar{\theta}\) is a vector of posterior means. Obviously, a model with the smallest DIC is the best model.

6. Results

6.1. Comparison between mean functions and error distributions

Various models with different mean functions and error distributions are fitted to the claim data. For each type of mean function, we set the two shape parameters, \(p\) and \(q\), of the GT distribution to take different values, which signify different distributions that are nested within the GT family. In the simulation study, we set \(p = 50\) to approximate the uniform distribution and \(q = 50\) to approximate the EP distribution which includes the normal and Laplace distributions. From (10), these values for \(p\) and \(q\) approximate the limiting cases of \(p \rightarrow \infty\) and \(q \rightarrow \infty\) well. Moreover, more general families of distributions can be obtained by setting either \(p\) or \(q\) or both to be random. For example, \(p = 1\) and \(q\) is random, and \(p\) is random and \(q = 0.5, 1\) and \(2\), respectively, are used in the simulation study. Models are ranked according to DIC. These shape parameters are assigned a non-informative Ga(0.001, 0.001) prior distribution if they are not fixed.

Table 3 exhibits the DIC values for a wide choice of error distributions for the log-ANOVA, log-ANCOVA and state space models. For the log-ANOVA model, the most appropriate error distribution is the GT distribution with \(p = 1\) and \(q\) being random. The posterior mean of \(q\) is 2.46. For the log-ANCOVA and state space models, the GT error distribution with a random \(p\) and \(q = 2\)
6.2. Comparison between simple and threshold models

To allow for structural changes in the trend of claim amount over time, threshold model that assumes a structural change at policy year $T$ is suggested. For the loss reserves data used throughout this paper, state space model is shown to perform better than the log-ANOVA and log-ANCOVA models across a wide choice of error distributions in general. Therefore, the state space threshold model and the GT error distribution with unknown $p$ and $q$ are used to re-analyze the claim data in this section. The threshold $T$ is fixed within the range $(1, n)$ for the policy year, not too close to either sides of the range. The mean function takes two different sets of parameters – one set before $T$ and the other set on and after $T$. Table 4 exhibits the posterior means and posterior

### Table 4: Posterior means of $p$ and $q$ (standard deviations in parentheses) and $DIC$ for different state space threshold models.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>$T = 4$</th>
<th>$T = 5$</th>
<th>$T = 6$</th>
<th>$T = 7$</th>
<th>$T = 8$</th>
<th>$T = 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>1.19 (0.16)</td>
<td>0.84 (0.10)</td>
<td>0.79 (0.07)</td>
<td>0.81 (0.16)</td>
<td>0.84 (0.08)</td>
<td>0.79 (0.10)</td>
</tr>
<tr>
<td>$q$</td>
<td>2.12 (0.45)</td>
<td>2.77 (0.48)</td>
<td>3.12 (0.64)</td>
<td>2.63 (0.61)</td>
<td>2.56 (0.47)</td>
<td>2.93 (0.58)</td>
</tr>
<tr>
<td>$DIC$</td>
<td>290.3</td>
<td>269.8</td>
<td>261.6</td>
<td>228.3</td>
<td>240.3</td>
<td>280.2</td>
</tr>
</tbody>
</table>

is chosen and the posterior means of $p$ are 1.20 and 1.12, respectively. These $p$ and $q$ values correspond to the GT distributions that are heavier-tailed than the normal distribution. The $DIC$ values for these three models are 336.0, 332.6 and 321.1, respectively and the state space model with a random $p$ and $q = 2$ is preferred amongst the models studied.

#### Table 3: Model comparison using $DIC$ for ANOVA, ANCOVA and state space models with different error distributions. Posterior means of the shape parameters of the error distributions are reported while bold type values correspond to fixed shape parameter values.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$p$</th>
<th>$q$</th>
<th>$DIC$</th>
<th>$p$</th>
<th>$q$</th>
<th>$DIC$</th>
<th>$p$</th>
<th>$q$</th>
<th>$DIC$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>2</td>
<td>50</td>
<td>542.4</td>
<td>2</td>
<td>50</td>
<td>529.9</td>
<td>2</td>
<td>50</td>
<td>528.3</td>
</tr>
<tr>
<td>Cauchy</td>
<td>2</td>
<td>0.5</td>
<td>375.0</td>
<td>2</td>
<td>0.5</td>
<td>340.8</td>
<td>2</td>
<td>0.5</td>
<td>356.7</td>
</tr>
<tr>
<td>EP</td>
<td>1.24</td>
<td>50</td>
<td>430.4</td>
<td>1.24</td>
<td>50</td>
<td>416.9</td>
<td>1.24</td>
<td>50</td>
<td>418.3</td>
</tr>
<tr>
<td>Laplace</td>
<td>1</td>
<td>50</td>
<td>375.6</td>
<td>1</td>
<td>50</td>
<td>359.0</td>
<td>1</td>
<td>5</td>
<td>350.9</td>
</tr>
<tr>
<td>Uniform</td>
<td>50</td>
<td>2</td>
<td>582.1</td>
<td>50</td>
<td>2</td>
<td>617.2</td>
<td>50</td>
<td>2</td>
<td>800.4</td>
</tr>
<tr>
<td>GT $q=0.5$</td>
<td>3.75</td>
<td>0.5</td>
<td>383.2</td>
<td>3.54</td>
<td>0.5</td>
<td>357.9</td>
<td>3.52</td>
<td>0.5</td>
<td>374.0</td>
</tr>
<tr>
<td>GT $q=1$</td>
<td>2.01</td>
<td>1</td>
<td>357.5</td>
<td>1.90</td>
<td>1</td>
<td>333.4</td>
<td>1.93</td>
<td>1</td>
<td>332.3</td>
</tr>
<tr>
<td>GT $q=2$</td>
<td>1.13</td>
<td>2</td>
<td>345.0</td>
<td>1.20</td>
<td>2</td>
<td>332.6</td>
<td>1.12</td>
<td>2</td>
<td>321.1</td>
</tr>
<tr>
<td>GT $p=1$</td>
<td>1</td>
<td>2.46</td>
<td>336.0</td>
<td>1</td>
<td>3.09</td>
<td>333.5</td>
<td>1</td>
<td>2.49</td>
<td>321.9</td>
</tr>
<tr>
<td>GT</td>
<td>0.93</td>
<td>3.45</td>
<td>337.3</td>
<td>1.16</td>
<td>2.26</td>
<td>333.9</td>
<td>1.06</td>
<td>2.73</td>
<td>323.4</td>
</tr>
</tbody>
</table>

#### Table 4: Posterior means of $p$ and $q$ (standard deviations in parentheses) and $DIC$ for different state space threshold models.
ROBUST BAYESIAN ANALYSIS OF LOSS RESERVES DATA

Figure 2: Change of DIC against threshold year, T, for threshold state space models.

Figure 3: Box-plots of the posterior samples of (a) \( \alpha_{1,1}, \ldots, \alpha_{1,6} \) and \( \alpha_{2,2}, \ldots, \alpha_{2,12} \), (b) \( \mu_1 \) and \( \mu_2 \), (c) \( p \) and \( q \) and (d) \( \sigma \).
standard deviations of $p$, $q$ and $DIC$ of various state space threshold models. Figure 2 displays the change of $DIC$ versus $T$. The state space model with threshold year $T = 7$ (year 1984) is shown to be the most appropriate model ($DIC = 228.3$). The posterior means of $p$ and $q$ are $\hat{p} = 0.81$ and $\hat{q} = 2.63$, respectively. This threshold model is chosen for the loss reserves prediction.

Figure 3 displays the box-plots of the posterior samples of all model parameters except the $\beta_j$s because there are too many parameters. The alpha and beta parameters in (8) and (9) are given by (6) and (7) respectively before and after the threshold year of 1984. The beta parameters for each policy year form a periodic trend across lag years: they increase up to lag years 4 to 6 and then decrease. Moreover the beta parameters at higher lag years have larger variability due to the insufficiency of observations to estimate the parameters accurately. Box-plots of the alpha parameters also show trends of increasing variances across policy years before and after 1984 (the first five box-plots and the remaining box-plots respectively).

Table 5 reports the predicted claims in the upper triangle. Claims in italic refer to the predicted claims from policies written on and after 1984. Claims underlined are predicted from claims which are extremely large or small as compared with their neighboring claims.

The predicted claims in Table 5 show two different trends of claims before and after 1984. For claims from policies written before 1984, the predicted claims increase with the lag year till the 4th lag year and drop slowly thereafter. For policies written on and after 1984, the predicted claims start from

<table>
<thead>
<tr>
<th>Year</th>
<th>$E(Y)$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>3494</td>
<td>8373</td>
<td>9300</td>
<td>9881</td>
<td>7567</td>
<td>4606</td>
<td>3365</td>
<td>2387</td>
<td>1595</td>
<td>553</td>
<td>1013</td>
<td>443</td>
<td>342</td>
<td>397</td>
<td>332</td>
<td>218</td>
<td>58</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>3964</td>
<td>6630</td>
<td>7364</td>
<td>7824</td>
<td>5991</td>
<td>3647</td>
<td>2665</td>
<td>1890</td>
<td>1263</td>
<td>438</td>
<td>802</td>
<td>351</td>
<td>371</td>
<td>314</td>
<td>326</td>
<td>24</td>
<td>46</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>4082</td>
<td>9165</td>
<td>10179</td>
<td>10815</td>
<td>8282</td>
<td>5042</td>
<td>3684</td>
<td>2613</td>
<td>1746</td>
<td>605</td>
<td>1109</td>
<td>485</td>
<td>374</td>
<td>434</td>
<td>363</td>
<td>24</td>
<td>70</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>4091</td>
<td>9517</td>
<td>10571</td>
<td>11231</td>
<td>8600</td>
<td>5236</td>
<td>3825</td>
<td>2710</td>
<td>1812</td>
<td>628</td>
<td>1152</td>
<td>504</td>
<td>388</td>
<td>451</td>
<td>377</td>
<td>26</td>
<td>70</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>4032</td>
<td>9662</td>
<td>10702</td>
<td>11402</td>
<td>8731</td>
<td>6315</td>
<td>4893</td>
<td>3784</td>
<td>1840</td>
<td>638</td>
<td>1169</td>
<td>511</td>
<td>294</td>
<td>458</td>
<td>403</td>
<td>24</td>
<td>70</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>3959</td>
<td>9388</td>
<td>10427</td>
<td>11079</td>
<td>8483</td>
<td>6164</td>
<td>3773</td>
<td>2676</td>
<td>1788</td>
<td>620</td>
<td>1136</td>
<td>497</td>
<td>383</td>
<td>432</td>
<td>363</td>
<td>22</td>
<td>68</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>2802</td>
<td>5067</td>
<td>7723</td>
<td>9293</td>
<td>9284</td>
<td>9311</td>
<td>8434</td>
<td>4362</td>
<td>4377</td>
<td>1720</td>
<td>1162</td>
<td>927</td>
<td>243</td>
<td>283</td>
<td>237</td>
<td>22</td>
<td>41</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>2542</td>
<td>4976</td>
<td>6757</td>
<td>8131</td>
<td>8122</td>
<td>8147</td>
<td>7379</td>
<td>3816</td>
<td>3829</td>
<td>1505</td>
<td>1016</td>
<td>800</td>
<td>209</td>
<td>242</td>
<td>200</td>
<td>20</td>
<td>24</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>2315</td>
<td>4572</td>
<td>6209</td>
<td>7471</td>
<td>7464</td>
<td>7486</td>
<td>6781</td>
<td>3507</td>
<td>3519</td>
<td>1382</td>
<td>953</td>
<td>732</td>
<td>195</td>
<td>226</td>
<td>186</td>
<td>18</td>
<td>22</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>2144</td>
<td>4148</td>
<td>5633</td>
<td>6777</td>
<td>6770</td>
<td>6791</td>
<td>6151</td>
<td>3181</td>
<td>3192</td>
<td>1238</td>
<td>844</td>
<td>663</td>
<td>174</td>
<td>201</td>
<td>168</td>
<td>16</td>
<td>29</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>2124</td>
<td>4064</td>
<td>5519</td>
<td>6640</td>
<td>6653</td>
<td>6653</td>
<td>6026</td>
<td>2118</td>
<td>2129</td>
<td>1244</td>
<td>811</td>
<td>654</td>
<td>171</td>
<td>197</td>
<td>164</td>
<td>16</td>
<td>27</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>2162</td>
<td>4104</td>
<td>5573</td>
<td>6706</td>
<td>6699</td>
<td>6719</td>
<td>6286</td>
<td>2076</td>
<td>2073</td>
<td>1211</td>
<td>845</td>
<td>666</td>
<td>171</td>
<td>198</td>
<td>164</td>
<td>16</td>
<td>29</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>2228</td>
<td>4528</td>
<td>6148</td>
<td>7398</td>
<td>7390</td>
<td>7473</td>
<td>6641</td>
<td>3405</td>
<td>3620</td>
<td>1371</td>
<td>933</td>
<td>726</td>
<td>201</td>
<td>222</td>
<td>165</td>
<td>18</td>
<td>32</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>2554</td>
<td>5047</td>
<td>6854</td>
<td>8247</td>
<td>8239</td>
<td>8060</td>
<td>7388</td>
<td>3834</td>
<td>7986</td>
<td>1520</td>
<td>1040</td>
<td>812</td>
<td>211</td>
<td>245</td>
<td>204</td>
<td>20</td>
<td>25</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>2647</td>
<td>5269</td>
<td>7155</td>
<td>8609</td>
<td>8561</td>
<td>8475</td>
<td>7742</td>
<td>4016</td>
<td>8244</td>
<td>1588</td>
<td>1090</td>
<td>858</td>
<td>222</td>
<td>254</td>
<td>214</td>
<td>21</td>
<td>27</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>2738</td>
<td>5473</td>
<td>7423</td>
<td>8814</td>
<td>8928</td>
<td>8875</td>
<td>8101</td>
<td>4211</td>
<td>8509</td>
<td>1651</td>
<td>1146</td>
<td>885</td>
<td>231</td>
<td>269</td>
<td>225</td>
<td>22</td>
<td>29</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>2727</td>
<td>5506</td>
<td>7496</td>
<td>8924</td>
<td>8993</td>
<td>8938</td>
<td>8175</td>
<td>4241</td>
<td>7750</td>
<td>1661</td>
<td>1154</td>
<td>893</td>
<td>235</td>
<td>271</td>
<td>224</td>
<td>22</td>
<td>29</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>2782</td>
<td>5548</td>
<td>7524</td>
<td>9038</td>
<td>9054</td>
<td>9097</td>
<td>8127</td>
<td>4279</td>
<td>4037</td>
<td>1687</td>
<td>1163</td>
<td>894</td>
<td>236</td>
<td>272</td>
<td>224</td>
<td>22</td>
<td>28</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Run-off triangle of the predicted claims (upper triangle) and projected claims (lower triangle) using the state space threshold model with $T = 7$. Values in italic are the predicted or projected claims for policies written on and after 1984. Values in black are the predicted claims and those in grey are the projected claims. Values underlined in black are the outlying predicted claims. Values underlined in grey are the projected claims with standard errors larger than half of the values.
lower levels, rise to the lower maximums at the 6th lag year and drop more rapidly thereafter. In general, the predicted claims (Table 5) lie close to the observed claims (Table 2) showing that the chosen model fits the data well.

Projection for the outstanding claims in the lower triangle of Table 2 using the chosen model is reported in Table 5. The values are the posterior medians of the projected claims using the $K = 1,000$ sets of parameter estimates. The projected total outstanding claims, as the posterior median of the 1,000 sums of projected claims, is 296,159 dollars with an estimated standard error of 123,867 dollars. The projected total outstanding claims for each policy year and their standard errors are reported in Table 6. Note that the projected claims $\tilde{Y}_{ij}$, $i \geq 7$, $j \geq 13$ are estimated using

$$\beta_{2,0} = \beta_{1,i-1,j} + \nu_{2,i}, \quad i \geq 8, \quad j \geq 13$$

and $\beta_{2,7j} = \beta_{1,j}$, $j \geq 13$ because their lag year effects $\beta_{2,ij}$ are not given by the model. Moreover it should be noted that some projected claims are underlined in Table 5 because their standard errors are more than half of the estimated values. As discussed above, the beta parameters for higher lag year have higher variability because fewer observations are available to estimate these parameters accurately.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>11</td>
<td>86</td>
<td>131</td>
<td>515</td>
<td>928</td>
<td>868</td>
<td>1558</td>
<td>2450</td>
<td>3539</td>
<td>4729</td>
<td>10760</td>
<td>20295</td>
<td>31810</td>
<td>42016</td>
<td>52743</td>
<td>60016</td>
<td>61717</td>
</tr>
<tr>
<td>t.e.</td>
<td>69</td>
<td>161</td>
<td>210</td>
<td>229</td>
<td>254</td>
<td>219</td>
<td>11082</td>
<td>10828</td>
<td>8996</td>
<td>8697</td>
<td>8561</td>
<td>10202</td>
<td>12729</td>
<td>12688</td>
<td>12604</td>
<td>14927</td>
<td>17500</td>
</tr>
</tbody>
</table>

**TABLE 6**: Total projected outstanding claims for each policy year and their standard errors.

**FIGURE 4**: Observed and predicted relative frequencies for residuals using the chosen model.
Figure 4 plots the observed relative frequencies and the expected probabilities using the density function (10) for the residuals $r_{ij} = \ln y_{ij} - \mu_{ij}$ where $\mu_{ij}$ is given by (8) and (9). As the observed relative frequencies are very close to the expected probabilities, the chosen model provides a very good fit to the loss reserves data. The mean, median, standard deviation, skewness and kurtosis for the residuals are –0.1655, –0.00715, 1.1306, –5.651 and 46.189 respectively. The residuals are quite leptokurtic and negatively skewed due to the existence of several extremely small outliers.

6.3. Detection of Outliers

The SMU representation of the GT distribution not only simplifies the model implementation using the Bayesian approach but also allows detection of possible outliers using the parameter $\psi_{ij} = u_{ij}^{1/p} S_{ij}^{1/2}$. An unusually large $\psi_{ij}$ value indicates that the observed claim amount is a possible outlier. The posterior medians of $\psi_{ij}$ for the chosen model are displayed in Figure 5. They identify observations 14, 29, 33, 34 and 35 (64 and 65 are marginal), labeled across a row from top to bottom in Table 2, to be possible outliers since their $\psi_{ij} > 10$ ($\psi_{ij}$ for observations 64 and 65 are 8.94 and 8.70 respectively). The level of 10 is chosen to identify a moderate number of observations as outliers. Table 7 gives a summary of these outlying observations. These observations, grey in color and underlined in Table 2, correspond to the claim amounts which are substantially lower than their neighboring values. They may lead to an underestimation for the loss reserves and hence lower the solvency of an insurance company and increase its risk of bankruptcy. However, distorting effects from these outliers are down-weighted and hence inference is protected when the GT error distribution is used.

<table>
<thead>
<tr>
<th>Outlier index</th>
<th>Policy year $i$</th>
<th>Lag year $j$</th>
<th>$\psi_{ij}$</th>
<th>$y_{ij}$</th>
<th>$E(y_{ij})$</th>
<th>Post. $p$-value $p_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1978</td>
<td>14</td>
<td>40.05</td>
<td>0.01</td>
<td>392</td>
<td>0.0000</td>
</tr>
<tr>
<td>29</td>
<td>1979</td>
<td>11</td>
<td>11.55</td>
<td>35</td>
<td>798</td>
<td>0.0106</td>
</tr>
<tr>
<td>33</td>
<td>1979</td>
<td>15</td>
<td>12.69</td>
<td>6</td>
<td>244</td>
<td>0.0080</td>
</tr>
<tr>
<td>34</td>
<td>1979</td>
<td>16</td>
<td>10.81</td>
<td>1</td>
<td>29</td>
<td>0.0099</td>
</tr>
<tr>
<td>35</td>
<td>1979</td>
<td>17</td>
<td>13.25</td>
<td>0.01</td>
<td>3</td>
<td>0.0000</td>
</tr>
<tr>
<td>64</td>
<td>1981</td>
<td>13</td>
<td>8.94</td>
<td>38</td>
<td>391</td>
<td>0.0197</td>
</tr>
<tr>
<td>65</td>
<td>1981</td>
<td>14</td>
<td>8.70</td>
<td>45</td>
<td>450</td>
<td>0.0213</td>
</tr>
<tr>
<td>100</td>
<td>1984</td>
<td>7</td>
<td>1.30</td>
<td>11920</td>
<td>8493</td>
<td>0.1712</td>
</tr>
<tr>
<td>165</td>
<td>1992</td>
<td>4</td>
<td>2.50</td>
<td>15546</td>
<td>8467</td>
<td>0.2595</td>
</tr>
</tbody>
</table>

Table 7: Posterior medians of $\psi_{ij}$, observed values $y_{ij}$, fitted values $E(y_{ij})$ and the posterior $p$-values $p_{ij}$ of the extreme observations.
Figure 5: Posterior medians of the parameters in the chosen model.

Figure 6 plots the observed versus predicted claims (in black for the chosen model). The plot shows that observations 100 and 165 (grey, italic and underlined in Table 2) are potential outliers as the observed claim amounts are much higher than their predicted values. However, their $\psi_{ij}$ values are 1.3
and 2.5, respectively indicating that $\psi_{ij}$ is less sensitive to large outliers. One possible explanation lies on the log-linear model: the log function is more sensitive to low values than large values. As effects from large outliers are less likely to be down-weighted by the mixing parameters in the SMU representation, loss reserves may be over-estimated. However the problem of over-estimating loss reserves may lower the profit of an insurance company but it may not weaken its solvency. Hence the risk of bankruptcy may not be seriously affected by the over-estimation.

Table 7 displays the posterior medians of $\psi_{ij}$, observed claims $y_{ij}$ and posterior expected claims $E(y_{ij})$ and posterior predictive $p$-values, $p_{ij}$, for the possible outliers. The posterior $p$-values (Meng, 1994 and Spiegelhalter, et al., 2004) are approximated using the following equation

$$p_{ij} = \begin{cases} 2 F_u(\ln(y_{ij})), & r_{ij} < 0, \\ 2[1 - F_u(\ln(y_{ij}))], & r_{ij} > 0, \end{cases}$$

where $F_u(\cdot)$ is an approximated cumulative distribution function of the density function in (10) and $\mu_{ij}$ is given by (8) and (9). First, we set the approximated density function $f_u(\cdot)$ to be $GT(m_k | \mu, \sigma, p, q)$ in (10) at a number of mass points $m_k$ within $\mu \pm 5.5\sigma$ and with an increment of 0.01. Then the approximated cumulative distribution function $F_u(\ln(y_{ij}))$ is given by the sum of the area of the trapeziums such that

$$F_u(\ln(y_{ij})) = \sum_{k \leq k'} 0.005[GT(m_k | \mu, \sigma, p, q) + GT(m_k + 0.01 | \mu, \sigma, p, q)]$$

and $m_k < \ln(y_{ij}) < m_{k'} + 0.01$. All except observations 100 and 165, written in italic in Table 7, give negative residuals, $r_{ij} = \ln y_{ij} - \mu_{ij}$. The posterior $p$-values $p_{ij}$ indicate that these extreme observations, except observations 100 and 165, are outliers.

6.4. Comparison with chain ladder method

The chosen model produces more accurate prediction of claims than the popular chain ladder method which indirectly estimates the incremental loss $y_{ij}$ by projecting the cumulative loss $S_{ij}$

$$\hat{S}_{ij} = \hat{S}_{i, j-1} \times f_j, \quad i = 2, \ldots, n, \quad j = n - i + 2, \ldots, n$$

where $S_{ij} = \sum_{k=1}^{i} y_{ik}$ and $S_{i, n+1-i} = \hat{S}_{i, n+1-i}$ using the projective factors

$$f_j = \frac{\sum_{i=1}^{n+1-j} S_{ij}}{\sum_{i=1}^{n+1-j} S_{i, j-1}}, \quad j = 2, \ldots, n.$$
TABLE 8: Run-off triangle of predicted claims (black) and projected claims (grey) using the chain ladder method.

<table>
<thead>
<tr>
<th>Year</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
<th>Y5</th>
<th>Y6</th>
<th>Y7</th>
<th>Y8</th>
<th>Y9</th>
<th>Y10</th>
<th>Y11</th>
<th>Y12</th>
<th>Y13</th>
<th>Y14</th>
<th>Y15</th>
<th>Y16</th>
<th>Y17</th>
<th>Y18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3323</td>
<td>7160</td>
<td>10531</td>
<td>11618</td>
<td>7967</td>
<td>5582</td>
<td>2164</td>
<td>1086</td>
<td>3700</td>
<td>807</td>
<td>178</td>
<td>326</td>
<td>489</td>
<td>284</td>
<td>220</td>
<td>548</td>
<td>214</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>3785</td>
<td>8156</td>
<td>13531</td>
<td>9052</td>
<td>4663</td>
<td>32180</td>
<td>-549</td>
<td>1017</td>
<td>-38</td>
<td>34</td>
<td>372</td>
<td>101</td>
<td>214</td>
<td>399</td>
<td>-77</td>
<td>-54</td>
<td>-22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4877</td>
<td>10078</td>
<td>11688</td>
<td>9548</td>
<td>7505</td>
<td>6450</td>
<td>4164</td>
<td>1464</td>
<td>442</td>
<td>141</td>
<td>470</td>
<td>632</td>
<td>372</td>
<td>395</td>
<td>443</td>
<td>256</td>
<td>-46</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>5288</td>
<td>11384</td>
<td>7437</td>
<td>16182</td>
<td>8890</td>
<td>4769</td>
<td>4647</td>
<td>1505</td>
<td>2005</td>
<td>-61</td>
<td>1837</td>
<td>2048</td>
<td>13</td>
<td>-20</td>
<td>9</td>
<td>-10</td>
<td>92</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>4943</td>
<td>14653</td>
<td>12513</td>
<td>12389</td>
<td>6408</td>
<td>3594</td>
<td>-252</td>
<td>582</td>
<td>2703</td>
<td>1142</td>
<td>351</td>
<td>60</td>
<td>276</td>
<td>227</td>
<td>130</td>
<td>49</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>3600</td>
<td>7775</td>
<td>8682</td>
<td>9725</td>
<td>720</td>
<td>7008</td>
<td>2702</td>
<td>2067</td>
<td>2433</td>
<td>679</td>
<td>449</td>
<td>688</td>
<td>490</td>
<td>322</td>
<td>219</td>
<td>113</td>
<td>43</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>3642</td>
<td>7848</td>
<td>8409</td>
<td>12191</td>
<td>7485</td>
<td>4186</td>
<td>3439</td>
<td>10430</td>
<td>3284</td>
<td>2954</td>
<td>1603</td>
<td>315</td>
<td>673</td>
<td>307</td>
<td>272</td>
<td>140</td>
<td>53</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2463</td>
<td>5027</td>
<td>5745</td>
<td>8738</td>
<td>6448</td>
<td>5493</td>
<td>7216</td>
<td>3923</td>
<td>2928</td>
<td>31</td>
<td>2984</td>
<td>883</td>
<td>287</td>
<td>242</td>
<td>214</td>
<td>111</td>
<td>42</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2267</td>
<td>4885</td>
<td>7680</td>
<td>7307</td>
<td>5599</td>
<td>7232</td>
<td>5494</td>
<td>5469</td>
<td>3729</td>
<td>1840</td>
<td>766</td>
<td>491</td>
<td>291</td>
<td>245</td>
<td>217</td>
<td>112</td>
<td>42</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2069</td>
<td>4329</td>
<td>3956</td>
<td>6267</td>
<td>6213</td>
<td>5700</td>
<td>7172</td>
<td>4487</td>
<td>5259</td>
<td>2727</td>
<td>937</td>
<td>464</td>
<td>275</td>
<td>232</td>
<td>205</td>
<td>106</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>1860</td>
<td>4008</td>
<td>7009</td>
<td>3698</td>
<td>7113</td>
<td>4091</td>
<td>5188</td>
<td>5717</td>
<td>1953</td>
<td>910</td>
<td>803</td>
<td>298</td>
<td>236</td>
<td>199</td>
<td>176</td>
<td>91</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2331</td>
<td>5023</td>
<td>3190</td>
<td>6009</td>
<td>5117</td>
<td>10118</td>
<td>8915</td>
<td>3699</td>
<td>2274</td>
<td>1066</td>
<td>932</td>
<td>462</td>
<td>273</td>
<td>230</td>
<td>204</td>
<td>105</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2314</td>
<td>4986</td>
<td>4962</td>
<td>4514</td>
<td>6385</td>
<td>11242</td>
<td>9712</td>
<td>3159</td>
<td>2395</td>
<td>1112</td>
<td>981</td>
<td>486</td>
<td>288</td>
<td>243</td>
<td>215</td>
<td>111</td>
<td>42</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2607</td>
<td>5617</td>
<td>3602</td>
<td>10906</td>
<td>6596</td>
<td>8534</td>
<td>4821</td>
<td>3048</td>
<td>2317</td>
<td>1076</td>
<td>949</td>
<td>471</td>
<td>279</td>
<td>235</td>
<td>208</td>
<td>107</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2595</td>
<td>5592</td>
<td>6234</td>
<td>7288</td>
<td>15640</td>
<td>7136</td>
<td>5968</td>
<td>3710</td>
<td>2820</td>
<td>1310</td>
<td>1156</td>
<td>573</td>
<td>329</td>
<td>286</td>
<td>253</td>
<td>131</td>
<td>49</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>3155</td>
<td>6796</td>
<td>4703</td>
<td>10078</td>
<td>7247</td>
<td>5860</td>
<td>4818</td>
<td>3646</td>
<td>2316</td>
<td>1075</td>
<td>949</td>
<td>470</td>
<td>278</td>
<td>235</td>
<td>208</td>
<td>107</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2626</td>
<td>5688</td>
<td>6739</td>
<td>6070</td>
<td>6765</td>
<td>5470</td>
<td>4499</td>
<td>2844</td>
<td>2162</td>
<td>1004</td>
<td>856</td>
<td>439</td>
<td>260</td>
<td>219</td>
<td>194</td>
<td>100</td>
<td>28</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2527</td>
<td>6093</td>
<td>7162</td>
<td>8640</td>
<td>7243</td>
<td>5856</td>
<td>4815</td>
<td>3644</td>
<td>2314</td>
<td>1075</td>
<td>948</td>
<td>470</td>
<td>278</td>
<td>234</td>
<td>208</td>
<td>107</td>
<td>40</td>
<td>13</td>
</tr>
</tbody>
</table>

The prediction of \( S_{ij} \) on higher lag years, which is further away from the latest known claim amount requires the product of more projection factors and hence becomes less reliable as less data are available. We predict \( S_{ij} \) in the upper triangle of Table 2 using only one year ahead projective factor so that

\[
\hat{S}_{ij} = S_{i,j-1} \times f_j, \quad 2 \leq j \leq n - i + 1
\]

and \( \hat{S}_{i1} = S_{i1} \). We project \( S_{ij} \) in the lower triangle by

\[
\hat{S}_{ij} = S_{i,n-i+1} \times \prod_{k=n-i+2}^{j} f_k, \quad i = 2, \ldots, 18, \quad j > n - i + 1.
\]

The predicted and projected claims are reported in Table 8. Note that there are some negative values. They refer to predicted and projected claims which are very close to zero. The MSE for the predicted claims using the chain ladder method is 3,908,069 which is much larger than 1,584,478 for the chosen model. Moreover Figure 6 reveals that the predicted claims using the chosen model (black diamonds) are closer to the observed claims than those using the chain ladder method (grey triangles). Thus the GT model provides more reliable prediction of loss reserves. Accurate prediction of loss reserves is very important to insurance companies since the levels of loss reserves have a dramatic impact on the profitability and solvency of insurance companies.

7. CONCLUSION

In this paper, log-linear models with three different forms of mean function are used to model loss reserves. For the data set analyzed, the GT error distribution
performs better than the normal and other well-known heavy-tailed distributions such as the Student-$t$ and EP distributions. Meanwhile, the time-recursive state space model with threshold in 1984 provides the best fit to the data. Prediction and projection of loss reserves are based on Bayesian approach. Expressing the GT distribution into the uniform scale mixtures form makes Gibbs sampler easier to set up and enables outlier diagnostics using the mixing parameters. Model selection relies on the deviance information criterion. Comparing with the popular chain ladder method, the chosen GT model is shown to provide more accurate predicted claims. The chosen model is then used to project outstanding claims of the lower triangle in Table 2.

As the log-linear model is more sensitive to low values than large values, the residuals for the data set analyzed are negatively skewed. To allow for a skewed error distribution, one may consider the skewed heavy-tailed distributions including the skewed-$t$ distribution or the scale mixtures of beta distribution. Further investigation into the skewed error distributions will surely improve the modeling of loss reserves data.

ACKNOWLEDGEMENT

The authors would like to thank Deborah Street and anonymous referees for their very constructive comments, Connie P.Y. Lam for calculating the DICs and Wai-Yin Wan for running some WinBUGS programs. The second author is supported by an internal grant of the University of Technology Sydney, Australia.

REFERENCES


DAWID, A.P. (1973) Posterior expectations for large observations, Biometrika, 60, 664-666.


JENNIFER S.K. CHAN
School of Mathematics and Statistics
The University of Sydney,
NSW 2006, Australia
E-Mail: jchan@maths.usyd.edu.au

S.T. BORIS CHOY
Department of Mathematical Sciences
University of Technology Sydney,
P.O. Box 123, Broadway,
NSW 2007, Australia
E-Mail: boris.choy@uts.edu.au

UDI E. MAKOV
Department of Statistics
University of Haifa
Haifa, 31905
Israel
E-Mail: makov@stat.haifa.ac.il