Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty

Ioannis Ntzoufras[1] and Petros Dellaportas[2]

Abstract

This paper deals with the prediction of the amount of outstanding automobile claims that an insurance company will pay in the near future. We consider various competing models using Bayesian theory and Markov chain Monte Carlo methods. Claim counts are used to add a further hierarchical stage in the model with log-normally distributed claim amounts and its corresponding state space version. This way, we incorporate information from both the outstanding claim amounts and counts data resulting in new model formulations. Implementation details and illustrations with real insurance data are provided.

1. Introduction

Because of various administrative reasons, insurance companies often do not pay outstanding claims as soon as they occur. Instead, claims are settled with a time delay that may be years or, in some extreme cases, decades. Reserving for outstanding claims is of central interest in actuarial practice and has attracted the attention of many researchers because of the challenging stochastic uncertainties involved. In this paper, we investigate possible extensions of models that deal with outstanding liabilities. These models assume that the outstanding claims will be considered to be only the "reported but not settled" claims, the year of notification is the same as the accident year, and there are no partial payments. We were motivated by our real data example described in Section 4, which refers to car accidents in Greece. Under Greek legislation, these claims must be reported within three working days, and the proportion of claims that are paid in more than one installment is minimal. Other cases such as claims settled with subpayments or "incurred but not reported" claims, including claims that have been already settled but may be reopened, are not considered in this paper.

Mathematically, the problem can be formulated as follows. There exist data with a structure given by Table 1. A_i, $i = 1, 2, \ldots, r$ denote the years of accident notification and B_j, $j = 1, 2, \ldots, r$ denote the number of years from notification to settlement with 1 indicating settlement in the same year. For example, the cell ij contains the amount Y_{ij} that the company paid with a delay of $j - 1$ years for accidents reported during the year i. Furthermore, the numbers of claims for each cell are also available for certain types of insurance claims. The claim counts have the same triangular form as the claim amounts; in Table 2, n_{ij} represents claim counts that an insurance company paid with a delay of $j - 1$ years for accidents reported in year i, and T_i denotes the total number of accidents that have or will be observed. Finally, the inflation factor for each cell is denoted by inf_{ij}; it is used to deinflate the claim amounts and is also assumed to be known.

Many models and techniques have been proposed to predict the lower triangles of Tables 1 or 2. For classical statistics approaches, see Taylor and Ashe (1983), De Jong and Zehnwirth (1983), Renshaw (1989, 1994), Verrall (1989, 1991, 1993, 1994, 1996), Renshaw and Verrall (1998), and Haberman and Renshaw (1996). For a Bayesian perspective, see Jewell (1989), Verrall (1990), Makov et al. (1996), Haastrup and Arjas (1996), Alba et al. (1998), and Scollnik (2001).

[1] Ioannis Ntzoufras is a Lecturer, Department of Business Administration, University of the Aegean, 8 Michalon Street, 82100, Chios, Greece.

[2] Petros Dellaportas is an Associate Professor, Department of Statistics, Athens University of Economics and Business, 76 Patission Street, 10434 Athens, Greece. e-mail: petros@aueb.gr.

Table 1
Structure of Claim Amounts Data

		B				
		B_1	B_2	…	B_{r-1}	B_r
A	A_1	Y_{11}	Y_{12}	…	$Y_{1,r-1}$	Y_{1r}
	A_2	Y_{21}	Y_{22}	…	$Y_{2,r-1}$	
	$\vdots$	$\vdots$	$\vdots$			
	A_{r-1}	$Y_{r-1,1}$	$Y_{r-1,2}$			
	A_r	Y_{r1}				

In this paper, we propose a Bayesian approach to investigate various models for the outstanding claims problem. Using the Bayesian paradigm did not emanate from the need to use prior information as in Verrall (1990) but, rather, from its computation flexibility that allows us to handle complex models. For example, Models 2 and 4 proposed in Sections 3.2 and 3.4, respectively, would be extremely laborious with a non-Bayesian treatment. Markov chain Monte Carlo (MCMC) sampling strategies are used to generate samples for each posterior distribution of interest. This approach strengthens our Bayesian choice, because as it will be evident in the presentation of the real data analysis in Section 4, posterior summaries of interest can be presented in a very appealing graphical way that provides the best insight. A key feature in the modelling approach we propose is the simultaneous use of data in Tables 1 and 2. In this way, we are essentially modelling "payments per claim finalized" (PPCF) in delay year j having origin in year i; see Taylor and Ashe (1983). This has the advantage that an increase in accidents, which is expressed via an increase of claim counts in Table 2, results in an increase of total claims in Table 1. Modelling of both claim counts and amounts has been advocated in the past; see Norberg (1986), Hesselager (1991), Neuhaus (1992), and Haastrup and Arjas (1996) for application in individual claim data and Taylor and Ashe (1983) for aggregated data.

Because T_i are assumed to be known, we model the claim counts for year i using a multinomial distribution. If T_i are not known, the assumption is easily relaxed, and the claim counts data can be modelled as Poisson distributed. For a related modelling approach, see Verall (2000).

In the next section, we briefly describe the Bayesian theory and the MCMC methodolody. Section 3 presents the models under consideration. In Section 4, an illustrated example with data from an insurance company is presented. Finally, a brief discussion is given in Section 5.

2. BAYESIAN MODELLING VIA MCMC

The approach in this paper is Bayesian, and the tool adopted to implement the Bayesian paradigm is MCMC methodology. In summary, MCMC algorithms construct Markov chains that have as their stationary distribution the required posterior distribution of the parameters. Thus, simulation of such chains provides, when convergence is reached, a (dependent) sample of the posterior distribution that can be used for any posterior summary statistics of interest. Details of the technicalities involved in MCMC can be found, for example, in Smith and Roberts (1993), Besag and Green (1993), and Gilks et al. (1996). MCMC algorithms have appeared in Actuarial literature by Carlin (1992), Rosenberg and Young (1999), and Scollnik (1998, 2001). To avoid extended discussion on MCMC, the following section focuses only on the modelling aspects of the problem and implementation details for one of the models proposed (form of the full conditional densities, Metropolis Hastings sampling, etc.) are given in the Appendix. Technicalities used include sampling from the usual log-concave densities met in generalized linear models (Dellaportas and Smith, 1993), MCMC strategies for problems with constrained parameter problems (Gelfand et al., 1992b), ways to handle the missing values problem (Gelfand et al., 1990), and finally, usual transformation of the MCMC output to the actual parameters of interest. For the latter, let us give some additional information. For every model,

Table 2
Structure of Claim Counts Data

		B					
		B_1	B_2	…	B_{r-1}	B_r	
A	A_1	n_{11}	n_{12}	…	$n_{1,r-1}$	n_{1r}	T_1
	A_2	n_{21}	n_{22}	…	$n_{2,r-1}$		T_2
	$\vdots$	$\vdots$	$\vdots$				$\vdots$
	A_{r-1}	$n_{r-1,1}$	$n_{r-1,2}$				T_{r-1}
	A_r	n_{r1}					T_r

Table 3
Deinflated Claim Amounts in Thousands Drachmas
(Partial Total: Amounts of Claim Settled until 1995)

	Year	B 1	2	3	4	5	6	7	Partial Total
A	1989	527003	183260	90539	50471	37875	10110	21165	920423
	1990	715247	341364	166001	99845	108648	91958		1073129
	1991	810593	257122	87318	72538	81336			1308907
	1992	1012181	339664	156531	181253				1689629
	1993	1459202	448651	188698					2096551
	1994	1689751	563683						2253434
	1995	1698534							1698534

the missing entries in Table 1 (Y_{ij}, $i + j > r + 1$) are treated as parameters. Having obtained the MCMC output samples, it is straightforward to obtain the posterior sample for the parameters of interest $Y_{i.} = \sum_{j=r+2-i}^{r} Y_{ij}$ which express the outstanding claims of year i by adding the generated values of Y_{ij}, $i + j > r + 1$.

An important issue in MCMC implementation is convergence diagnostics. Apart from the usual informal, but necessary, visual checks of the MCMC output which include experiments with different dispersed starting values, we applied three different convergence diagnostics that indicated convergence has been achieved. For details about this important aspect of MCMC methodology, see the recent review papers by Cowles and Carlin (1996) and Brooks and Roberts (1999).

3. Modelling Approaches

In this section, we propose new modelling aspects expressed in terms of four models of the outstanding claims problem. For comparison purposes, we present a Bayesian analysis of two models used in the past, the log-normal (Model 1) and the state space model (Model 3). We enhance these models by simultaneously modelling claim amounts and counts, using the total claim counts to specify appropriate parameter constraints. These modifications result in Models 2 and 4.

3.1 Model 1: Log-Normal Model

The simplest model for the data in Table 1 is a log-normal (anova-type) model. This model was investigated by Renshaw (1989), Verrall (1991, 1993, 1996), and Renshaw and Verrall (1998). Verrall (1990) produced Bayes estimates for the parameters of this model. The model is given by the formulation

$$Z_{ij} = \log \frac{Y_{ij}}{inf_{ij}}, Z_{ij} \sim N(\mu_{ij}, \sigma^2),$$

$$\mu_{ij} = \mu + \alpha_i + \beta_j, \quad i, j = 1, \ldots, r \quad (1)$$

where Z_{ij} are called log-adjusted claim amounts and $N(\mu_{ij}, \sigma^2)$ denotes the normal distribution with mean μ_{ij} and variance σ^2. In Taylor and Ashe (1983) and Verrall (1991, 1993, 1996), the alternative parameterization $Z_{ij} = \log(Y_{ij}) - log(inf_{ij} \times E_i)$ with $Z_{ij} \sim N(\mu_{ij}, \sigma^2)$ is used, where E_i is a measure of exposure (for example size of portfolio for year i). This reparametrization can be adopted for all following models. Finally, (1) requires appropriate constraints to achieve identifiability of the parameters, so here we adopt the usual corner constraints, that is $\alpha_1 = \beta_1 = 0$. Consequently, expression (1) assumes that the expected log-adjusted claim amount, μ_{ij}, originated in year i and was paid with a delay of $j - 1$ years, is modelled via a linear predictor that consists of the expected log-adjusted claim amount μ originated at and paid within the first accident notification year, a factor which reflects expected changes because of the origin year α_i and a factor depending on the delay pattern β_j. These constraints imply that the parameters to be estimated are μ, σ^2, $\alpha = (\alpha_2, \ldots, \alpha_r)^T$, $\beta = (\beta_2, \ldots, \beta_r)^T$, and $Z^L = \{\log(Y_{ij}) : i + j > r + 1\}$.

Table 4
Claim Counts (Partial Total: Totals of Claims Settled until 1995; T_i: Total of Claims Occurred)

	Year	B 1	2	3	4	5	6	7	Totals Partial	T_i
A	1989	6622	1943	489	138	61	223	66	9542	9542
	1990	6943	2133	632	154	162	390		10414	10496
	1991	8610	2216	736	651	256			12469	12601
	1992	9791	3167	1570	624				15152	15565
	1993	11722	3192	1773					16687	17735
	1994	13684	3664						17348	19746
	1995	13068							13068	18600

To complete the Bayesian formulation we specify the priors

$$\mu \sim N(0, \sigma_\mu^2),\ \alpha_i \sim N(0, \sigma_{\alpha_i}^2),\ \beta_j \sim N(0, \sigma_{\beta_j}^2),$$
$$i, j = 2, \ldots, r,\ \tau = \sigma^{-2} \sim G(a_\tau, b_\tau)$$

with $G(a, b)$ denoting the gamma distribution with probability density function $f(\tau) = b^a\tau^{a-1}e^{-b\tau}/\Gamma(a)$. For the kind of problems we are interested in, vague diffuse proper priors (Kass and Wasserman, 1996) are produced by using

$$\sigma_\mu^2 = 1000,\ \sigma_{\alpha_i}^2 = 100,\ i = 2, \ldots, r,$$
$$\sigma_{\beta_j}^2 = 100, j = 2, \ldots, r,\ a_\tau = b_\tau = 0.001.$$

A disadvantage of the above model is that it does not use any information from the observed counts. That is, any prediction of the missing claim amounts will be based only on the observed claim amounts. As a result, a source of information for a year (or cell), such as a sudden increase of accidents, will not affect the prediction of the claim amounts.

3.2 Model 2: Log-Normal and Multinomial Model

Here, we suggest a two stage hierarchical model that uses the data in Tables 1 and 2. Assuming $n_{ij} > 0$ for all i, j, this model can be written as

$$Z_{ij} = log\,\frac{Y_{ij}}{inf_{ij}},\ Z_{ij} \sim N(\mu_{ij}, \sigma^2),$$
$$\mu_{ij} = \mu + \alpha_i + \beta_j + log(n_{ij}),\ (n_{i1}, n_{i2}, \ldots, n_{ir})^T$$
$$\sim \text{Multinomial}(\pi_1, \pi_2, \ldots, \pi_r; T_i),$$
$$\log(\pi_j/\pi_1) = \beta_j^* \quad (2)$$

where $(n_{i1}, n_{i2}, \ldots, n_{ir})^T$ are the number of claims originated at year i, and π_j is the probability for a claim to be settled with a delay of $j - 1$ years. For both stages of the model, we use, as in Model 1, corner constraints. Compared to (1), the linear predictor in the first stage has been enhanced with the term $\log(n_{ij})$. As a result, μ represents the expected log-adjusted amount per claim finalized originating at and paid within the first accident notification year, and α_i, β_j reflect expected differences from μ due to the origin and delay years, respectively. For the second stage, β_j^* represents the log-odds of an accident to be paid with a delay of $j - 1$ years versus an accident paid without any delay.

The second (multinomial) stage of Model 2 is equivalent to the log-linear model

$$n_{ij} \sim Poisson(\lambda_{ij}),\ \log(\lambda_{ij}) = \mu^* + \alpha_i^* + \beta_j^*$$

under the constraints $\sum_{j=1}^{r} n_{ij} = n_{i.} = T_i$ (which also

Table 5
Inflation Factor

Year	1989	1990	1991	1992	1993	1994	1995	1996
Inflation (%)	100.0	120.4	143.9	166.6	190.6	214.2	235.6	257.0

Table 6
Summary Diagnostics

Model	Diagnostics		Length Considered		
	Geweke[1]	Raftery and Lewis[2]	Burn-in	Lag	Sample Size
1	1.54	13606	1,000	2	20,000
2	1.68	8920	1,000	5	20,000
3	1.86	3913	10,000	100	20,000
4	2.01	16016	10,000	1000	20,000

[1] Maximum absolute value over all parameters for Geweke diagnostic.
[2] Maximum iterations required for the 2.5% percentile.

imply that $\sum_{j=1}^{r} (\lambda_{ij} = \lambda_{i.} = T_i)$, where μ^* and α_i^* are nuisance parameters; for more details, see Agresti (1990). The two-stage formulation of this model results in parameters for estimation μ, σ^2, $\alpha = (\alpha_2, \ldots, \alpha_r)^T$, $\beta = (\beta_2, \ldots, \beta_r)^T$, $Z^L = \{log(Y_{ij}) : i + j > r + 1\}$, $\beta^* = (\beta_2^*, \ldots, \beta_r^*)^T$ and $N^L = \{n_{ij} : i + j > r + 1\}$.

Under the assumption that $n_{ij} > 0$ for all i, j, it is precise to assume that n_{ij} follows a "truncated at zero" $Poisson(\lambda_{ij})$. However, for the size of the data we are interested in, the above distribution is practically identical to $Poisson(\lambda_{ij})$. Had we assumed that T_i is unknown, we would have used the above log-linear model without constraints on λ_{ij} and n_{ij}. This could be useful, for example, if some kind of exposure measure is available, say size of portfolio. Then, Model 2 without the constraints on λ_{ij} and n_{ij} is appropriate for predicting "incurred but not reported claims."

We suggest similar prior distributions as Model 1

$$\mu \sim N(0, \sigma_\mu^2),\ \alpha_i \sim N(0, \sigma_{\alpha_i}^2),\ i = 2, \ldots, r,$$
$$\beta_j \sim N(0, \sigma_{\beta_j}^2),\ j = 2, \ldots, r,$$
$$\tau = \sigma^{-2} \sim G(a_\tau, b_\tau),\ \beta_j^* \sim N(0, \sigma_{\beta_j^*}^2),\quad j = 2, \ldots, r.$$

The same values for σ_μ^2, $\sigma_{\alpha_i}^2$, $\sigma_{\beta_j}^2$ as in Model 1 can be used. For the additional parameters β_j^*, we suggest $\sigma_{\beta_j^*}^2 = 100$, for $j = 2, \ldots, r$. Note that the interpretation of parameters α and β are different in Models 1 and 2, so if informative priors were used, the parameters of the prior distributions should not be the same but adjusted accordingly.

3.3 Model 3: State Space Modelling of Claim Amounts

An alternative modelling perspective for this kind of problem is the state space (or dynamic linear) models where the parameters depend on each other in a time-recursive way. A general description of MCMC in dynamic models is given by Gamerman (1998). Carter and Kohn (1994) describe how to use a Gibbs sampler for general state space models, and Carlin (1992) applies a Gibbs sampler for state space models for actuarial time series. For an application of state space models in the claim amounts problem, see De Jong and Zehnwirth (1983) and Verrall (1989, 1994). The state space model can be written as

$$Z_{ij} = log\,\frac{Y_{ij}}{inf_{ij}},\ Z_{ij} \sim N(\mu_{ij}, \sigma^2),\ \mu_{ij} = \mu + \alpha_i + \beta_{ij} \tag{3}$$

Table 7
Posterior Mean and Standard Deviation for Total Outstanding Claim Amounts of Each Accident Year (Million Drachmas; Adjusted for Inflation)

Model	Year					
	1990	1991	1992	1993	1994	1995
1	34 (16)	66 (23)	214 (63)	404 (114)	763 (227)	1430 (629)
2	31 (20)	13 (6)	95 (37)	300 (114)	635 (250)	1229 (621)
3	30 (13)	76 (31)	220 (76)	495 (199)	847 (337)	1522 (558)
4	31 (17)	20 (11)	107 (48)	319 (151)	538 (246)	1080 (538)

Table 8
Posterior Mean and Standard Deviation of Total Claim Amounts to be Paid in Each Future Year (Million Drachmas; Adjusted for Inflation)

Model	Year						Total
	1996	1997	1998	1999	2000	2001	
1	1225 (333)	674 (305)	470 (134)	299 (102)	153 (62)	89 (57)	2911 (730)
2	1072 (407)	576 (214)	369 (155)	185 (91)	65 (34)	36 (29)	2303 (718)
3	1226 (310)	743 (270)	576 (329)	357 (238)	191 (174)	95 (117)	3188 (1014)
4	923 (309)	480 (215)	357 (333)	193 (177)	92 (139)	50 (161)	2097 (852)

with the recursive associations

$$\beta_{ij} = \beta_{i-1,j} + v_i,\ v_i \sim N(0,\ \sigma_v^2),\ i = 2,\ \ldots,\ r,$$
$$j = 2,\ \ldots,\ r,\ \alpha_i = \alpha_{i-1} + h_i,$$
$$h_i \sim N(0,\ \sigma_h^2),\ i = 2,\ \ldots,\ r$$

and corner constraints $\alpha_1 = \beta_{i1} = 0, i = 1, 2, \ldots, r$. This model formulation implies that $\mathrm{Var}(Z_{ij}) = \sigma^2$ if $i = 1$ or $j = 1$ and $Var(Z_{ij}) = \sigma^2 + (i - 1)(\sigma_v^2 + \sigma_h^2)$ for $i > 1$ and $j > 1$. The parameters of this model are μ, σ^2, σ_ν^2, σ_h^2, $\alpha = (\alpha_2, \ldots, \alpha_r)^T$, $\beta^T = (\beta_2^T, \ldots, \beta_r^T)$, where $\beta_i^T = (\beta_{i2}, \ldots, \beta_{ir})$ for $i = 2, \ldots, r$, and $Z^L = \{log(Y_{ij}) : i + j > r + 1\}$.

Comparing Model 3 with Model 1, we first note that β_j has been replaced by β_{ij}. Thus, the delay effect on the log-adjusted claim amounts changes with the origin year. Second, the introduced recursive associations express the belief that the parameters α_i and β_j evolve through time via known stochastic mechanisms. In fact, these mechanisms are determined by the disturbance terms v_i and h_i; as σ_v^2 approaches zero, (3) resem-

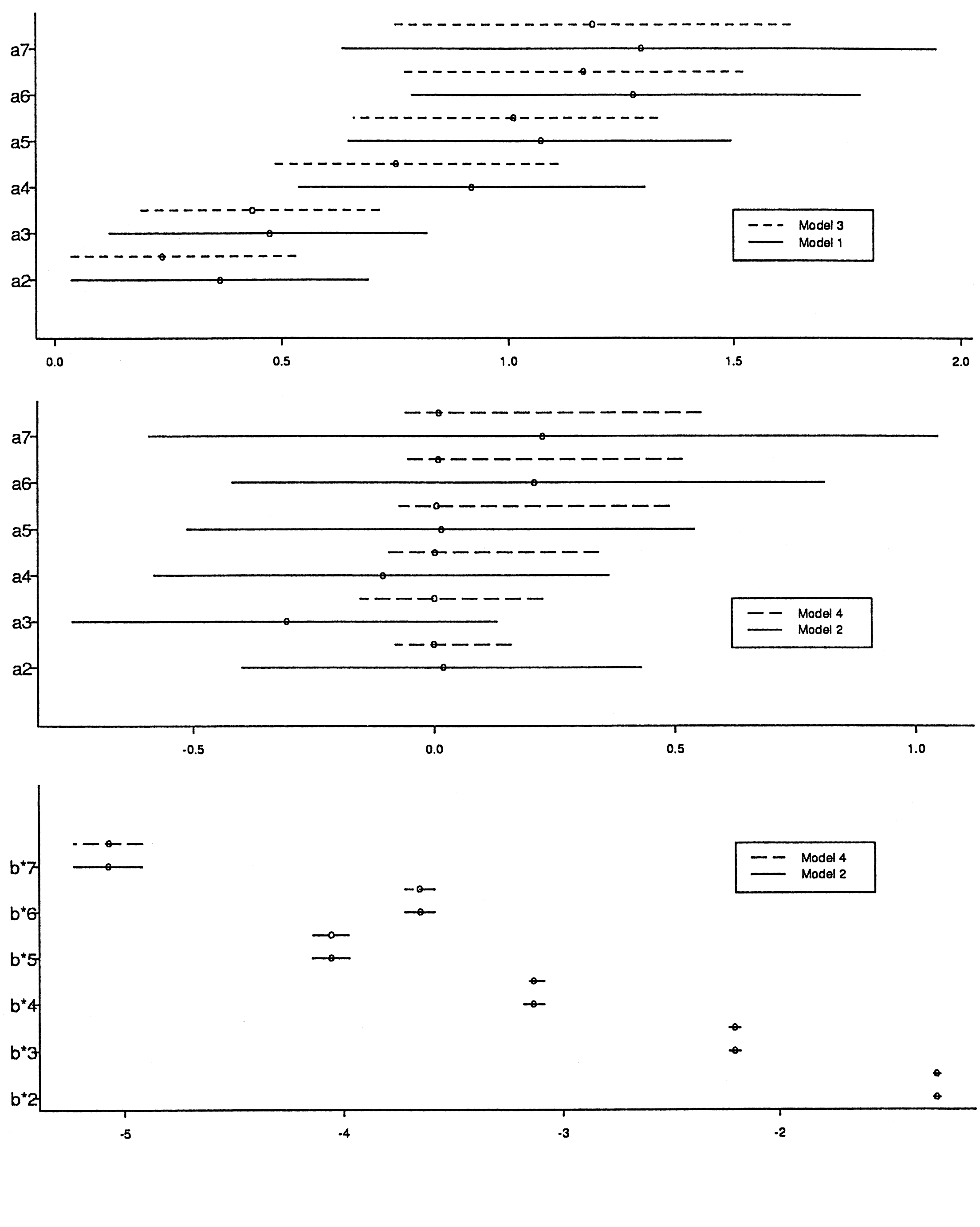
Table 9
Posterior Mean (Standard Deviation) for Model Parameters and L-criterion for Each Model

Model	μ	σ^2	σ_ν^2	σ_h^2	$\hat{L}_m$
1	13.05(0.17)	0.08(0.04)			536,757
3	13.14 (0.13)	0.04 (0.04)	0.07 (0.09)	0.10 (0.11)	533,273
2	4.64 (0.20)	0.13 (0.05)			603,936
4	4.56 (0.15)	0.07 (0.06)	0.12 (0.14)	0.01 (0.02)	582,045

Model		$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$
1	α_j	0.36 (0.17)	0.47 (0.18)	0.92 (0.19)	1.07 (0.21)	1.28 (0.25)	1.29 (0.33)
3		0.25 (0.13)	0.44 (0.13)	0.77 (0.16)	1.01 (0.16)	1.16 (0.18)	1.19 (0.21)
2		0.02 (0.20)	−0.31 (0.22)	−0.11 (0.24)	0.01 (0.27)	0.20 (0.31)	0.23 (0.41)
4		0.02 (0.06)	0.04 (0.09)	0.08 (0.13)	0.14 (0.19)	0.16 (0.20)	0.17 (0.22)
1	β_j	−1.08 (0.16)	−1.94 (0.18)	−2.24 (0.19)	−2.43 (0.21)	−3.33 (0.25)	−3.09 (0.33)
3	β_{1j}	−1.06 (0.15)	−1.85 (0.18)	−2.30 (0.16)	−2.54 (0.18)	−3.61 (0.30)	−3.17 (0.25)
	β_{2j}	−1.03 (0.15)	−1.90 (0.16)	−2.36 (0.18)	−2.49 (0.17)	−3.12 (0.32)	−3.18 (0.36)
	β_{3j}	−1.10 (0.14)	−2.05 (0.18)	−2.31 (0.16)	−2.39 (0.19)	−3.12 (0.42)	−3.17 (0.45)
	β_{4j}	−1.11 (0.14)	−1.94 (0.15)	−2.04 (0.27)	−2.38 (0.33)	−3.12 (0.50)	−3.17 (0.53)
	β_{5j}	−1.12 (0.14)	−1.99 (0.15)	−2.04 (0.38)	−2.39 (0.43)	−3.12 (0.56)	−3.17 (0.59)
	β_{6j}	−1.09 (0.15)	−1.99 (0.31)	−2.04 (0.47)	−2.39 (0.51)	−3.12 (0.62)	−3.17 (0.65)
	β_{7j}	−1.09 (0.31)	−1.99 (0.41)	−2.04 (0.54)	−2.39 (0.57)	−3.12 (0.68)	−3.17 (0.71)
2	β_j	0.17 (0.20)	0.30 (0.22)	0.99 (0.24)	1.43 (0.26)	−0.44 (0.31)	1.13 (0.41)
4	β_{1j}	0.10 (0.18)	0.49 (0.31)	1.16 (0.32)	1.63 (0.36)	−0.52 (0.23)	1.21 (0.30)
	β_{2j}	0.15 (0.18)	0.38 (0.23)	1.08 (0.27)	1.30 (0.19)	−0.19 (0.33)	1.22 (0.45)
	β_{3j}	0.16 (0.17)	0.22 (0.17)	0.58 (0.35)	1.27 (0.20)	−0.19 (0.47)	1.22 (0.57)
	β_{4j}	0.09 (0.17)	0.09 (0.21)	0.93 (0.19)	1.27 (0.40)	−0.19 (0.58)	1.22 (0.67)
	β_{5j}	0.17 (0.17)	0.06 (0.23)	0.92 (0.40)	1.27 (0.52)	−0.19 (0.67)	1.22 (0.75)
	β_{6j}	0.20 (0.18)	0.06 (0.41)	0.92 (0.53)	1.28 (0.63)	−0.19 (0.75)	1.22 (0.83)
	β_{7j}	0.21 (0.39)	0.06 (0.54)	0.93 (0.63)	1.28 (0.72)	−0.19 (0.83)	1.22 (0.89)
2	β_j^*	−1.28 (0.01)	−2.20 (0.01)	−3.13 (0.02)	−4.06 (0.04)	−3.65 (0.03)	−5.07 (0.08)
4	β_j^*	−1.28 (0.01)	−2.20 (0.01)	−3.13 (0.02)	−4.05 (0.04)	−3.65 (0.03)	−5.07 (0.08)

Figure 1

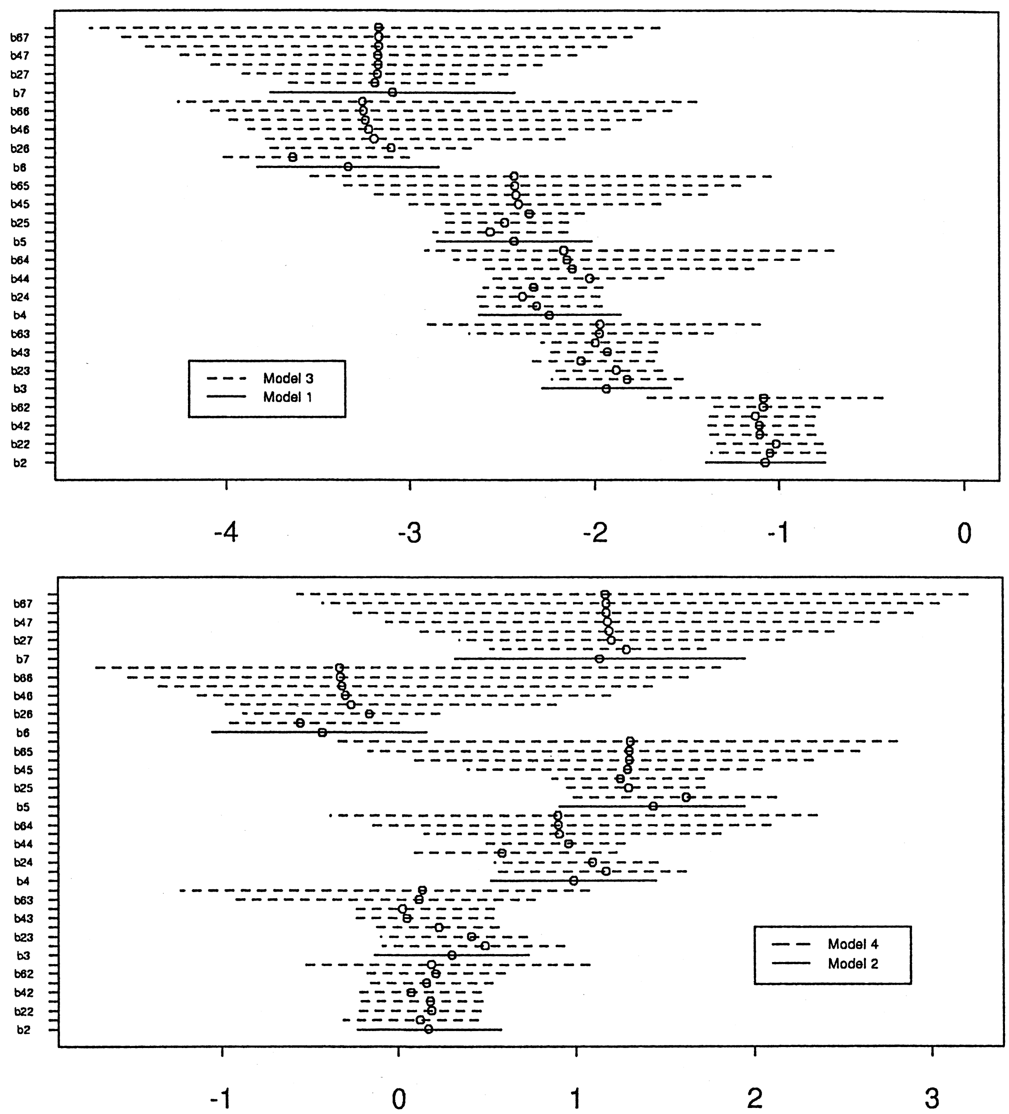
95% Posterior Credible Intervals for α and β^*; Dots Indicate the Posterior Median.



bles Model 1; whereas, when σ_h^2 approaches zero, the parameters α_i tend to zero. The corner constraints imply that μ is the expected log-adjusted claim amount for the first origin year paid without delay, and α_i and β_{ij} are interpreted accordingly.

In (3), we only need to define prior distributions for the first state space parameters; for more details, see Carlin et al. (1992), Carlin (1992), and Gamerman (1998). We propose priors $\beta_{1j} \sim N(0, \sigma^2_{\beta_{1j}})$ and $\mu \sim N(0, \sigma^2_{\mu})$ with $\sigma^2_{\beta_{1j}} = 100$ and $\sigma^2_{\mu} = 1000$. The prior for the precision $\tau = \sigma^{-2}$ is a $G(a_\tau, b_\tau)$ density as in Model 1. We additionally use non-informative Gamma priors for the parameters $\sigma_\nu^{-2} \sim G(a_\nu, b_\nu)$ and $\sigma_h^{-2} \sim G(a_h, b_h)$ with proposed values $a_\nu = b_\nu = a_h = b_h = 10^{-10}$. Finally, as in Model 1, we note that

Figure 2
95% Posterior Credible Intervals for β; Dots Indicate the Posterior Median.



this model does not use any information from claim counts.

3.4 Model 4: State Space Modelling of Average Claim Amount per Accident

Here we generalize Model 3 by incorporating information from data in Table 2. Assuming that $n_{ij} > 0$ for all i, j, we suggest

$$Z_{ij} = log\, \frac{Y_{ij}}{inf_{ij}}, Z_{ij} \sim N(\mu_{ij}, \sigma^2),$$

$$\mu_{ij} = \mu + \alpha_i + \beta_{ij} + log(n_{ij}), (n_{i1}, n_{i2}, \ldots, n_{ir})^T$$

$$\sim \text{Multinomial}(\pi_1, \pi_2, \ldots, \pi_r; T_i),$$

$$\log(\pi_j/\pi_1) = \beta_j^*, \beta_1^* = 0 \quad (4)$$

with the recursive associations

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

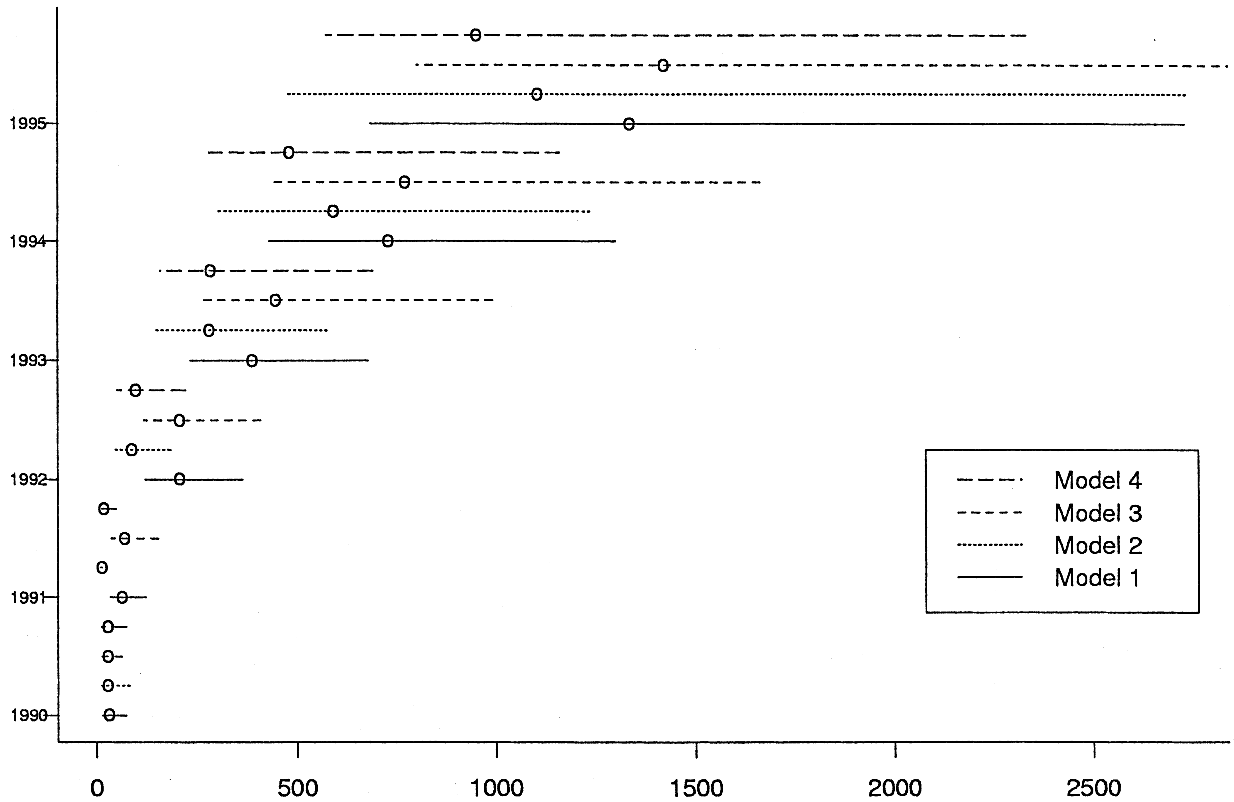
$$i = 2, \ldots, r, j = 2, \ldots, r,$$

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

The variance of Z_{ij} for $i, j > 1$ conditional on n_{ij} is given by $\text{Var}(Z_{ij}|n_{ij}) = \sigma^2 + (i-1)(\sigma_v^2 + \sigma_h^2)$. In analogy with Model 2, we have added the term

Figure 3

95% Posterior Credible Intervals for Total Outstanding Deinflated Claim Amounts (in Million Drachmas) for Each Accident Year; Dots Indicate the Posterior Median.



$log(n_{ij})$ in the linear predictor. Thus, μ represents the log-adjusted amount per claim finalized for the first origin year paid without delay, and α_i, β_{ij} are interpreted accordingly. The multinomial second stage formulation is interpreted exactly as in Model 2. Thus, the parameters of the model are μ, σ^2, σ_v^2, σ_h^2, $\alpha = (\alpha_2, \ldots, \alpha_r)^T$, $\beta^T = (\beta_2^T, \ldots, \beta_r^T)$, where $\beta_i^T = (\beta_{i2}, \ldots, \beta_{ir})$ for $i = 2, \ldots, r$, and $Z^L = \{log(Y_{ij}) : i + j > r + 1\}$, $\beta^* = (\beta_2^*, \ldots, \beta_r^*)^T$, and $N^L = \{n_{ij} : i + j > r + 1\}$.

The priors can be defined similarly as in Models 2 and 3.

3.5 Model Selection

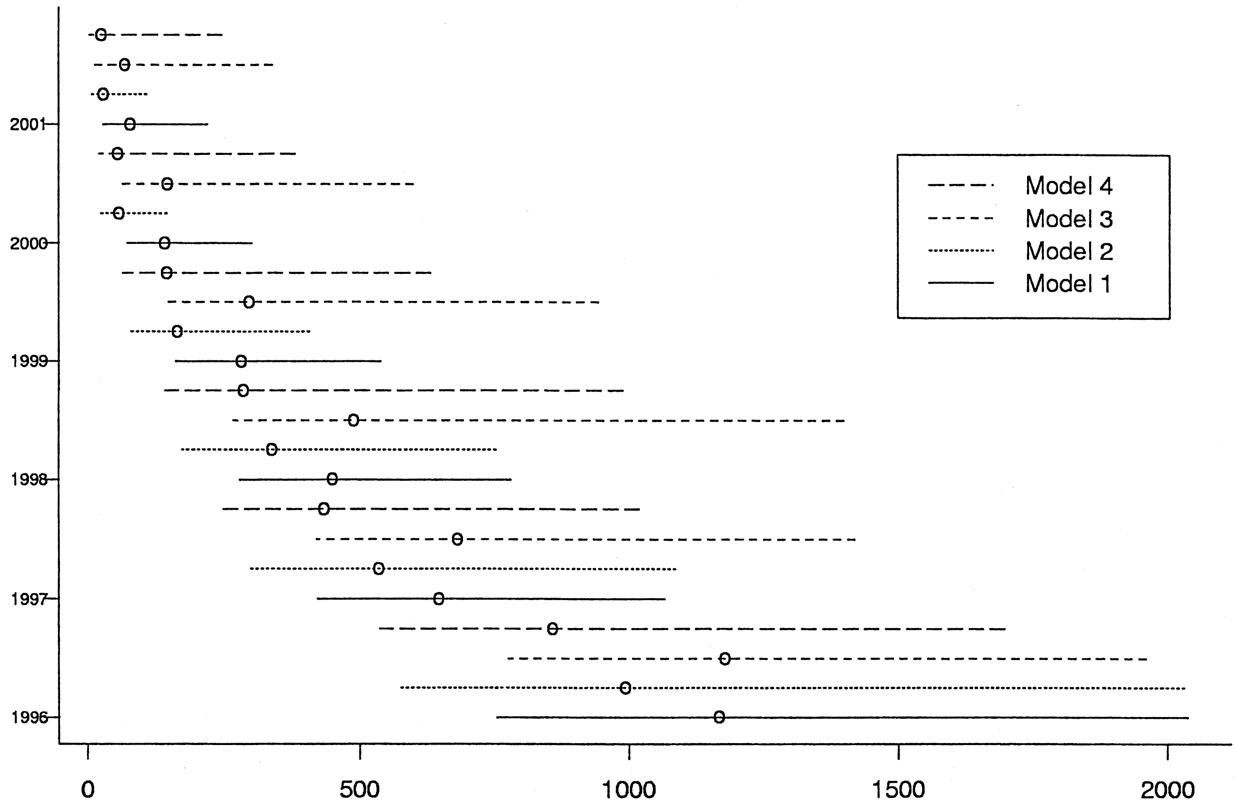
Although emphasis in this paper is firmly on drawing inferences assuming one of the above proposed models, it is of particular interest to comment on the ability to choose the most appropriate model for a given data set. It would be appealing to state that Models 2 and 4 are preferable to Models 1 and 3 because they incorporate additional information via the claim counts data set. However, it is clear that the focus in such applications is not data fitting but data prediction. Thus, the only sensible way to compare models is to compare their predictive ability; this can only be achieved with access to a series of complete tables with all entries of claim counts and amounts available.

One way to compare models based on their predictive performance is to use the predictive measures based on the L-criterion of Laud and Ibrahim (1995). Assume that all claim amounts were available, but only the top left triangular entries of Tables 1 and 2 were used for the estimation of the outstanding liabilities. The MCMC estimated Laud-Ibrahim criterion selects the model that minimizes

$$\hat{L}_m^2 = \frac{1}{K-b} \sum_{k=b+1}^{K} \sum_{(i,j)\in\mathcal{L}} ([y_{ij}^{(m)}]_k - y_{ij})^2$$

Figure 4

95% Posterior Credible Intervals for Total Outstanding Deinflated Claim Amounts (in Million Drachmas) to be Paid in Each Year; Dots Indicate the Posterior Median.



where $[y_{ij}^{(m)}]_k$ are the MCMC generated missing claim amounts of cell ij in iteration k resulted from model m, $\mathscr{L}$ includes the cells (i, j) that represent available future observations (usually the lower triangular matrix of Table 1) which are not available at the time of estimation. K is the total number of iterations, and b is the number of iterations discarded as burn-in.

A more formal approach to the predictive type criteria was given by Gelfand and Ghosh (1998) who adopt a completely formal decision-theoretic setting. See also Gelfand et al. (1992a) for a series of other diagnostics based on predictive performance.

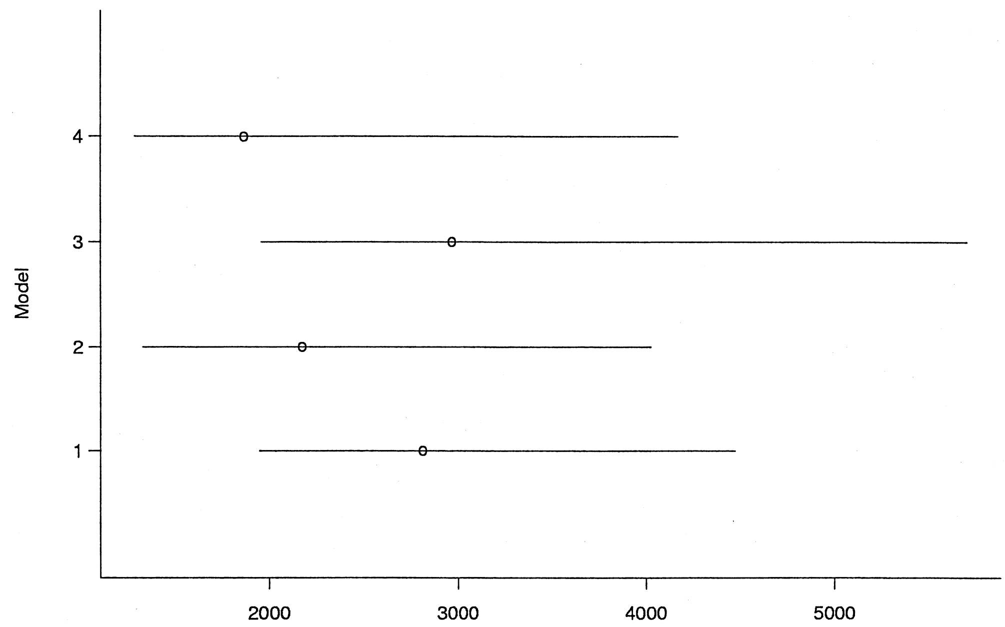
4. Illustrated Example

The following data came from a major insurance company in Greece. Tables 3 to 5 give the deinflated claim amounts, claim counts, total counts for car accidents, and inflation factors used by the insurance company and provided by the bank of Greece. Because of their nature, the main source of delay is due to claims that are notified but settled after the accident year. Liabilities that have arisen but reported later are assumed to be minimal. Moreover, for our data, the assumption of no partial payments is plausible because only a small proportion of car accident claims are paid in more than one installment.

The analysis of the data above was initiated by deinflating the original claim amounts data using the inflation factors of Table 5. Therefore, the resulting predictive amounts presented in this section should be multiplied by the corresponding inflation factor to represent amounts for a specific year (for example multiply by 257/100 = 2.57 to get the inflated amount for year 1996).

To save computer space, we kept only lagged values of the Markov chains. These values were chosen after inspecting the autocorrelation func-

Figure 5
95% Posterior Credible Intervals for Total Outstanding Deinflated Claim Amounts in Million Drachmas; Dots Indicate the Posterior Median.



tion of the parameters of the initial run, and they represent the lag intervals with which the sample should be collected so that the resulting sample is nearly non-correlated. The convergence diagnostics proposed by Geweke (1992), Heidelberg and Welsch (1983), and Raftery and Lewis (1992) indicated convergence; see Table 6 for summary values.

Posterior summaries of Models 1 through 4 are given in Tables 7 and 8. Note the striking difference of our proposed Models 2 and 4 when compared with the existing approaches expressed by Models 1 and 3 for outstanding claim amounts for 1991 and 1992. This deviation is easily explainable if we carefully examine the data in Table 4. The remaining outstanding claims for 1991 are only 132 and account for the 1.05% of the total claim counts (12,601). This percentage is comparably much smaller than the corresponding outstanding claim counts of 1989 and 1990, which were 3.03 and 4.48%, respectively. This decrease is being taken into account by our models, and the produced estimates for 1991 are appropriately adjusted.

In fact, Models 2 and 4 give, for all years, lower predictions than Models 1 and 3. A careful look at Table 4, which Models 1 and 3 do not take into account, reveals that the additional information of the claim count pattern indicates a decrease in the company's outstanding liabilities. As a further example, consider years 1994 and 1995. Although the deinflated claim amounts reported in these years are about the same (1689751 and 1698534, respectively; see Table 3), there is a considerable decrease in total unsettled claim counts after the first delay year ($19746 - 13684 = 6062$ and $18600 - 13068 = 5532$, respectively). Considering that a rough estimate for the mean payments per claim are $1689751/13684 = 123.48$ and $1698534/13068 = 129.98$, respectively, the outstanding liabilities for 1994 are expected to be $6062 \times 123.48 = 748535.8$, about 4% more than the corresponding expected liabilities for year 1995 ($5532 \times 129.98 = 719049.4$). Therefore, Models 2 and 4 reduce the estimated unsettled amounts of year 1995 in comparison to year 1994 while Models 1 and 3 give similar predictions for both years.

Table 9 gives the posterior summaries of the model parameters for all models. Again, all presented summaries refer to corner constrained parameters. Additionally, the posterior summaries (Figures 1–5) provided depict the posterior median, 2.5 and 97.5% percentiles of model parameters and estimated total claim amounts. These provide the best insight of the parameters of interest and emphasize the power of the Bayesian sampling scheme we used; for example, skew posterior densities could not be revealed by techniques employed in classical statistics. For the data we analyzed, we noticed that the state space Models 3 and 4 did not differ very much from the simpler Models 1 and 2. This may informally imply that no dynamic term is needed when modelling the total claim amounts.

Finally, we have compared the predictive performance of our proposed models by using the L-type criterion described in Section 3.5 (see Table 9). We have used real data for the years 1996–2000 from the insurance company; unfortunately, we did not obtain permission to report these data here. The L criterion suggests that, for these data, the predictive ability of Models 1 and 2 seems to be better than that of Models 3 and 4.

5. DISCUSSION

In this paper, we analyzed the well-known problem of insurance outstanding claims companies'. All models were investigated using Bayesian theory and MCMC methodology.

The models fitted can be divided in two categories. The first category contains models that use only the information from claim amounts (Table 1), while the second exploits both claim amounts and counts (Tables 1 and 2). Thus, the enriched family attempts to model the average payment per claim finalized/paid; this is the approach we advocate, and we believe that it improves the model predictive behavior.

The models dealt with in this paper can be generalized by adding other factors in the first (log-normal) stage. For example, we may assume that the variance of Z_{ij} depends on the claim counts of the corresponding cell. Because our suggested models are already multiplicative in the error, this adjustment will improve, at least in our data, only slightly to fit.

Finally, we would like to mention that the Bayesian paradigm used in this paper did not utilize the advantage of using informative prior densities. By illustrating our results with non-informative priors, we only provide a yardstick for comparison with other approaches. However, any prior knowledge can be incorporated in our models using the usual quantification arguments.

ACKNOWLEDGMENTS

We would like to thank Professor Nickos Frangos for introducing us to the problem and the insurance company for providing us the data.

REFERENCES

AGRESTI, A. 1990. "Categorical Data Analysis." John Wiley and Sons, USA.

DE ALBA, E., M. JUARES, AND T.M. MORENO. 1998. "Bayesian Estimation of IBNR Reserves," *Transactions of the International Congress of Actuaries*. Birmingham, England.

BESAG, J., AND P.J. GREEN. 1993. "Spatial Statistics and Bayesian Computation," *Journal of the Royal Statistical Society B* 55: 25–37.

BROOKS, S.P., AND G.O. ROBERTS. 1999. "Diagnosing Convergence of Markov Chain Monte Carlo Algorithms," *Statistics and Computing* 8: 319–35.

CARLIN, B.P. 1992. "State Space Modelling of Non-standard Actuarial Time Series," *Insurance: Mathematics and Economics* 11: 209–22.

CARLIN, B.P., N.G. POLSON, AND D.S. STOFFER. 1992. "A Monte Carlo Approach to Non-normal and Nonlinear State-Space Modelling," *Journal of the American Statistical Association* 87: 493–500.

CARTER, C.K., AND R. KOHN. 1994. "On Gibbs Sampler for State Space Models," *Biometrika* 81: 541–53.

COWLES, M.K. AND B.P. CARLIN. 1996. "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review," *Journal of the American Statistical Association* 91: 883–904.

DE JONG, P., AND B. ZEHNWIRTH. 1983. "Claims Reserving, State Space Models and Kalman Filter," *The Journal of the Institute of Actuaries* 110: 157–81.

DELLAPORTAS, P., AND A.F.M. SMITH. 1993. "Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling," *Applied Statistics* 42: 443–60.

GAMERMAN, D. 1998. "Markov Chain Monte Carlo for Dynamic Generalized Linear Models," *Biometrika* 85: 215–27.

GELFAND, A.E., D. DEY, AND H. CHANG. 1992a. "Model Determination Using Predictive Distributions with Implementation via Sampling Based Methods (with Discussion)," *Bayesian Statistics 4* (J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds.), Oxford: University Press, 147–67.

GELFAND, A.E., AND S.K. GHOSH. 1998. "Model Choice: A Minimum Posterior Predictive Loss Approach," *Biometrika* 85: 1–11.

GELFAND, A.E., S.E. HILLS, A. RACINE-POON, AND A.F.M. SMITH. 1990. "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association* 85: 972–85.

GELFAND, A.E., A.F.M. SMITH, AND T.-M. LEE. 1992b. "Bayesian Analysis of Constrained Parameter and Truncated Data

Problems Using Gibbs Sampling," *Journal of the American Statistical Association* 87: 523–32.

GEWEKE, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments," *Bayesian Statistics 4* (J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds.), Oxford: University Press, 169–93.

GILKS, W.R., S. RICHARDSON, AND D.J. SPIEGELHALTER. 1996. "Markov Chain Monte Carlo in Practice." Chapman and Hall, UK.

GILKS, W.R., AND P. WILD. 1992. "Adaptive Rejection Sampling for Gibbs Sampling." *Applied Statistics* 41: 337–48.

HAASTRUP S., AND E. ARJAS. 1996. "Claims Reserving in Continuous Time; A Nonparametric Bayesian Approach," *Astin Bulletin* 26: 139–64.

HABERMAN, S., AND A.E. RENSHAW. 1996. "Generalized Linear Models and Actuarial Science," *The Statistician* 45: 407–36.

HEIDELBERGER, P., P. AND WELSCH. 1983. "Simulation Run Length Control in the Presence of an Initial Transient," *Operations Research* 31: 1109–44.

HESSELAGER, O. 1991. "Prediction of Outstanding Claims: An Hierarchical Credibility Approach," *Scandinavian Actuarial Journal* 1991, 25–47.

JEWELL, W.S. 1989. "Predicting IBNYR Events and Delays," *Astin Bulletin* 19: 25–55.

KASS, R.E., AND L. WASSERMAN. 1996. "The Selection of Prior Distributions by Formal Rules," *Journal of American Statistical Association* 91: 1343–70.

LAUD, P.W., AND J.G. IBRAHIM. 1995. "Predictive Model Selection," *Journal of the Royal Statistical Society B* 57: 247–62.

MAKOV, U.E., A.F.M. SMITH, AND Y.H. LIU. 1996. "Bayesian Methods in Actuarial Science," *The Statistician* 45: 503–15.

NEUHAUS, W. 1992. "IBNR Models with Random Delay Distributions," *Scandinavian Actuarial Journal* 1992: 97–107.

NORBERG, R. 1986. "A Contribution to Modelling of IBNR Claims," *Scandinavian Actuarial Journal* 1986: 155–203.

RAFTERY, A.L., AND S. LEWIS. 1992. "How Many Iterations in the Gibbs Sampler?" *Bayesian Statistics 4* (J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds.). Oxford: University Press, 763–74.

RENSHAW, A.E. 1989. "Chain Ladder and Interactive Modelling." *The Journal of the Institute of Actuaries* 116: 559–87.

RENSHAW, A. 1994. "On the Second Moment Properties and the Implementation of Certain GLIM Based Stochastic Claims Reserving Models," *Actuarial Research Paper 65.* Department of Actuarial and Statistics, City University, London, UK.

RENSHAW, A., AND R. VERRALL. 1998. "A Stochastic Model Underlying the Chain-Ladder Technique," *British Actuarial Journal* 4: 903–26.

ROSENBERG, M.A., AND V.R. YOUNG. 1999. "A Bayesian Approach to Understanding Time Series Data," *North American Actuarial Journal* 3: No. 2, 130–43.

SCOLLNIK, D.P.M. 1998. "On the Analysis of Truncated Generalized Poisson Distribution Using a Bayesian Method," *Astin Bulletin* 28: No. 1, 135–52.

SCOLLNIK, D.P.M. 2001. "Actuarial Modeling with MCMC and BUGS," *North American Actuarial Journal* 5: No. 2, 96–124.

SMITH, A.F.M., AND G.O. ROBERTS. 1993. "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society B* 55: 3–23.

TAYLOR, G.C., AND F.R. ASHE. 1983. "Second Moments of Estimates of Outstanding Claims," *Journal of Econometrics* 23: 37–61.

VERRALL, R. 1989. "State Space Representation of the Chain Ladder Linear Model," *The Journal of the Institute of Actuaries* 116: 589–610.

VERRALL, R. 1990. "Bayes and Empirical Bayes Estimation for the Chain Ladder Model," *Astin Bulletin* 20: 217–43.

VERRALL, R. 1991. "Chain Ladder and Maximum Likelihood," *The Journal of the Institute of Actuaries* 118: 489–99.

VERRALL, R. 1993. "Negative Incremental Claims: Chain Ladder and Linear Models," *The Journal of the Institute of Actuaries* 120: 171–83.

VERRALL, R. 1994. "A Method of Modelling Varying-off Evolutions in Claims Reserving," *Astin Bulletin* 24: 325–32.

VERRALL, R. 1996. "Claims Reserving and Generalized Additive Models," *Insurance: Mathematics and Economics* 19: 31–43.

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

APPENDIX

The conditional posterior distributions that are needed for the MCMC implementation of Model 2 are given here in detail. The procedure is similar for the other models discussed in the paper. Iterative samples from these conditional densities provide, after some burn-in period and by using an appropriate sample lag, the required samples from the posterior density.

Model 2 includes parameters μ, $\alpha = (\alpha_2, \ldots, \alpha_r)^T$, $\beta = (\beta_2, \ldots, \beta_r)^T$, σ^{-2} from stage one, and $\beta^* = (\beta_2^T, \ldots, \beta_r^*)^T$ from stage two. The claim amounts and counts are divided in known/observed (data) for $i + j \leq r + 1$ and unknown missing (parameters) for $i + j > r + 1$. Denote by N^U and Z^U, the observed chain counts and amounts, respectively, by N^L and Z^L, the missing claim counts and amounts, respectively, and by N and Z, the matrices containing both observed and missing claim counts and amounts, respectively. Assuming that the missing data N^L and Z^L are a further set of parameters, the parameter vector is given by (μ, α, β, σ^{-2}, β^*, N^L, Z^L), and the data vector is given by (N^U, Z^U). Using Bayes theorem and denoting by f the prior, conditional and marginal densities, the posterior distribution is given by

$$
\begin{aligned}
f(\mu, \alpha, \beta, \sigma^{-2}, \beta^*, N^L, Z^L|N^U, Z^U) &\propto \\
& f(N^U, Z^U|\mu, \alpha, \beta, \sigma^{-2}, \beta^*, N^L, Z^L) \\
& f(\mu, \alpha, \beta, \sigma^{-2}, \beta^*, N^L, Z^L) \propto \\
& f(N^U, Z^U|\mu, \alpha, \beta, \sigma^{-2}, \beta^*) \\
& f(N^L, Z^L|\mu, \alpha, \beta, \sigma^{-2}, \beta^*) \\
& f(\mu)f(\alpha)f(\beta)f(\sigma^{-2})f(\beta^*) \propto \\
& f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(Z^L|\mu, \alpha, \beta, \sigma^{-2}, N) \\
& f(N|\beta^*)f(\mu)f(\alpha)f(\beta)f(\sigma^{-2})f(\beta^*) \qquad (5)
\end{aligned}
$$

where $\beta = (\beta_2, \ldots, \beta_r)$, $\alpha = (\alpha_2, \ldots, \alpha_r)$ and $\beta^* = (\beta_2^*, \ldots, \beta_r^*)$. Note that the parameters α_1, β_1, and β_1^* have been removed from the likelihood and the posterior distribution since are set to zero in corner constraints. To make the MCMC algorithm more efficient, we may integrate out all the missing parameters Z^L from (5) and therefore sample all other parameters from

$$
\begin{aligned}
& f(\mu, \alpha, \beta, \sigma^{-2}, \beta^*, N^L|N^U, Z^U) \\
& = \int f(\mu, \alpha, \beta, \sigma^{-2}, \beta^*, N^L, Z^L|N^U, Z^U)dZ^L \propto \\
& f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(N|\beta^*) \\
& \qquad f(\mu)f(\alpha)f(\beta)f(\sigma^{-2})f(\beta^*) \qquad (6)
\end{aligned}
$$

and the missing (future) claim amounts from the conditional predictive density

$$f(Z^L|\mu, \alpha, \beta, \sigma^{-2}, Z^U, N).$$

The full conditional distributions are therefore given by

1. $f(\mu|.) \propto f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(\mu)$
2. $f(\alpha|.) \propto f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(\alpha)$
3. $f(\beta|.) \propto f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(\beta)$
4. $f(\sigma^{-2}|.) \propto f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)f(\sigma^{-2})$
5. $f(Z^L|.) \propto f(Z^L|\mu, \alpha, \beta, \sigma^{-2}, N)$
6. $f(\beta^*|.) \propto f(N|\beta^*)f(\beta^*)$
7. $f(N^L|.) \propto f(N|\beta^*)$

The conditional distribution $f(Z|\mu, \alpha, \beta, \sigma^{-2}, N)$ is the "augmented" likelihood for the first stage, assuming that there are no missing data in the claim amount table while $f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N)$ is the "observed" likelihood of the available data (upper triangle of claim amount data). Therefore,

$$
f(Z|\mu, \alpha, \beta, \sigma^{-2}, N) = (2\pi\sigma^2)^{-r^2/2}\, exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^{r} \sum_{j=1}^{r} [Z_{ij} - \mu - \alpha_i - \beta_j - log(n_{ij})]^2 \right)
$$

and

$$
f(Z^U|\mu, \alpha, \beta, \sigma^{-2}, N) = (2\pi\sigma^2)^{-|U|/2}\, exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^{r} \sum_{j=1}^{r-i+1} [Z_{ij} - \mu - \alpha_i - \beta_j - \log(n_{ij})]^2 \right),
$$

where $|U|$ is the number of cells of the available data (upper triangle) given by $|U| = r(r + 1)/2$.

Then, assuming a multinomial distribution we have for all $i = 1, \ldots, r$

$$
f(n_{i1}, \ldots, n_{ir}|\pi_1, \ldots, \pi_r) = \frac{T_i!}{\prod_{j=1}^{r} n_{ij}!} \prod_{j=1}^{r} \pi_j^{n_{ij}} \qquad (7)
$$

with $n_{ir} = T_i - \sum_{j=1}^{r-1} n_{ij}$ and $\sum_{j=1}^{r} \pi_j = 1$. Thus, the full likelihood $f(N|\beta^*)$ of the second stage, assuming no missing claim counts, can be written as

$$
f(N|\beta^*) = \prod_{i=1}^{r} f(n_{i1}, \ldots, n_{ir}|\pi_1, \ldots, \pi_r).
$$

Using the equation $\log(\pi_j/\pi_1) = \beta_j^*$ and hence $\pi_j = e_j^{\beta^*}/\sum_{k=1}^{r} e_k^{\beta^*}$, we have

$$
f(N|\beta^*) = exp\left(\sum_{i=1}^{r} log(T_i!) - \sum_{i=1}^{r}\sum_{j=1}^{r} log(n_{ij}!) + \sum_{j=2}^{r} n_{.j}\beta_j^* - n_{..}\, log\left(\sum_{k=1}^{r} e_k^{\beta^*} \right) \right),
$$

where $n_{..} = \sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij} = \sum_{i=1}^{r} T_i$ and $n_{.j} = \sum_{i=1}^{r} n_{ij}$. Thus, the resulting conditional distributions are

$$
1. \quad f(\mu|.) = \left(\frac{M}{|U| + \sigma^2/\sigma_\mu^2}, \frac{\sigma^2}{|U| + \sigma^2/\sigma_\mu^2} \right), \qquad (8)
$$

where

$$M = \sum_{i=1}^{r} \sum_{j=1}^{r-i+1} \{Z_{ij} - \alpha_i - \beta_j - \log(n_{ij})\}$$

$$= Z_{..} - \ln_{..} - \sum_{i=1}^{r} (r - i + 1)\alpha_i$$

$$- \sum_{j=1}^{r} (r - j + 1)\beta_j,$$

with $Z_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r-i+1} Z_{ij}$ and $ln_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r-i+1} log(n_{ij})$.

2. $$f(\alpha_i|.) = N\left(\frac{A_i}{(r - i + 1) + \sigma^2/\sigma^2_{\alpha_i}}, \frac{\sigma^2}{(r - i + 1) + \sigma^2/\sigma^2_{\alpha_i}}\right), i = 2, \ldots, r, \quad (9)$$

where

$$A_i = \sum_{j=1}^{r-i+1} \{Z_{ij} - \mu - \beta_j - \log(n_{ij})\}$$

$$= Z_{i.} - \ln_{i.} - (r - i + 1)\mu - \sum_{j=1}^{r-i+1} \beta_j,$$

with $Z_{i.} = \sum_{j=1}^{r-i+1} Z_{ij}$ and $\ln_{i.} = \sum_{j=1}^{r-i+1} \log(n_{ij})$.

3. $$f(\beta_j|.) = N\left(\frac{B_j}{(r - j + 1) + \sigma^2/\sigma^2_{\beta_j}}, \frac{\sigma^2}{(r - j + 1) + \sigma^2/\sigma^2_{\beta_j}}\right), j = 2, \ldots, r, \quad (10)$$

where

$$B_j = \sum_{i=1}^{r-j+1} \{Z_{ij} - \mu - \alpha_i - \log(n_{ij})\}$$

$$= Z_{.j} - ln_{.j} - (r - j + 1)\mu - \sum_{i=1}^{r-j+1} \alpha_i,$$

with $Z_{.j} = \sum_{i=1}^{r-j+1} Z_{ij}$ and $ln_{.j} = \sum_{i=1}^{r-j+1} \log(n_{ij})$.

4. $$f(\tau = \sigma^{-2}|.) = G(a_\tau + |U|/2, b_\tau + SS/2), \quad (11)$$

with $SS = \sum_{i=1}^{r} \sum_{j=1}^{r-i+1} (Z_{ij} - \mu_{ij})^2$ and $\mu_{ij} = \mu + \alpha_i + \beta_j + log(n_{ij})$. Recall that $\alpha_1 = \beta_1 = 0$.

5. $$f(Z_{ij}|.) = N(\mu_{ij}, \sigma^2), i = 2, \ldots, r, \quad j = r - i + 2, \ldots, r, \quad (12)$$

with $\mu_{ij} = \mu + \alpha_i + \beta_j + log(n_{ij})$.

6. $$f(\beta^*_j|.) \propto exp\left\{\beta^*_j n_{.j} - n_{..} \, log\left(\sum_{k=1}^{r} e^{\beta^*_k}\right) - 0.5\beta^{*2}_j/\sigma^2_{\beta^*_j}\right\}, j = 2, \ldots, r, \quad (13)$$

where $n_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij}$, $n_{.j} = \sum_{i=1}^{r} n_{ij}$ and $\beta^*_1 = 0$.

To obtain a sample from (13), we may use either Metropolis-Hastings algorithm or Gilks and Wild (1992) adaptive rejection sampling for log-concave distributions. Both methods provide similar convergence rates. We chose to use random walk Metropolis algorithm; see Smith and Roberts (1993). Each proposal density was taken to be normal with variance tuned to achieve an acceptance rate between 20 and 30%. The variance of the proposal density of each β^*_j, for $j = 2, \ldots, r$ was taken to be 0.023^2, 0.035^2, 0.05^2, 0.085^2, 0.065^2, and 0.14^2, respectively.

7. The last cell n_{ir} of each multinomial distribution is not essentially an extra random variable but a function of all previous claim counts because $n_{ir} = T_i - \sum_{j=1}^{r-1} n_{ij}$. Therefore, it is more convenient to rewrite (7) in the equivalent form

$$f(n_{i1}, \ldots, n_{i,r-1}|\pi_1, \ldots, \pi_r) = \frac{T_i!}{(T_i - = \sum_{j=1}^{r-1} n_{ij})! \prod_{j=1}^{r-1} n_{ij}!} \prod_{j=1}^{r-1} \pi_j^{n_{ij}} \pi_r^{(T_i - \sum_{j=1}^{r-1} n_{ij})}.$$

Now, n_{ir} is eliminated from the likelihood, and the updating of n_{ij} and n_{ir} is simultaneous. We generate each missing count n_{ij} for $j = r - i + 2, \ldots, r - 1, i = 3, \ldots, r$ from

$$f(n_{ij}|.) \propto f(N|\beta^*).$$

The constraint $T_i = \sum_{j=1}^{r} n_{ij}$ reduces the above posterior to

$$f(n_{ij}|.) \propto f(n_{i1}, \ldots, n_{i,r-1}|\beta^*)$$
$$\propto \frac{\pi_j^{n_{ij}}}{n_{ij}!} \frac{\pi_r^{\Omega_{ij}-n_{ij}}}{(\Omega_{ij} - n_{ij})!}$$

where $\Omega_{ij} = T_i - \sum_{k \neq j,r} n_{ik}$.

Therefore we sample from $f(n_{ij}|.)$ from

$$Binomial(p_j, \Omega_{ij}),\ p_j = \frac{\pi_j}{\pi_j + \pi_r} = (1 + exp(\beta_r^* - \beta_j^*))^{-1}.$$

Alternatively, if we assume that T_i are not available then we may use Poisson model for claim counts given by

$$n_{ij} \sim Poisson(\lambda_{ij})$$
$$log(\lambda_{ij}) = \mu^* + \alpha_i^* + \beta_j^*.$$

We impose constraints on new model parameters: $\alpha_1^* = 0$ and $\beta_1^* = 0$. We sample μ^*, α_i^*, β_j^* from the corresponding conditional posterior distributions and the missing claim amounts from the corresponding Poisson distribution instead of Binomial. The approach is similar if any other discrete distribution is adopted.

IMPLEMENTATION OF FOUR MODELS FOR OUTSTANDING LIABILITIES IN WINBUGS: A DISCUSSION OF A PAPER BY NTZOUFRAS AND DELLAPORTAS

DAVID P. M. SCOLLNIK*

ABSTRACT

Ntzoufras and Dellaportas (2002) described four models for outstanding claim amounts of the "reported but not settled" variety. Two of the models incorporated claim-counts data in addition to the claim amounts themselves in order to add a hierarchical stage in the usual log-normal and state-space models. The purpose of this discussion is to describe how the models presented in Ntzoufras and Dellaportas may be implemented using WinBUGS. The use of WinBUGS to implement the Bayesian analysis of a number of other actuarial models was considered by Scollnik (2001).

1. INTRODUCTION

Ntzoufras and Dellaportas (2002) described four models for outstanding claim amounts of the "reported but not settled" variety. Two of the models incorporated claim-counts data in addition to the claim amounts themselves. These authors suggested that the four models be implemented using Bayesian theory and the Markov chain Monte Carlo (MCMC) simulation method and provided some implementation details and illustrations with real insurance data. However, the implementation details and formulas provided are intimidating and cover only one of the four models, and the intricacies of implementing the Gibbs sampler (a version of MCMC) as described would be a formidable task for most actuarial practitioners.

Scollnik (2001) described how a variety of actuarial models can be implemented and analyzed in accordance with the Bayesian paradigm using MCMC via the BUGS (Bayesian inference Using Gibbs Sampling) suite of software packages. BUGS is a specialized software package for implementing MCMC-based analyses of full probability models in which all unknowns are treated as random variables. Its programming language is easily understood and allows the user to make a straightforward specification of the full probability model under consideration. Because the sole purpose of the BUGS package is to implement MCMC-based analyses of full probability models, its associated programming language was very efficiently designed for this task. Consequently, a MCMC analysis written in FORTRAN or another computer language and consisting of several hundreds of lines of code often can be reduced toliterally one or two dozen lines of code in

This work was supported by a grant from Natural Sciences and Engineering Council of Canada (NSERC).

* David P. M. Scollnik, A.S.A., Ph.D., is Associate Professor, Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4, e-mail, scollnik@ucalgary.ca.

BUGS. The Windows version of BUGS is known as WinBUGS, and it provides a graphical interface to the BUGS language. The BUGS Project Web site is found at www.mrc-bsu.cam.ac.uk/bugs.

The purpose of this discussion is to describe how the four models presented in Ntzoufras and Dellaportas (2002) can be implemented using WinBUGS. Illustrative WinBUGS code segments for two representative (i.e., the first and fourth) models are included below. Illustrative WinBUGS program files for all four models are available at www.math.ucalgary.ca/~scollnik/abcd/outstand. For an interpretation of each model, see Ntzoufras and Dellaportas. These interpretations are omitted in the interest of brevity.

2. THE FIRST MODEL

2.1 Description of the Model in WinBUGS

The first model is a log-normal (anova-type) model given by the formulation

$$Z_{ij} = \log Y_{ij} \sim \text{normal}(\mu_{ij}, \sigma^2), \tag{2.1}$$

$$\mu_{ij} = \mu + \alpha_i + \beta_j, \tag{2.2}$$

with $i, j = 1, \ldots, r$. The Y_{ij} values are the claim amounts paid by the insurance company with a delay of $j - 1$ years for accidents reported in the year i. The model is to be fitted to the data appearing in Table 1, for which $r = 7$. Note that the Y_{ij} values have already been deinflated. These are the same data utilized in Ntzoufras and Dellaportas (2002). Since the data are reported in thousands of drachmas, Equation (2.2) will be modified by subtracting log(1,000) from its right-hand side. The net effect on inferences will be to shift the location of the posterior distribution for μ by log(1,000), or about 6.9078, units to the right compared to that reported in Ntzoufras and Dellaportas. The model described so far is coded in WinBUGS as follows:

```
for ( i in 1 : r ) {
  for( j in 1 : r ) {
    y[ i, j ] ~ dlnorm( mu.y[ i, j ], tau )
    mu.y[ i, j ] <- mu + alpha[ i ]
      + beta[ j ] - log( 1000 )
  }
}
sigma2 <- 1 / tau
```

Note that the use of <− indicates a deterministic assignment. The symbol <− itself is made up of a "less than" symbol, <, followed immediately by a minus sign, −. Stochastic assignments are indicated with the tilde, ˜. Observe that `dlnorm` denotes a log-normal distribution and that its second parameter τ is the precision parameter (inverse variance) rather than variance of the underlying normal distribution. Hence, the definition of `sigma2` on the last line.

For identifiability of the α_i and β_j parameters, Ntzoufras and Dellaportas (2002) impose the corner constraints $\alpha_1 = \beta_1 = 0$. In WinBUGS this is accomplished easily with the lines

```
alpha[ 1 ] <- 0
 beta[ 1 ] <- 0
```

It is possible to impose sum-to-zero parametrizations in place of the corner constraints for identifiability: that is, we can require $\Sigma_i \alpha_i = \Sigma_j \beta_j = 0$.

Table 1
Claim Amounts (Thousands of Drachmas, Adjusted for Inflation)

Year of Origination	Years until Settlement ($j - 1$)						
	0	1	2	3	4	5	6
1989	527,003	183,260	90,539	50,471	37,875	10,110	21,165
1990	594,059	237,289	99,640	52,390	50,723	39,028	
1991	810,593	257,122	87,318	72,538	81,336		
1992	1,012,181	339,664	156,531	181,253			
1993	1,459,202	448,651	188,698				
1994	1,689,751	563,683					
1995	1,698,534						

Ntzoufras and Dellaportas originally made use of this approach in an earlier draft of their paper (see Ntzoufras and Dellaportas, 1998). In WinBUGS the sum-to-zero parametrizations are implemented with the lines

```
alpha[ 1 ] <- - sum( alpha[ 2 : r ] )
 beta[ 1 ] <- - sum( beta[ 2 : r ] )
```

Regardless of the identifiability constraints imposed, to complete the definition of the full probability model it remains to assign prior distributions to the unknown parameters. Ntzoufras and Dellaportas (2002) supposed that

$$\mu \sim \text{normal}(0, \sigma_\mu^2), \tag{2.3}$$

$$\alpha_i \sim \text{normal}(0, \sigma_{\alpha_i}^2),\ i = 2, \ldots, r, \tag{2.4}$$

$$\beta_j \sim \text{normal}(0, \sigma_{\beta_j}^2),\ j = 2, \ldots, r, \tag{2.5}$$

$$\tau = \sigma^{-2} \sim \text{gamma}(a_\tau, b_\tau). \tag{2.6}$$

Here `gamma( a, b )` denotes the gamma distribution with mean ab^{-1} and variance ab^{-2}. Vague diffuse proper priors are produced by using $\sigma_\mu^2 = 1{,}000$, $\sigma_{\alpha_i}^2 = 100$, $i = 2, \ldots, r$, $\sigma_{\beta_j}^2 = 100$, $j = 2, \ldots, r$, and $a_\tau = b_\tau = 0.001$. In WinBUGS this portion of the model is coded with the following lines:

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

mu ~ dnorm( 0.0, tau.mu )
tau ~ dgamma( 0.001, 0.001 )

tau.mu <- 1 / 1000
tau.alpha <- 1 / 100
tau.beta <- 1 / 100
```

Observe that `dnorm` denotes a normal distribution, and, as was the case with `dlnorm`, its second parameter is a precision or inverse variance—hence, the definitions of `tau.mu`, `tau.alpha`, and `tau.beta`.

Scollnik (2001) provides a complete discussion addressing the specifics of model checking, data loading, compiling, initializing, and running a simulation coded in WinBUGS. To implement the MCMC simulation of model 1, these instructions should be applied to the WinBUGS program for this model. This file is available from the web site referenced in Section 1. At this time it will only be noted that the data are defined in WinBUGS as follows:

```
list( r = 7 )

 y[,1]       y[,2]      y[,3]      y[,4]      y[,5]      y[,6]      y[,7]
 527003     183260      90539      50471      37875      10110      21165
 594059     237289      99640      52390      50723      39028         NA
 810593     257122      87318      72538      81336         NA         NA
1012181     339664     156531     181253         NA         NA         NA
1459202     448651     188698         NA         NA         NA         NA
1689751     563683         NA         NA         NA         NA         NA
1698534         NA         NA         NA         NA         NA         NA
```

Observe that a missing value is demarked with an `NA`. Of course, these are the values for which posterior inferences are desired.

2.2 An Alternate Parametrization

In the "Seeds" example included in the *BUGS Examples Volume 1* document, available from the documentation page at the BUGS Project Web site referenced previously, an alternate method is presented for implementing the sum-to-zero parametrization for the α_i and β_j parameters; this alternative is presented below.

Suppose we wish the α_i parameters to have marginal prior distributions

$$\alpha_i \sim \text{normal}(0, \sigma_\alpha^2),\ i = 1, \ldots, r,$$

and the α_i's must sum to zero. Introduce independent variables α_i^x with marginal distributions

$$\alpha_i^x \sim \text{normal}(0, \sigma_\alpha^2), i = 1, \ldots, r,$$

and set

$$\alpha_i = \sqrt{\frac{r}{r-1}}\left(\alpha_i^x - \frac{1}{r}\sum_{j=1}^{r}\alpha_j^x\right), i = 1, \ldots, r.$$

Then simple algebra and elementary normal distribution theory lead to the result that

$$\alpha_i = \sqrt{\frac{r}{r-1}}\left(\frac{r-1}{r}\alpha_i^x - \frac{1}{r}\sum_{j \neq i}\alpha_j^x\right) \sim \text{normal}(0, \sigma_{\alpha_i}^2), \quad (2.7)$$

for $i = 1, \ldots, r$, as required. It is evident that the α_i parameters sum to zero, by construction. The same construction for the β_j parameters can be repeated. In WinBUGS this method is coded as follows:

```
for( i in 1 : r ) {
  alpha.x[i] ~ dnorm( 0.0, tau.alpha )
  alpha[i] <- sqrt( r / (r - 1) )
      * ( alpha.x[i] - mean( alpha.x[] ) )
  beta.x[i] ~ dnorm( 0.0, tau.beta )
  beta[i] <- sqrt( r / (r - 1) )
      * ( beta.x[i] - mean( beta.x[] ) )
}
```

In practice the identifiability constraints and parametrizations used for the α_i and β_j parameters generally will not have too great an impact on the final inferences developed on the basis of the simulation results for the missing claim amounts.

3. THE FOURTH MODEL

3.1 Description of the Model in WinBUGS

The fourth model incorporates information on the claim numbers settled to date and the number of claims reported but not yet settled through the formulation

$$Z_{ij} = \log Y_{ij} \sim \text{normal}(\mu_{ij}, \sigma^2), \quad (3.1)$$

$$\mu_{ij} = \mu + \alpha_i + \beta_{ij} + \log(n_{ij}), \quad (3.2)$$

for $i, j = 1, \ldots, r$. Here the n_{ij} values are the claim numbers settled by the insurance company with a delay of $j - 1$ years for accidents reported in the year i (see Table 2). It is apparent from Equation (3.2) that the "beta" parameters now are assumed to vary with both i and j. Further, the fourth model assumes a state-space formulation for both the α_i and β_{ij} parameters so that

$$\beta_{ij} = \beta_{i-1,j} + v_i, v_i \sim \text{normal}(0, \sigma_v^2), \quad i, j = 2, \ldots, r, \quad (3.3)$$

$$\alpha_i = \alpha_{i-1} + h_i, h_i \sim \text{normal}(0, \sigma_h^2), \quad i = 2, \ldots, r, \quad (3.4)$$

with corner constraints $\alpha_1 = 0$ and $\beta_{i1} = 0$, for $i = 1, \ldots, r$. Modeling these recursions in WinBUGS is very easy to do, as shown here:

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

Modeling the Y_{ij} values in WinBUGS is a little trickier than it was for model 1, as this task is now complicated by the necessity of simultaneously modeling the claim numbers n_{ij} as well. Ntzoufras

Table 2

Claim Counts

Year of Origination	Years until Settlement ($j - 1$)							Total
	0	1	2	3	4	5	6	
1989	6,622	1,943	489	138	61	223	66	9,542
1990	6,943	2,133	632	154	162	390		10,496
1991	8,610	2,216	736	651	256			12,601
1992	9,791	3,167	1,570	624				15,565
1993	11,722	3,192	1,773					17,735
1994	13,684	3,664						19,746
1995	13,068							18,600

and Dellaportas (2002) assume a multinomial distribution for the claim numbers associated with each accident year; that is,

$$(n_{i1}, n_{i2}, \ldots, n_{ir}) \sim \text{multinomial}(\pi_1, \pi_2, \ldots, \pi_r; T_i), \quad i = 1, \ldots, r, \quad (3.5)$$

where π_j is the probability for a claim to be settled with a delay of $j - 1$ years and T_i is the total number of claims reported in year i, so that $T_i = n_{i1} + n_{i2} + \cdots + n_{ir}$. The values of T_i are assumed to be known, as in Table 2, which reports these values in its last column. Although WinBUGS does permit the definition of a multinomial random variable with the use of `dmulti`, the multinomial data so defined must either be completely observed or completely unobserved.

Of course, the actual claim numbers are only partially observed for most years of claim origination, as Table 2 illustrates. This is not a great problem, as we can partition the vector of claim numbers associated with a particular year of origination into a marginal multinomial variable for the vector of claim numbers observed in the years to date along with the total number of claims remaining to be settled, and a conditional multinomial variable for the vector of claim numbers settled in future years given the number already settled. Perhaps this is best explained in the context of a concrete example like Table 2. The process is also simplified if we prepare the data in Table 2 for entry into WinBUGS as illustrated here:

```
n[,1]   n[,2]   n[,3]   n[,4]   n[,5]   n[,6]   n[,7]   n[,8]
 6622    1943     489     138      61     223      66       0
 6943    2133     632     154     162     390      82       0
 8610    2216     736     651     256     132      NA      NA
 9791    3167    1570     624     413      NA      NA      NA
11722    3192    1773    1048      NA      NA      NA      NA
13684    3664    2398      NA      NA      NA      NA      NA
13068    5532      NA      NA      NA      NA      NA      NA
```

Observe that a diagonal has been added to the data, running from the top right downwards to the left. The entries in the diagonal are the number of reported claims associated with particular years of origination that remain to be settled. Consider the fourth row of data in this table. The vector of counts (9791, 3167, 1570, 624, 413) is a realization of a

$$\text{multinomial}(\pi_1, \pi_2, \pi_3, \pi_4, 1 - \Sigma_{i=1}^{4} \pi_i; T_4) \quad (3.6)$$

random variable. Observe that the value of 413 was calculated as $T_4 - \Sigma_{i=1}^{4} n_{4i}$. The three NA values in the fourth row correspond to the unobserved number of reported claims to be settled in each of the future years. Conditioning on the known number of claims settled in previous years, these unknown counts have a conditional distribution that is

$$\text{multinomial}\left(\frac{\pi_5}{\pi_5 + \pi_6 + \pi_7}, \frac{\pi_6}{\pi_5 + \pi_6 + \pi_7}, \frac{\pi_7}{\pi_5 + \pi_6 + \pi_7}; 413\right). \quad (3.7)$$

Each row containing NA values can be partitioned into marginal and conditional multinomial variables in an analogous way; this is shown in the WinBUGS code in the following example. Note that caution must be exercised in order to ensure that the indexings in the `for` loops properly accommodate the diagonal added to the claim counts.

```
# Years 1 and 2 (years with no unknown cell counts).

for( i in 1 : 2 ) {
  for( j in 1 : r ) {
    y[ i, j ] ~ dinorm( mu.y[ i, j ], tau )
    mu.y[ i, j ] <- mu + alpha[ i ] + beta[ i, j ] + log( n[ i, j ] ) - log( 1000 )
  }
}

# Years 3 through r (years with some unknown cell counts).

for( i in 3 : r ) {
  for( j in 1 : (r+1-i) ) {
    y[ i, j ] ~ dlnorm( mu.y[ i, j ], tau )
    mu.y[ i, j ] <- mu + alpha[ i ] + beta[ i, j ] + log( n[ i, j ] ) - log( 1000 )
  }
  #
  # The index of "n" is shifted here, as one diagonal must be skipped.
  #
  for( j in (r+2-i) : r ) {
    y[ i, j ] ~ dlnorm( mu.y[ i, j ], tau )
    mu.y[ i, j ] <- mu + alpha[ i ] + beta[ i, j ] + log( n[ i, j + 1 ] ) - log( 1000 )
  }
}

# The first two rows of "n" are known. The first row was completely
# observed, and the single missing cell in the second row is deduced
# from the given row total.

n[ 1, 1:r ] ~ dmulti( pi[], T[1] )
n[ 2, 1:r ] ~ dmulti( pi[], T[2] )

# The third through seventh rows must each be split into the cells
# with known values and the cells with unobserved values.

# Use a marginal multinomial node for the cells with known values
# in a given row.

for ( i in 3 : 4 ) {
  for( j in 1 : (r+1-i) ) {
    pi.marg[ i, j ] <- pi[ j ]
  }
  pi.marg[ i, r+2-i ] <- 1 - sum( pi.marg[ i, 1:(r+1-i) ] )
  n[ i, 1:(r+2-i) ] ~ dmulti( pi.marg[ i, 1:(r+2-i) ], T[i] )
}

# Use a conditional multinomial node for the cells with unknown values
# in a given row, given the total of these counts (which IS known).

for( i in 3 : r ) {
  for( j in (r+3-i) : (r+1) ) {
    pi.cond[ i, j ] <- pi[ j - 1 ] / pi.marg[ i, r+2-i ]
  }
  n[ i, (r+3-i):(r+1) ] ~
    dmulti( pi.cond[ i, (r+3-i):(r+1) ], n[ i, r+2-i ] )
}
```

Next, a prior distribution on the π_j values will be defined. Ntzoufras and Dellaportas (2002) assume that

$$\log\left(\frac{\pi_j}{\pi_1}\right) = p_j \sim \text{normal}(0, \sigma_p^2), j = 2, \ldots, r. \quad (3.8)$$

The value of π_1 is fixed by the requirement that the π_j values sum to unity. Since $\pi_j = \pi_1 e^{p_j}$, for $j = 2, \ldots, r$, we have

$$1 = \pi_1 + \pi_2 + \ldots + \pi_r = \pi_1(1 + e^{p_1} + \ldots + e^{p_r}). \quad (3.9)$$

Thus,

$$\pi_1 = \frac{1}{(1 + e^{p_1} + \ldots + e^{p_r})} \quad (3.10)$$

and

$$\pi_j = \frac{e^{p_j}}{(1 + e^{p_1} + \ldots + e^{p_r})}, \quad (3.11)$$

for $j = 2, \ldots, r$. This is coded in WinBUGS as follows:

```
for( i in 1 : (r − 1)) {
  # If you like, use
  # p.exp[ i ] ~ dlnorm( 0.0, tau.p )
  # in place of the first two lines below.
  p[ i ] ~ dnorm( 0.0, tau.p )
  p.exp[ i ] <− exp( p[ i ] )
  pi[ i+1 ] <− p.exp[ i ] /
      ( 1 + sum( p.exp[ 1:(r−1) ] ) )
}
pi[ 1 ] <− 1 − sum( pi[ 2:r ] )
```

To complete the definition of the full probability model, it remains to assign prior distributions to the remaining unknown parameters. Ntzoufras and Dellaportas (2002) supposed that

$$\mu \sim \text{normal}(0, \sigma_\mu^2), \quad (3.12)$$

$$\beta_{1j} \sim \text{normal}(0, \sigma_{\beta_{1j}}^2), j = 2, \ldots, r, \quad (3.13)$$

$$\tau = \sigma^{-2} \sim \text{gamma}(a_\tau, b_\tau), \quad (3.14)$$

$$\tau_v = \sigma_v^{-2} \sim \text{gamma}(a_v, b_v), \quad (3.15)$$

$$\tau_h = \sigma_h^{-2} \sim \text{gamma}(a_h, b_h). \quad (3.16)$$

Ntzoufras and Dellaportas produced vague diffuse proper priors by using $\sigma_\mu^2 = 1{,}000$, $\sigma_{\beta_{1j}}^2 = 100$, $j = 2, \ldots, r$, $a_\tau = b_\tau = 0.001$, and $a_v = b_v = a_h = b_h = 10^{-10}$. For the additional parameters p_j, they used $\sigma_p^2 = 100$, for $j = 2, \ldots, r$. The simulation here for this model will proceed similarly, except that we will set $a_v = b_v = a_h = b_h = 10^{-5}$ in an attempt to speed up convergence of the simulation. The same change was made to the third model. The corresponding WinBUGS code is omitted, as it is very similar to what has gone before.

3.2 A Comment Regarding the Fourth Model

There is one complication with the model description given above. Equation (3.2) implicity assumes that $n_{ij} > 0$ for all values of i and j. This restriction is not evident in Equation (3.5). However, provided that the number of reported but not settled claims is large, and the number of years being forecast is moderate, this should not cause a problem, as the simulation will tend not to generate years with zero settlements. In effect, the simulation of the unknown claim numbers is made from a truncated multinomial distribution for all intents and purposes.

4. ANALYSIS OF THE FOUR MODELS

Illustrative WinBUGS files for all four of the models discussed in Ntzoufras and Dellaportas (2002) are available from the Web site referenced in Section 1. Each file contains the following definitions, which will make it easier to monitor the simulated values of the predicted total claim amounts to be generated in future years:

```
for( i in 2 : r ) {
  outstand.row[ i ] <− sum( y[ i, ( r + 2 − i ) : r ] ) / 1000
}
outstand.cal[ 2 ] <−
  ( y[2,r] + y[3,r−1] + y[4,r−2] + y[5,r−3] + y[6,r−4] + y[7, r−5] ) / 1000
outstand.cal[ 3 ] <−
  ( y[3,r] + y[4,r−1] + y[5,r−2] + y[6,r−3] + y[7,r−4] ) / 1000
```

```
outstand.cal[ 4 ] <- ( y[4,r] + y[5,r-1] + y[6,r-2] + y[6,r-3] ) / 1000
outstand.cal[ 5 ] <- ( y[5,r] + y[6,r-1] + y[7,r-2] ) / 1000
outstand.cal[ 6 ] <- ( y[6,r] + y[7,r-1] ) / 1000
outstand.cal[ 7 ] <- ( y[7,r] ) / 1000
outstand.cal[ 8 ] <- sum( outstand.cal[ 2 : 7 ] )
```

The `outstand.row` values correspond to the total predicted claim amounts associated with the different years of origination (the rows in Table 1), whereas the `outstand.cal` values correspond to the total predicted claim amounts paid in future calendar years (along the unobserved diagonals in Table 1). The sums have been divided by 1,000 simply to keep the size of the numbers down.

Scollnik (2001) provides a complete discussion addressing the specifics of model checking and data loading, compiling and initializing, running, monitoring for convergence, and summarizing the results from a simulation coded in WinBUGS. These steps have been implemented for each of the four WinBUGS model files in turn. For each model four independent simulations were run and each simulation was allowed to burn in for 20,000 iterations. Each simulation was then run for an additional 100,000 iterations, over which the values of the unknown variables of interest (the missing data and unknown model parameters) were monitored and then summarized. Thus, the summaries are based on $4 \times 100{,}000 = 400{,}000$ simulated values for each unknown variable of interest, for each model. Convergence was monitored using the Gelman-Rubin diagnostic statistic as it is implemented in WinBUGS. A portion of the simulation-based inferences that resulted are given in Tables 3 and 4. Summaries for some additional parameters have been saved in the four WinBUGS files, and additional summaries, particularly graphical summaries such as histograms and estimated posterior density plots for each monitored parameter, can also be developed. See Scollnik (2001) for details.

These results can be compared with those reported in Tables 7 and 8 in Ntzoufras and Dellaportas (2002). The two sets of inferences are in general agreement, although there are some discrepancies in the results reported for models 3 and 4. Even in this case, though, the 95% posterior probability intervals reported in Tables 3 and 4 include the values for the posterior means reported in Tables 7 and 8 in Ntzoufras and Dellaportas. It is difficult to discern precisely why the inferences for these latter two models are not in closer agreement. The discrepancies may be attributed in part to simulation error, rounding error, or a model coding error either in the WinBUGS file or on the part of Ntzoufras and Del-

Table 3

Posterior Mean, Standard Deviation, and 95% Probability Intervals for Total Outstanding Claim Amounts of Each Accident Year (Millions of Drachmas, Adjusted for Inflation)

Year	Model			
	1	2	3	4
1990	33.48 (15.97)	31.11 (19.36)	31.18 (13.68)	29.01 (15.23)
	(12.81, 72.3)	(9.08, 78.73)	(12.61, 64.39)	(10.01, 66.68)
1991	66.19 (22.79)	12.9 (5.944)	65.64 (20.72)	15.54 (6.054)
	(33.6, 120.4)	(5.35, 27.31)	(34.9, 114.9)	(7.165, 30.16)
1992	213.2 (62.08)	95.12 (38.82)	197.1 (52.38)	92.76 (29.55)
	(119.6, 359.5)	(43.88, 187.5)	(114.4, 318.9)	(49.56, 163.2)
1993	403.4 (111.8)	300.5 (112.1)	384.6 (94.44)	266.1 (74.17)
	(233.2, 664.9)	(145.0, 570.0)	(233.5, 601.8)	(153.7, 441.4)
1994	759.5 (220.5)	634.7 (245.9)	694.1 (176.5)	487.0 (134.2)
	(426.4, 1,280.0)	(297.56, 1,227.0)	(412.8, 1,101.0)	(289.5, 809.7)
1995	1,422.0 (535.6)	1,234.0 (611.5)	1,335.0 (457.9)	923.2 (324.6)
	(671.7, 2,711.0)	(469.9, 2,729.0)	(700.1, 2,354.0)	(518.8, 1,640.0)

Table 4
Posterior Mean, Standard Deviation, and 95% Probability Intervals for Total Claim Amounts to be Paid in Each Future Year (Millions of Drachmas, Adjusted for Inflation)

Year	Model 1	Model 2	Model 3	Model 4
1996	1,219.0 (330.6) (753.3, 2,009.0)	1,075.0 (390.6) (573.9, 2,020.0)	1,152.0 (294.1) (722.5, 1,825.0)	862.4 (240.3) (525.8, 1,419.0)
1997	671.4 (168.1) (421.5, 1,068.0)	577.1 (212.3) (296.9, 1,082.0)	630.2 (152.2) (395.9, 977.6)	450.5 (130.8) (265.4, 755.1)
1998	461.2 (127.2) (271.0, 762.1)	361.3 (147.5) (167.8, 719.5)	423.3 (106.5) (254.8, 670.3)	281.8 (85.89) (157.5, 487.7)
1999	298.0 (97.37) (160.3, 530.6)	185.1 (92.44) (76.73, 408.0)	274.6 (86.99) (148.4, 475.3)	141.3 (54.26) (70.6, 268.3)
2000	152.4 (58.66) (71.65, 292.6)	64.95 (33.78) (24.26, 146.9)	137.7 (50.64) (65.31, 257.5)	50.6 (21.08) (22.49, 101.4)
2001	88.33 (52.38) (27.38, 216.2)	36.33 (30.26) (7.68, 108.8)	79.4 (43.58) (26.5, 184.4)	27.26 (17.12) (7.969, 68.87)
Total	2,891.0 (613.8) (1,954.0, 4,325.0)	2,299.0 (673.6) (1,360.0, 3,897.0)	2,697.0 (563.3) (1,820.0, 3,949.0)	1,814.0 (407.8) (1,227.0, 2,756.0)

laportas. It should also be borne in mind that a slightly different prior density specification was used for models 3 and 4 than did Ntzoufras and Dellaportas. Also, the fact that the lognormal model is quite long tailed may explain the differences in the simulation-based estimates of the posterior means. Finally, note that the results for models 3 and 4 reported in Ntzoufras and Dellaportas (2002) also differ from those reported for the same data set and models in Ntzoufras and Dellaportas (1998).

REFERENCES

NTZOUFRAS, I., AND P. DELLAPORTAS. 1998. *Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty.* Technical Report 41, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

NTZOUFRAS, I., AND P. DELLAPORTAS. 2002. "Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty," *North American Actuarial Journal* 6(1): 113–128.

SCOLLNIK, D.P.M. 2001. "Actuarial Modeling with MCMC and BUGS." *North American Actuarial Journal* 5(2): 96–124. See www.math.ucalgary.ca/~scollnik/abcd/ for the related WinBUGS program files.

Additional discussions on this paper can be submitted until July 2002. 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.