# APPLICATION OF BAYESIAN MODELS WITH MARKOV CHAIN MONTE CARLO SIMULATION TO REAL CLAIMS DATA

Jackie Li BCom, MApplFin, FIAA

Centre for Actuarial Studies, The University of Melbourne, Australia

Email: k.li3@pgrad.unimelb.edu.au

## ABSTRACT

In this paper we demonstrate an application of Bayesian models with Markov chain Monte Carlo (MCMC) simulation to some real claims data (gross of reinsurance) of three Australian private lines of business via the use of BUGS (Bayesian Inference Using Gibbs Sampling). We take the accident period effect and the development period effect into account and consider parameter error and process error. We devise an approximation for the over-dispersed Poisson (ODP) assumption in BUGS and select suitable non-informative priors. We also make use of some formal model criticism procedures to validate the selected model structures. We then compare our estimates with the actual claim payments made after the valuation date. We find that BUGS generally provides parameter estimates consistent with those produced by generalised linear models (GLMs). Moreover, our estimates of the expected total outstanding claims liability plus the aggregate risk margin cover the actual claim payments appropriately.

# 1. INTRODUCTION

This paper demonstrates the application of Bayesian estimation with MCMC simulation to some real claims data for three Australian private lines of business. These lines are motor insurance, public liability insurance, and compulsory third party (CTP) insurance. The software BUGS is employed as a platform for the claims modelling and the reserving calculations. Scollnik (2001) provides a clear presentation on adopting BUGS in actuarial modelling. The accident period effect and the development period effect are taken into account in the modelling process, and parameter error and process error are included in the reserving calculations. Formal model criticism procedures are adopted to justify the selected model structures. In addition, GLMs are used to check the reasonableness of the parameter estimates generated by BUGS.

The source of the claims data and the corresponding limitations are stated in Section 2. The general procedures of claims modelling are set forth in Section 3. The application results of the three lines of business are analysed in Sections 4 to 6. Comparison of our results with those reported by the Australian general insurance industry is set forth in Section 7. Final discussion is set forth in Section 8. Some BUGS codings are also provided in the appendices.

As noted in Australian Guidance Note GN 353, Bayesian modelling is one of those techniques most likely to be useful in the quantification of risk margin and MCMC simulation is at present largely experimental. (Under Australian Prudential Standards

GPS 210 and 310, the outstanding claims liabilities must be valued at the 75th percentile, subject to a minimum of the mean plus one half of the standard deviation. The difference between the mean and the 75th percentile is called the risk margin, which is often expressed as a percentage of the mean.) Despite some papers such as Verrall (1990), Scollnik (2001), Ntzoufras and Dellaportas (2002), and de Alba (2002a, 2002b, 2003), the practical application of Bayesian models with MCMC simulation to general insurance liabilities is still in its early stage. Accordingly, in this paper, we explore the feasibility of modelling some real claims data with Bayesian models, making use of BUGS to implement MCMC simulation and exercising reasonable judgement in selecting the parameters and distributions. Various quantities including the mean, the 75th percentile, and the standard deviation of the outstanding claims liability are sampled from the MCMC simulation process for each line of business. (The entire empirical distribution of the outstanding claims liability can be formed via BUGS.)

## 2. CLAIMS DATA AND LIMITATIONS

We make use of the aggregated claims data published by Australian Prudential Regulation Authority (APRA) and its predecessor the Insurance and Superannuation Commission (ISC). Australian general insurers are statutorily required to provide various financial returns to APRA annually (and to the ISC previously) under the Insurance Act 1973. These returns have to be submitted according to the nature of the business. APRA (and the ISC previously) discloses some of these figures in aggregate terms in its publications titled ‘Selected Statistics on the General Insurance Industry’ (available on the Internet since 1996). No individual company’s information is exposed to the public due to privacy reasons.

We have collected the aggregated ‘Gross Claims Paid’ of the private sector from the ‘Claims Analysis’ section of the ‘Selected Statistics on the General Insurance Industry’. We have selected payment years 1983 to 1996, after which the figures are not subdivided by the accident year and so cannot be used to form run-off triangles. In addition, we have combined the ‘Domestic Motor’ and ‘Commercial Motor’ lines of business starting from payment year 1992 for motor insurance.

Adopting the approach as described in Hart et al (1996), all the gross of reinsurance claim payment figures have been converted to 31 December 1996 dollar values in accordance with the average weekly ordinary time earnings (AWOTE) before the modelling and calculations. This procedure is common in practice and is based on the assumption that wage inflation is the 'normal' inflation for the claims considered. The difference between this wage inflation and the actual claims inflation occurred is called 'superimposed' inflation, of which the significance may be revealed via examining the residuals. Moreover, the cells from development year 11 onwards of each run-off triangle are adjusted because each of the original figures refers to claims of several accident years. Further inflation and adjustment details are stated in Appendix I.

The following are some limitations of these claims and inflation data, in which the list is not meant to be exhaustive:

(1) Continual changes in reporting requirements lead to certain data inconsistencies.

(2) Insurers submit revised financial returns from time to time.

(3) Changes in accounting periods have led to incomplete financial returns.

(4) Data inconsistencies arise from system resolution issues that may have previously excluded some insurers from the publication or included an insurer twice due to changes in balance date.

(5) The definitions of some lines of business are modified over time.

(6) There have been some shifts of business from the private sector to the public sector and vice versa.

(7) The adjustments to the cells in the upper right corner of each run-off triangle are merely approximations.

(8) Actual claims development beyond the latest development year of the adjusted claims data is possible and the assessment of its uncertainty is highly judgemental.

(9) Despite its popularity, AWOTE may not be the most suitable inflation index for adjusting the claim amounts of the three lines of business considered. There are other wage inflation indices and other types of economic indices such as the consumer price index (CPI).

Under these data limitations, one needs to take extreme care in drawing any inference from the results of this paper. As the claims data are of the whole general insurance private sector in Australia, individual companies may experience claims run-off patterns different from those of the industry in aggregate. Moreover, as stated in the BUGS manual, 'MCMC sampling can be dangerous' and hence the application results need to be interpreted carefully, particularly for certain 'types of model that are currently not featured'. Overall, for the models tested in this paper, BUGS runs smoothly and provides sensible results in various aspects. We believe the application of BUGS here is justifiable.

The following three claims development triangles are formed for motor insurance, public liability insurance, and CTP insurance respectively. For each line of business, the upper left triangle of cells represents the past gross of reinsurance claim payments. The missing lower right triangle represents the gross of reinsurance outstanding claims liability to be assessed. For motor insurance, it is assumed that all claims will be settled in 13 years. For the other two lines, it is assumed that all claims will be settled in 14 years.

**Motor Insurance – Gross Claim Payments** (31 Dec 1996 dollar values of 1 million)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1984** | 1,102.99 | 271.65 | 8.15 | 1.09 | 0.55 | 0.15 | 0.07 | 0.39 | 0.26 | 0.01 | 0.02 | 0.01 | 0.19 |
| **1985** | 1,294.15 | 364.75 | 10.14 | 1.32 | 0.08 | -0.19 | 0.22 | 0.31 | 0.24 | 0.18 | 0.02 | 0.34 | |
| **1986** | 1,374.86 | 421.30 | 13.34 | 2.90 | 0.76 | 0.41 | 0.86 | 0.26 | 0.52 | 0.52 | 0.55 | | |
| **1987** | 1,663.80 | 439.24 | 20.76 | 4.10 | 2.81 | 1.71 | 1.70 | 0.45 | 0.26 | 0.18 | | | |
| **1988** | 1,671.97 | 419.87 | 13.41 | 2.69 | 2.44 | 1.85 | 1.33 | 0.09 | 0.14 | | | | |
| **1989** | 1,703.57 | 490.72 | 17.67 | 8.49 | 3.96 | 1.91 | 0.78 | 0.19 | | | | | |
| **1990** | 2,056.00 | 589.97 | 30.90 | 7.82 | 5.08 | 2.59 | 1.53 | | | | | | |
| **1991** | 2,009.07 | 423.54 | 16.36 | 11.77 | 3.49 | 4.22 | | | | | | | |
| **1992** | 1,825.12 | 426.93 | 124.93 | 11.39 | 3.78 | | | | | | | | |
| **1993** | 1,974.08 | 456.91 | 138.09 | 10.21 | | | | | | | | | |
| **1994** | 2,370.16 | 516.53 | 154.85 | | | | | | | | | | |
| **1995** | 2,515.28 | 632.95 | | | | | | | | | | | |
| **1996** | 3,065.13 | | | | | | | | | | | | |

**Public Liability Insurance – Gross Claim Payments** (31 Dec 1996 dollar values of 1 million)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1983** | 17.27 | 25.16 | 20.33 | 31.33 | 20.22 | 115.34 | 23.11 | 18.94 | 13.58 | 7.43 | 7.96 | 12.10 | 25.09 | 3.41 |
| **1984** | 18.14 | 26.14 | 23.47 | 29.37 | 25.05 | 18.66 | 21.84 | 15.70 | 12.42 | 6.38 | 17.29 | 36.97 | 4.98 | |
| **1985** | 27.12 | 32.53 | 28.67 | 31.35 | 27.54 | 29.01 | 22.89 | 15.69 | 12.09 | 11.85 | 52.81 | 7.34 | | |
| **1986** | 23.35 | 35.39 | 40.67 | 37.37 | 35.23 | 35.45 | 21.96 | 13.42 | 15.47 | 9.59 | 10.48 | | | |
| **1987** | 23.84 | 45.22 | 38.57 | 34.04 | 38.17 | 23.08 | 18.89 | 20.10 | 11.61 | 12.73 | | | | |
| **1988** | 33.35 | 48.88 | 41.93 | 34.92 | 22.27 | 25.28 | 29.15 | 14.74 | 12.85 | | | | | |
| **1989** | 29.34 | 65.25 | 42.75 | 25.17 | 33.36 | 32.06 | 20.65 | 17.36 | | | | | | |
| **1990** | 42.81 | 87.77 | 23.30 | 32.19 | 38.13 | 40.22 | 23.81 | | | | | | | |
| **1991** | 40.58 | 37.38 | 33.20 | 45.40 | 41.19 | 28.18 | | | | | | | | |
| **1992** | 12.11 | 44.77 | 53.39 | 59.36 | 43.68 | | | | | | | | | |
| **1993** | 31.41 | 59.03 | 47.05 | 53.34 | | | | | | | | | | |
| **1994** | 35.30 | 57.59 | 71.54 | | | | | | | | | | | |
| **1995** | 31.23 | 56.78 | | | | | | | | | | | | |
| **1996** | 36.69 | | | | | | | | | | | | | |

**CTP Insurance – Gross Claim Payments** (31 Dec 1996 dollar values of 1 million)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1983** | 2.88 | 10.20 | 15.53 | 20.52 | 17.32 | 15.23 | 7.25 | 7.64 | 5.78 | 5.48 | 7.25 | 3.88 | 1.34 | 2.71 |
| **1984** | 3.59 | 11.09 | 22.04 | 20.34 | 20.56 | 16.23 | 17.56 | 10.64 | 6.96 | 4.10 | 6.38 | 2.34 | 4.33 | |
| **1985** | 2.32 | 10.70 | 15.60 | 18.97 | 14.50 | 20.50 | 11.82 | 15.27 | 7.26 | 5.80 | 3.84 | 7.59 | | |
| **1986** | 2.26 | 9.46 | 14.95 | 16.25 | 19.73 | 15.97 | 19.10 | 9.28 | 8.25 | 5.63 | 12.46 | | | |
| **1987** | 2.62 | 10.79 | 13.41 | 20.05 | 17.89 | 21.47 | 16.00 | 11.83 | 10.18 | 7.76 | | | | |
| **1988** | 2.40 | 8.62 | 10.40 | 15.23 | 19.82 | 20.44 | 12.54 | 13.65 | 14.24 | | | | | |
| **1989** | 6.59 | 16.35 | 27.87 | 44.59 | 58.06 | 44.05 | 31.71 | 32.91 | | | | | | |
| **1990** | 33.94 | 31.20 | 65.27 | 118.45 | 184.31 | 115.37 | 86.18 | | | | | | | |
| **1991** | 35.25 | 38.35 | 103.98 | 194.44 | 141.36 | 116.76 | | | | | | | | |
| **1992** | 30.48 | 54.66 | 156.35 | 184.44 | 145.33 | | | | | | | | | |
| **1993** | 31.07 | 87.71 | 191.47 | 202.85 | | | | | | | | | | |
| **1994** | 46.76 | 104.89 | 238.32 | | | | | | | | | | | |
| **1995** | 46.47 | 117.93 | | | | | | | | | | | | |
| **1996** | 43.03 | | | | | | | | | | | | | |

The run-off patterns seem reasonable at a glance. For motor insurance, the claim payments develop quickly and most of them are settled in the first few development years. For public liability insurance, it takes much longer for the claim payments to develop. For CTP insurance, many of the claim payments are made only a few development years after occurrence. These features accord with the general perceptions of the claims development behaviour of the three lines of business.

In practice, it is more common to have the claims projection based on quarterly or monthly claims data for just a few development years for motor insurance.

For convenience, accident years 1984 to 1996 are sometimes referred to as accident years 1 to 13 for motor insurance, and accident years 1983 to 1996 are sometimes referred to as accident years 1 to 14 for the other two lines.

# 3. PROCEDURES OF CLAIMS MODELLING

This section provides a description of the procedures we adopt for modelling the past claims data with Bayesian techniques and projecting the future claims with MCMC simulation. Let $X_{i,j}$ be the incremental claim amount of accident period $i$ and development period $j$. Let $\alpha_i$ and $\beta_j$ be the parameters allowing for the accident period effect and the development period effect respectively, $\mu_{i,j} = \mathrm{E}\left(X_{i,j}\right)$, and $\eta_{i,j}$ be the linear predictor. Let $\hat{X}_{i,j}$ and $\hat{\eta}_{i,j}$ be the estimators of $\mu_{i,j}$ and $\eta_{i,j}$.

For each line of business, we start with two model structures that are similar to the over-dispersed Poisson (ODP) model structure and the gamma model structure as stated in Verrall (1999). We assume the unknown parameters are random variables following some vague non-informative prior distributions and the incremental claim payments $X_{i,j}$'s follow some pre-determined process distributions. The fixed parameters of the prior distributions are chosen to ensure BUGS runs smoothly. The mathematical expressions are shown as follows:

$X_{i,j} \sim \mathrm{Pn}(\mu_{i,j})$ for ODP or $X_{i,j} \sim \gamma(\mu_{i,j}, r)$ for gamma, $\eta_{i,j} = \ln(\mu_{i,j}) = v + \alpha_i + \beta_j$,

$v \sim \mathrm{N}(0, 1000)$, $\alpha_i \sim \mathrm{N}(0, 100)$, $\beta_j \sim \mathrm{N}(0, 100)$, $\alpha_1 = \beta_1 = 0$,

and $r \sim \mathrm{U}(0, 100)$ for gamma.

As reflected in the analyses later, the variance of the claim payments $X_{i,j}$'s is generally larger than the mean. The Poisson assumption alone is thus inappropriate (in which the mean and variance are equal). GLMs tackle this problem by setting the dispersion parameter $\phi$ to be greater than one, in which case the variance of $X_{i,j}$ is $\phi\,\mu_{i,j}$. To accommodate this over-dispersion with the Poisson assumption similarly in BUGS, we make the approximation that $X_{i,j} \sim \mathrm{N}\left(\mu_{i,j}\,,\phi\,\mu_{i,j}\right)$ and $\phi^{-1} \sim \gamma\left(1\,,10^{-2}\right)$. This approximation is based on the fact that a Poisson variable $X \sim \mathrm{Pn}\left(\mu\right)$ is approximately normally distributed as $X \sim \mathrm{N}\left(\mu\,,\mu\right)$ when $\mu$ tends to infinity (the gross of reinsurance claim payments considered here are substantial in this regard) and that ODP in GLMs is allowed for by making the dispersion parameter $\phi$ greater than one. Furthermore, we use the corresponding GLM parameter estimates to justify those generated from BUGS before projection. If the disparities between the GLM parameter estimates and the BUGS parameter estimates are material, then we use the former to help modify the prior distributions.

The BUGS parameter estimates are then grouped or linked so as to reduce the number of parameters. This exercise requires certain judgement about which parameters and how they are to be rearranged. The purpose is to achieve parsimony of parameters so as to reduce parameter error as a whole. BUGS provides a measure called DIC (deviance information criterion) that penalises both excessive use of parameters and poor data fitting (refer to the BUGS manual for details). We use DIC to assess whether the benefit of using fewer parameters outweighs the drawback of a poorer data fit. Overall, we attempt to discover the main underlying trends and prefer to select smooth patterns.

Apart from using DIC to compare the optimality of different parameter arrangements, we study various plots of residuals (for $i + j \leq n + 1$, where $n$ is the triangle dimension) to check whether there are any signs of distribution misfit, heteroscedasticity, or anomaly. The approach we adopt here is similar to that of traditional GLM modelling. Anscombe residuals as stated in McCullagh and Nelder (1989) are chosen because they are close to being normally distributed if the (GLM) model structure fits the data well and they are analytically solvable for the model structures considered in this paper. The residuals are 'standardised' to some extent by dividing them by the estimate of $\sqrt{\phi}$. The following means of examining (GLM) residuals are covered in McCullagh and Nelder (1989), Zehnwirth (1995), Taylor (2000), and Aitkin et al (2005):

(1) Plot the absolute values of the residuals against the scaled fitted values ($2\sqrt{\hat{X}_{i,j}}$ for ODP and $2\ln\left(\hat{X}_{i,j}\right)$ for gamma) – a positive trend reflects that the current variance of $X_{i,j}$ ($\phi\,\mu_{i,j}$ for ODP and $\phi\,\mu_{i,j}^2 = \mu_{i,j}^2 / r$ for gamma) assumed increases too slowly with the mean of $X_{i,j}$ (i.e. $\mu_{i,j}$) and a negative trend indicates the reverse.

(2) Form a Q-Q plot of the residuals vs $\Phi^{-1}\left((k-0.5)/n_r\right)$, where $\Phi^{-1}$ is the inverse standard normal cumulative distribution function, $k$ is the rank of the residual, and $n_r$ is the number of residuals – any deviance from the positive-slope diagonal line indicates some extent of non-normality of the residuals being unexplained by the model structure tested.

(3) Compare the empirical density of the residuals and the standard normal density – a significant difference between the two is a sign of non-normality of the residuals.

(4) Plot the residuals against the accident year, the development year, and the calendar year – a trend indicates a component dependent on that dimension being unexplained by the model structure tested, and any heteroscedasticity reflects that the actual variance depends on that dimension (weights can then be incorporated to reduce heteroscedasticity).

(5) Plot $\hat{\eta}_{i,j} + \left(X_{i,j} - \hat{X}_{i,j}\right) \left. \frac{d\eta_{i,j}}{d\mu_{i,j}} \right|_{\mu_{i,j}=\hat{X}_{i,j}}$ against $\hat{\eta}_{i,j}$ – an upward curvature shows that the power of $\mu_{i,j}$ in $\eta_{i,j}$ is too low and a downward curvature indicates the reverse.

If the residuals are evenly spread (i.e. not aggregated along a certain trend), there is no skewness in their empirical distribution, and there is no significant anomaly, the model structure can be regarded as being able to provide a reasonable fit over the past claims data. After all, claims modelling is not an exact science and remains partly an art. There are often several similarly optimal models (say, in terms of DIC) for a particular set of claims data. Hitherto, the residual diagnostic procedures are not completely formalised in the literature, despite the list of some model criticism procedures given above. A practitioner needs to strike a balance between different examinations of the residuals. Moreover, the future values of the parameters may vary by a large extent from those estimated as the conditions may change significantly over time. While scientific modelling is an expedient tool to help identify the past patterns, it cannot substitute practical judgement, in which one needs to determine the aspects of the modelling that are most important and also those that are most in doubt.

## 4. MOTOR INSURANCE

This section presents an analysis of the gross of reinsurance claim payments from accident years 1984 to 1996 of motor insurance in the Australian private sector. To start with, we employ the following model structure in BUGS (refer to Appendix II for an example of BUGS coding and further details of setting the priors), which is similar to the ODP assumption in GLMs, for $1 \leq i\, , j \leq 13$:

$X_{i,j} \sim \mathrm{N}(\mu_{i,j}\, , \phi\, \mu_{i,j})$ , $\eta_{i,j} = \ln(\mu_{i,j}) = v + \alpha_i + \beta_j$ , $v \sim \mathrm{N}(0\, , 1000)$ , $\alpha_i \sim \mathrm{N}(0\, , 100)$ ,

$\beta_j \sim \mathrm{N}(0\, , 100)$ , $\alpha_1 = \beta_1 = 0$ , and $\phi^{-1} \sim \gamma(1\, , 10^{-2})$ .

As noted earlier, the unknown parameters ($v$, $\alpha_i$, $\beta_j$, and $\phi$) are assumed to follow some vague non-informative prior distributions, of which the fixed parameters are chosen to allow BUGS to run smoothly. In addition, the gamma assumption is invalid for the motor insurance data here, which is indicated by the non-convergence of the corresponding GLM maximum likelihood estimation and by the awkward simulation results produced with BUGS. The gamma assumption is thus not considered in this section.

Figures 1 and 2 show the sample means of $\alpha_i$'s against the accident year and $\beta_j$'s against the development year respectively. As mentioned previously, we prefer to identify the main underlying trends and smooth them where possible. Accordingly, we find that two third-order polynomials fit the parameter estimates reasonably well, as shown in the two graphs. (We use the software CurveExpert to help identify the optimal polynomials via regression. This software is free for download from the Internet.)

The increase of the sample means of $\alpha_i$'s along the accident year suggests that the claim payments increase over time, along with the growth of the motor insurance business. The rapid decrease of the sample means of $\beta_j$'s along the first few development years is reasonable as most of the claim payments are settled within the first few years after occurrence for motor insurance.

**Figure 1 Sample Means of $\alpha_i$'s**

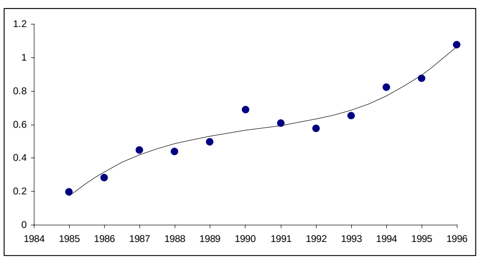
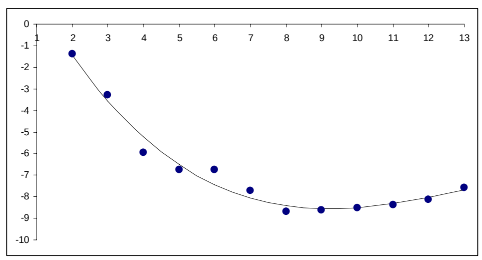



**Figure 2 Sample Means of $\beta_j$'s**




We then employ the following model structure in BUGS, in which two third-order polynomials are incorporated, for $1 \le i, j \le 13$:

$$X_{i,j} \sim \mathrm{N}(\mu_{i,j}, \phi\, \mu_{i,j}),\ \eta_{i,j} = \ln(\mu_{i,j}) = v + \alpha_i + \beta_j,\ v \sim \mathrm{N}(0, 1000),$$

$$\alpha_i = p_1 + p_2\, i + p_3\, i^2 + p_4\, i^3,\ \beta_j = p_5 + p_6\, j + p_7\, j^2 + p_8\, j^3,\ \alpha_1 = \beta_1 = 0,$$

$$p_k \sim \mathrm{N}(0, 100) \text{ for } 1 \le k \le 8, \text{ and } \phi^{-1} \sim \gamma(1, 10^{-2}).$$

Using fewer parameters in this model structure, DIC has decreased from 3,180.66 to 3,172.23. The benefit of a smaller number of parameters hence outweighs the drawback of less precise data fitting. Furthermore, Figures 3 to 10 demonstrate how we examine the Anscombe residuals. These plots include the residuals and their absolute values against the scaled fitted values, a Q-Q plot of the residuals against $\Phi^{-1}\left(\left(k-0.5\right)/n_r\right)$, the empirical density of the residuals against the standard normal density, the residuals against the accident year, development year, and calendar year, and a plot of $\hat{\eta}_{i,j}+\left(X_{i,j}-\hat{X}_{i,j}\right)/\hat{X}_{i,j}$ against $\hat{\eta}_{i,j}$. The straight line in Figure 4 indicates the major trend of the (absolute) residuals along the scaled fitted values. The zigzag lines in Figures 7 to 9 join the mean of the residuals of each unit along the *x*-axis and so these lines reflect the existence of any unmodelled trends along the accident year, development year, and calendar year.

**Figure 3 Residuals vs Fitted Values**

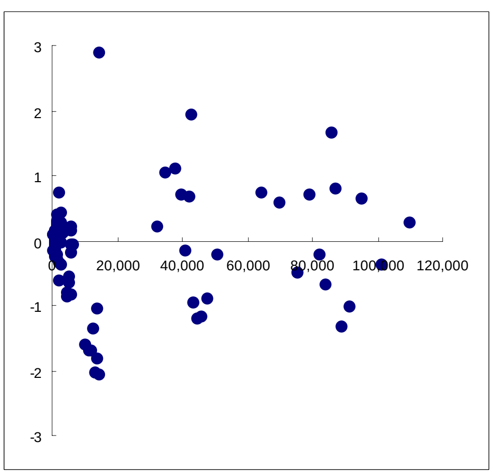



**Figure 4 Absolute Residuals vs Fitted Values**

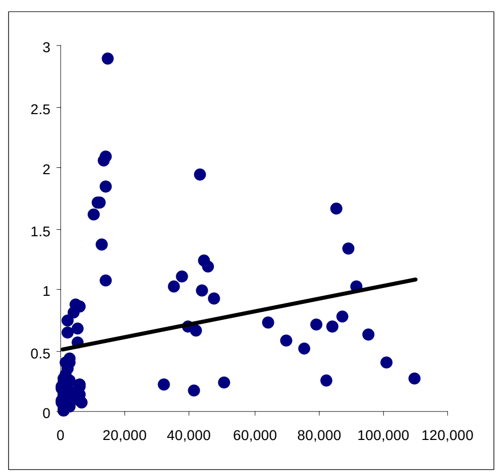


**Figure 5 Q-Q Plot of Residuals**

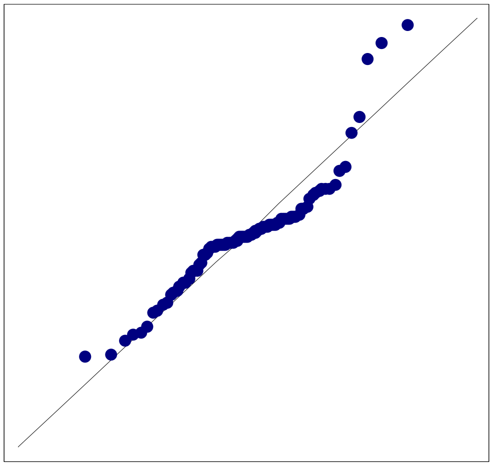


**Figure 6 Densities of Residuals and N(0,1)**

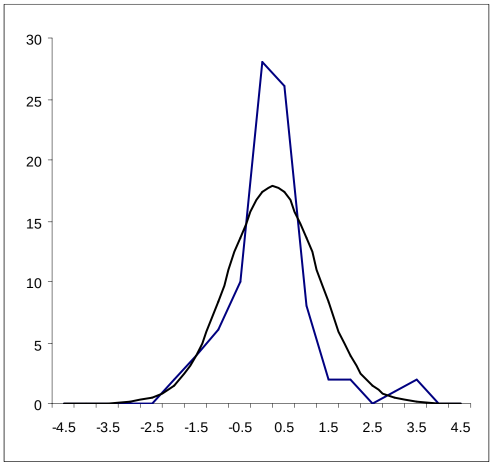



**Figure 7 Residuals vs Accident Year**

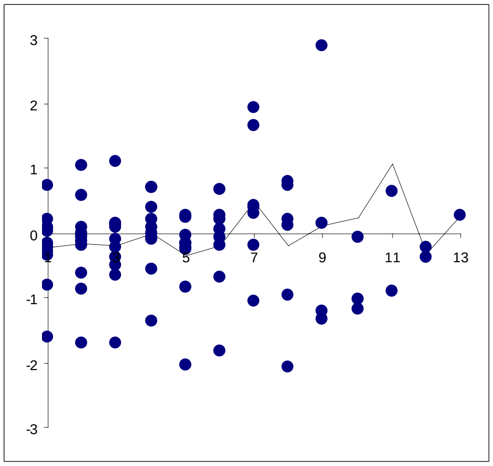



**Figure 8 Residuals vs Development Year**

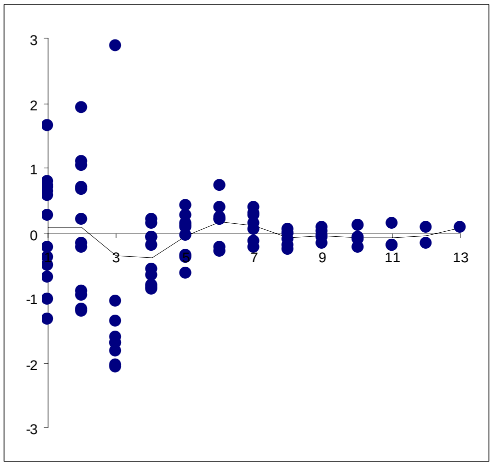



**Figure 9 Residuals vs Calendar Year**

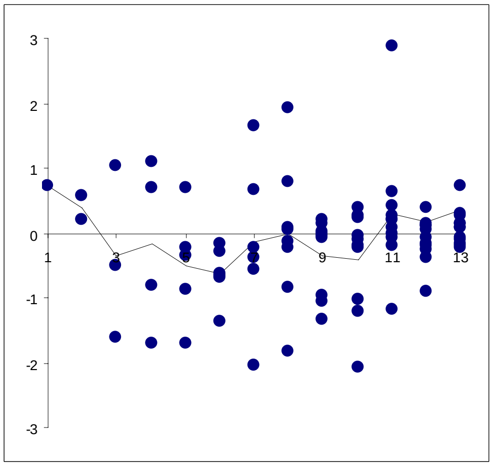



**Figure 10** $\hat{\eta}_{i,j} + \left(X_{i,j} - \hat{X}_{i,j}\right)/\hat{X}_{i,j}$ **vs** $\hat{\eta}_{i,j}$

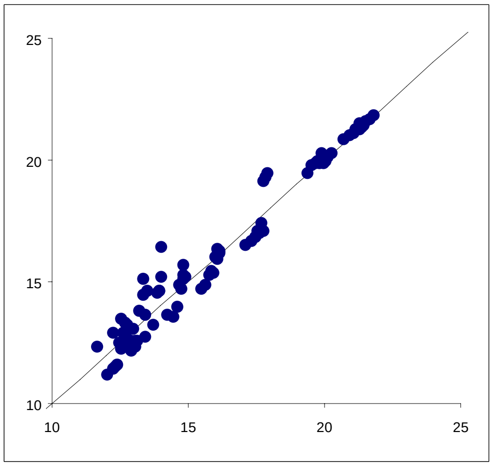


Figure 4 reflects that the variance of $X_{i,j}$ ($\phi\,\mu_{i,j}$ for ODP) assumed increases not as fast as the mean of $X_{i,j}$. Since the trend of the residuals is only slightly positive, the ODP assumption still sounds quite reasonable in this regard. In fact, one way to alleviate this type of problem is to increase the power of $\mu_{i,j}$ in the variance term, e.g. the variance is $\phi\,\mu_{i,j}^2$ for the gamma assumption. This gamma assumption, however, is invalid for the motor insurance data as discussed previously. (For GLMs, one may use a non-integral power of $\mu_{i,j}$ in the variance term, as demonstrated in Taylor and McGuire (2004). We find that a power of 1.8 still leads to convergence. In contrast, BUGS does not allow such an arrangement.)

Figures 5 and 6 show that the residuals are fairly close to being normally distributed, with a small degree of right skewness and some extent of kurtosis in their distribution. This skewness points out that the actual underlying probability distribution of $X_{i,j}$ should have a larger skewness than that assumed (e.g. the gamma distribution has a larger skewness than the over-dispersed Poisson distribution). Again, the effect is not overly significant and so can be disregarded as an approximation.

Figures 7 to 9 indicate that there are no obvious trends along the accident year, development year, and calendar year, though there are some mild fluctuations along the accident year and calendar year. Thus, the parameter arrangement seems justifiable. On the other hand, the uneven distribution of the residuals, especially along the development year, reflects that there is significant heteroscedasticity. It then becomes necessary to incorporate some weights into the model structure to alleviate this situation. Using the approach as stated in Taylor (2000), we adjust the structure with weights depending on the development year, using the reciprocal of the sample variance of the residuals of each development year. Figure 10 shows that there is no apparent curvature of the residuals along the positive-slope diagonal line. This phenomenon pinpoints that the power of $\mu_{i,j}$ in $\eta_{i,j}$ is reasonable, i.e. the linear relationship $\eta_{i,j} = \ln\left(\mu_{i,j}\right) = v + \alpha_i + \beta_j$ is warranted.

We now adjust the model structure with weights for $1 \le i\,,\, j \le 13$ as follows:

$X_{i,j} \sim \mathrm{N}\left(\mu_{i,j}\,, \phi\,\mu_{i,j} / w_j\right)$ , $w_j = 2 - 0.6j$ for $1 \le j \le 3$ ,

and $w_j = -5.9 + 2.8j$ for $4 \le j \le 13$ .

(We use CurveExpert to help identify the key patterns of the weights.)

The following graphs of residuals are produced from the adjusted structure. The incorporation of weights appears to have two main effects. The first effect is to make the modelling better and the residuals closer to being normally distributed, as reflected by the flat trend in Figure 12 and the adherence of the residuals to the diagonal line in Figure 13. The second effect is to greatly reduce heteroscedasticity, as shown in Figures 15 to 17. Hence the weight adjustment seems justifiable in this situation.

**Figure 11 Residuals vs Fitted Values**

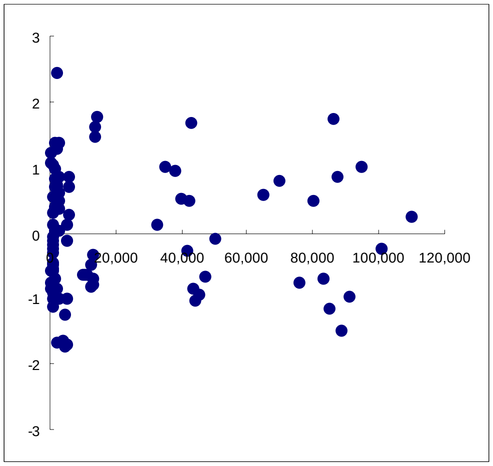



**Figure 12 Absolute Residuals vs Fitted Values**

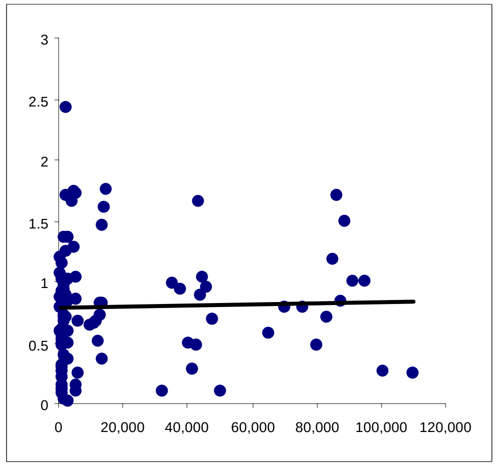



**Figure 13 Q-Q Plot of Residuals**

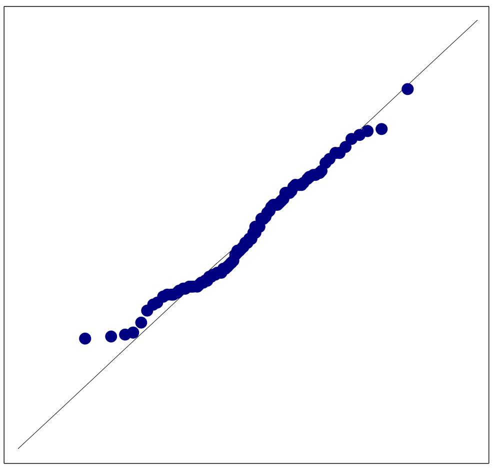


**Figure 14 Densities of Residuals and N(0,1)**

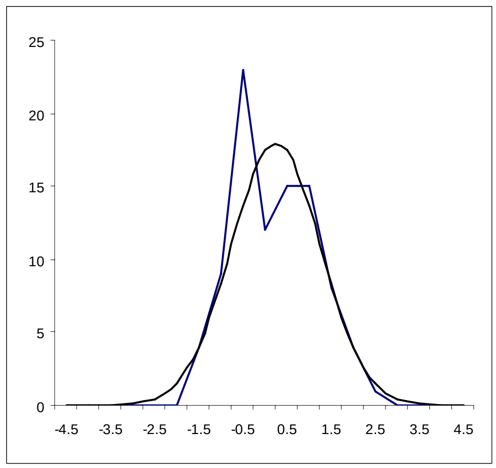



**Figure 15 Residuals vs Accident Year**

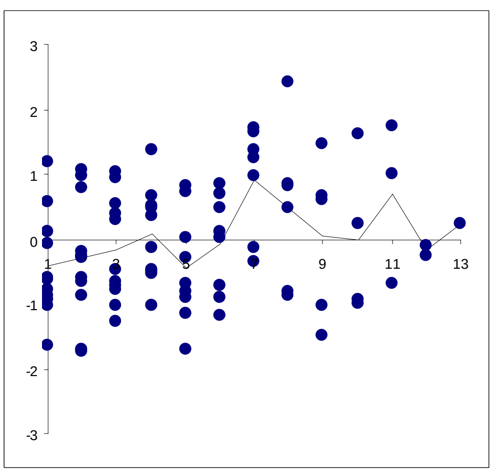



**Figure 16 Residuals vs Development Year**

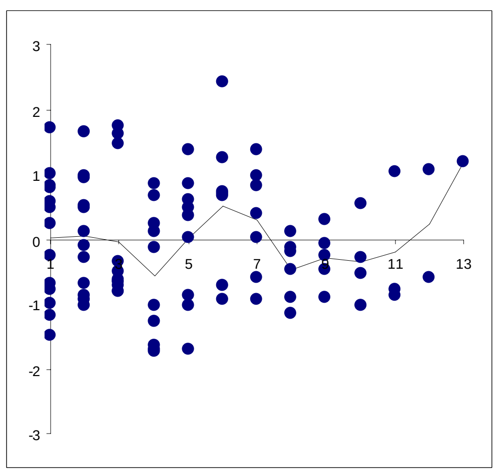


**Figure 17 Residuals vs Calendar Year**

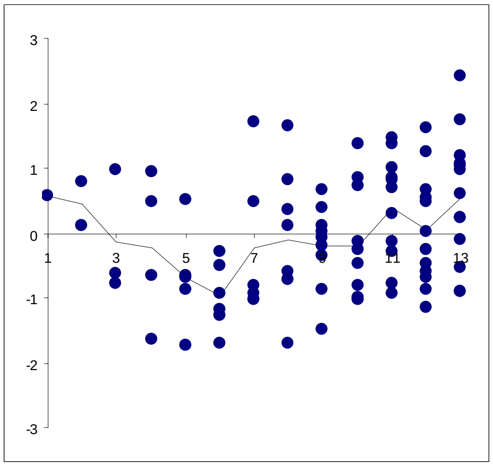
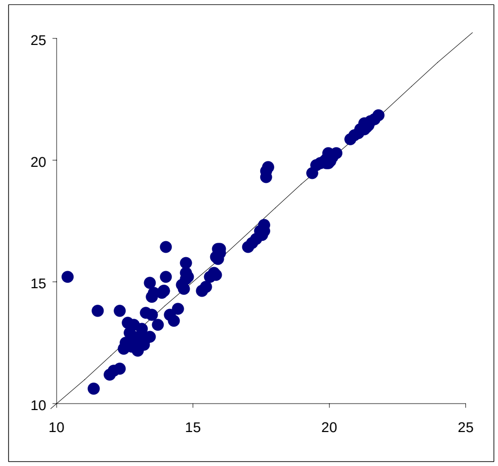



**Figure 18** $\hat{\eta}_{i,j} + \left(X_{i,j} - \hat{X}_{i,j}\right) / \hat{X}_{i,j}$ **vs** $\hat{\eta}_{i,j}$




There are two further points to note. First, the empirical density of the residuals in Figure 14 has two peaks. This problem is not material as the number of residuals is small (only 91 residuals in this case) and there is no skewness at the tails. Second, mild trends of the residuals along the few latest development years and along the few earliest calendar years do not cause serious problems, as the former case is related to a very small proportion of the total outstanding claims liability, and the latter case refers only to the past inflation effects several years ago and so is not quite related to the total outstanding claims liability to be assessed. From calendar year 6 onwards, however, there is a small positive trend of the residuals being unexplained by the model structure. This trend may reflect the existence of some superimposed inflation. Overall, the model fitting can hardly be perfect and one needs to strike a balance between the different ways to examine the residuals. These comments apply similarly to the analyses in the later sections.

Based on this adjusted structure, Table 1 below presents our estimates of the mean, one-half the coefficient of variation, and the 75th percentile margin (as a percentage of the mean) of the gross of reinsurance outstanding claims liability in 31 December 1996 dollar values.

**Table 1 Mean, Half of Coefficient of Variation, and 75th Percentile Margin of Outstanding Claims Liability of Motor Insurance**

| Accident Year | Mean | 0.5 CV | 75th Percentile Margin |
|---|---|---|---|
| **1985** | 42,650 | 125.9% | 155.3% |
| **1986** | 183,800 | 66.4% | 80.0% |
| **1987** | 470,900 | 42.7% | 53.0% |
| **1988** | 893,400 | 33.5% | 43.5% |
| **1989** | 1,443,000 | 27.4% | 35.8% |
| **1990** | 2,081,000 | 23.6% | 30.4% |
| **1991** | 2,964,000 | 21.3% | 27.9% |
| **1992** | 4,332,000 | 19.4% | 25.4% |
| **1993** | 7,479,000 | 16.8% | 21.8% |
| **1994** | 18,080,000 | 13.4% | 17.9% |
| **1995** | 80,770,000 | 31.6% | 44.0% |
| **1996** | 861,300,000 | 6.8% | 9.0% |
| **Total** | 980,039,750 | 6.6% | 8.9% |

# 5. PUBLIC LIABILITY

This section presents an analysis of the gross of reinsurance claim payments from accident years 1983 to 1996 of public liability insurance in the Australian private sector. Figures 19 to 22 show the sample means of $\alpha_i$'s against the accident year and $\beta_j$'s against the development year with the initial ODP and gamma model structures. For both model structures, we find that the $\alpha_i$'s can be fitted with a straight line and the $\beta_j$'s can be fitted with a quadratic function. For the approximate ODP structure, additional parameters are needed for accident years 1990, 1994, and 1995 and for development years 6, 11, 12, and 13, since the corresponding parameter estimates depart prominently from the rest. Similarly, for the gamma structure, extra parameters are required for accident years 1994 and 1995 and for development years 6, 11, 12, and 13. (We find that DIC is smaller if extra parameters are accommodated for these accident years and development years.)

The increase of the sample means of $\alpha_i$'s along the accident year suggests that the claim payments increase as the private sector of the public liability insurance business grows over time. The decrease of the sample means of $\beta_j$'s along the first few development years is gradual because it takes several years for the claim payments to develop after occurrence.

**Figure 19 Sample Means of $\alpha_i$'s (ODP)**

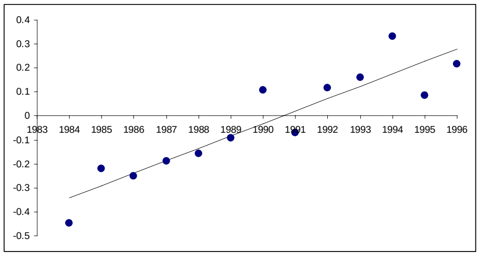



**Figure 20 Sample Means of $\beta_j$'s (ODP)**

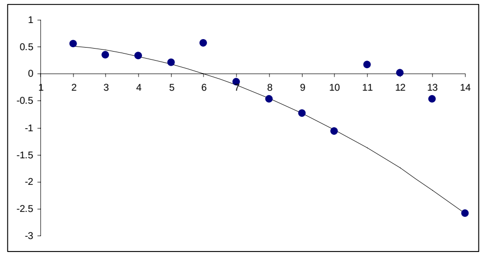



**Figure 21 Sample Means of $\alpha_i$'s (Gamma)**

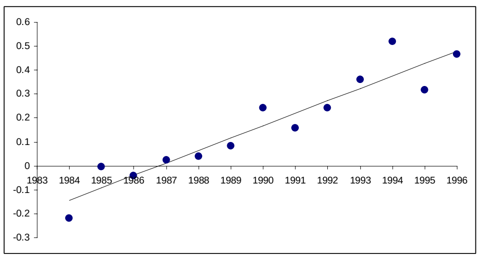
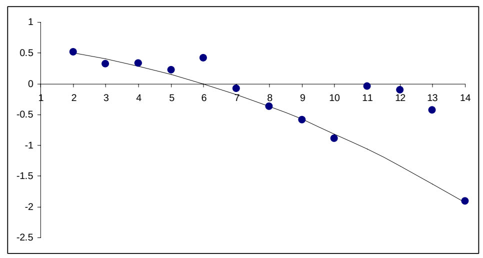



**Figure 22 Sample Means of $\beta_j$'s (Gamma)**




We find that for the ODP structure the variance of $X_{i,j}$ (i.e. $\phi\,\mu_{i,j}$) assumed increases slower than the mean of $X_{i,j}$. As mentioned previously, the power of $\mu_{i,j}$ in the variance term can be increased to tackle the problem. In this case, the gamma assumption can be considered, in which the variance is $\phi\,\mu_{i,j}^2$. We then find that for the gamma structure there is no trend of the absolute values of the residuals along the scaled fitted values, and so the gamma assumption looks more reasonable than the ODP assumption. We thus focus on the gamma assumption for the rest of this section.

After adjusting with weights for heteroscedasticity as in the previous section, we come up with the following model structure for $1 \le i$ , $j \le 14$ :

$X_{i,j} \sim \gamma(\mu_{i,j}, r\, w_j)$ , $w_j = 0.77 + 0.49 j$ for $1 \le j \le 10$ ,

and $w_j = 0.26 + 0.01 j$ for $11 \le j \le 14$ .

The following graphs of residuals are produced from the adjusted structure. As shown in Figures 27 to 29, there are no significantly consistent trends along the accident year, development year, and calendar year and there is no obvious heteroscedasticity after weight adjustment. In addition, the fluctuations of the residuals along the calendar year indicate that AWOTE may not be suitable for adjusting the claim payments for public liability insurance. Finally, Table 2 presents our estimation results.

**Figure 23 Residuals vs Fitted Values**

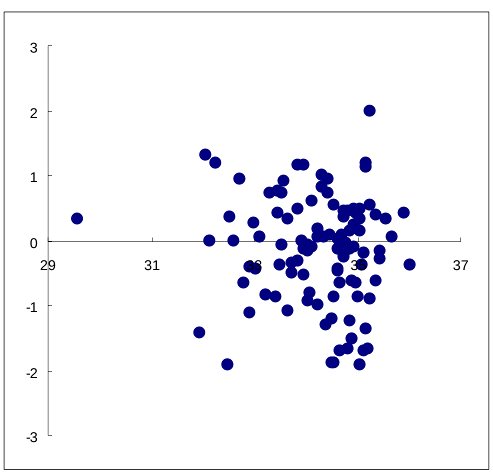



**Figure 24 Absolute Residuals vs Fitted Values**

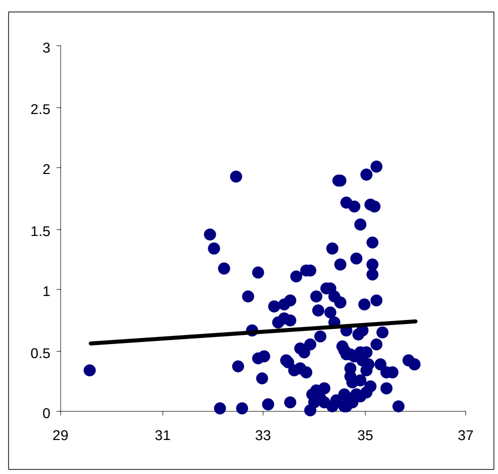


**Figure 25 Q-Q Plot of Residuals**

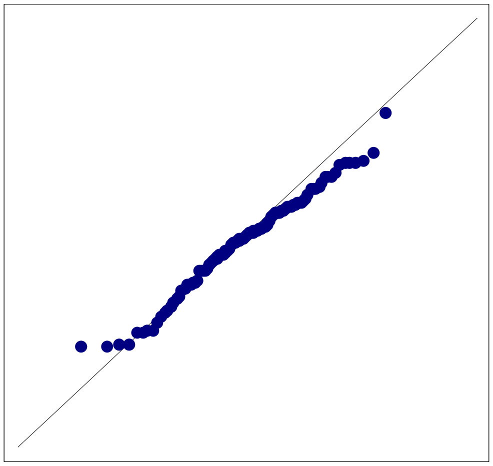


**Figure 26 Densities of Residuals and N(0,1)**

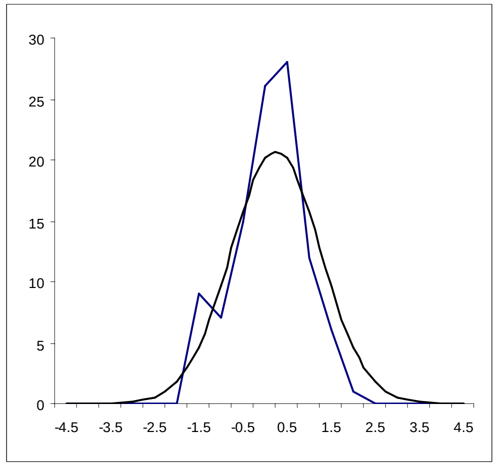



**Figure 27 Residuals vs Accident Year**

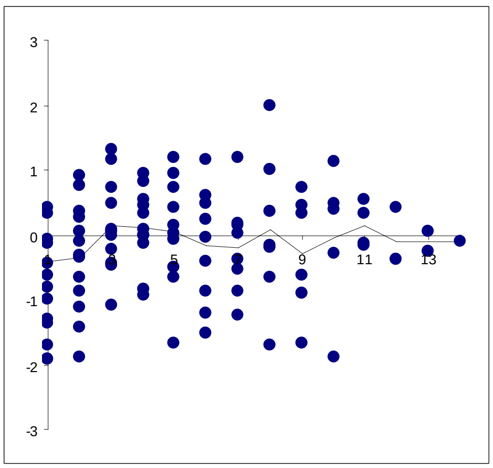



**Figure 28 Residuals vs Development Year**

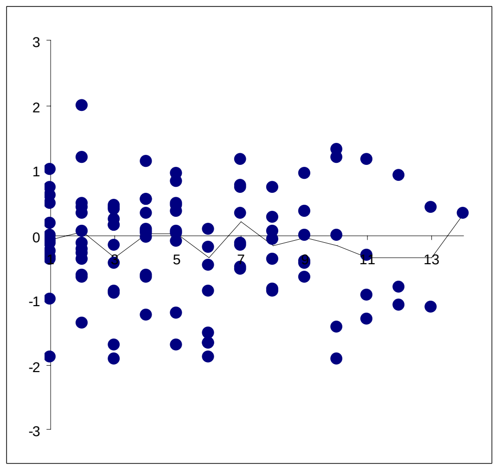



**Figure 29 Residuals vs Calendar Year**

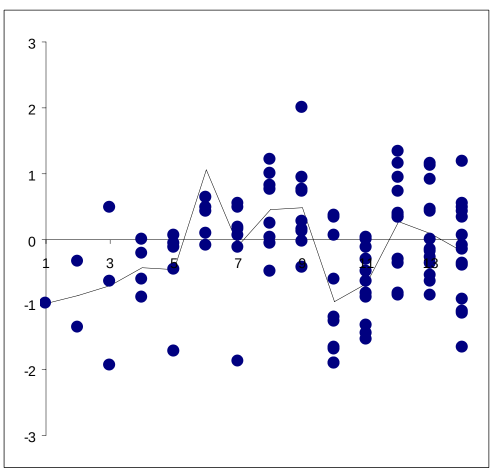
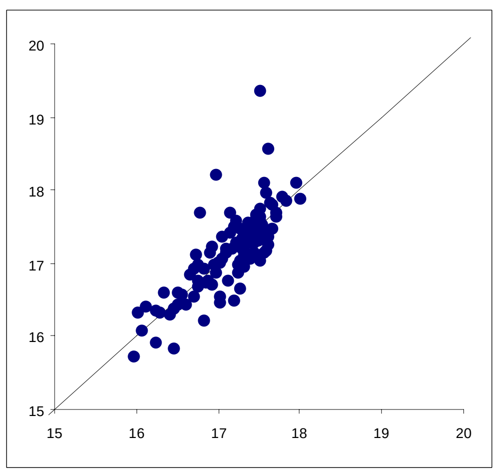



**Figure 30** $\hat{\eta}_{i,j} + \left(X_{i,j} - \hat{X}_{i,j}\right) / \hat{X}_{i,j}$ **vs** $\hat{\eta}_{i,j}$



## Table 2 Mean, Half of Coefficient of Variation, and 75th Percentile Margin of Outstanding Claims Liability of Public Liability Insurance

| Accident Year | Mean | 0.5 CV | 75th Percentile Margin |
|---|---|---|---|
| **1984** | 2,143,000 | 43.7% | 32.2% |
| **1985** | 20,500,000 | 77.3% | 18.7% |
| **1986** | 45,280,000 | 40.5% | 22.5% |
| **1987** | 77,140,000 | 32.1% | 22.6% |
| **1988** | 92,830,000 | 30.5% | 21.5% |
| **1989** | 111,200,000 | 25.9% | 18.7% |
| **1990** | 137,700,000 | 21.6% | 16.3% |
| **1991** | 171,300,000 | 20.4% | 13.9% |
| **1992** | 226,300,000 | 16.4% | 11.8% |
| **1993** | 280,400,000 | 13.6% | 11.8% |
| **1994** | 420,800,000 | 17.1% | 15.4% |
| **1995** | 421,700,000 | 22.3% | 16.2% |
| **1996** | 500,100,000 | 10.6% | 11.1% |
| **Total** | 2,507,393,000 | 11.1% | 8.6% |

# 6. COMPULSORY THIRD PARTY

This section demonstrates an analysis of the gross of reinsurance claim payments from accident years 1983 to 1996 of CTP insurance in the Australian private sector. Figures 31 to 34 show the sample means of $\alpha_i$'s against the accident year and $\beta_j$'s against the development year with the initial ODP and gamma model structures. For both structures, we find that the $\alpha_i$'s can be fitted with three separate straight lines and the $\beta_j$'s can be fitted with a fourth-order polynomial.

The increase of the sample means of $\alpha_i$'s along the accident year is split into a few segments, which indicate certain separate shifts of business into the private sector over time. The sample means of $\beta_j$'s increase then decrease along the development year as many of the claim payments are made only a few years after occurrence.

**Figure 31 Sample Means of $\alpha_i$'s (ODP)**

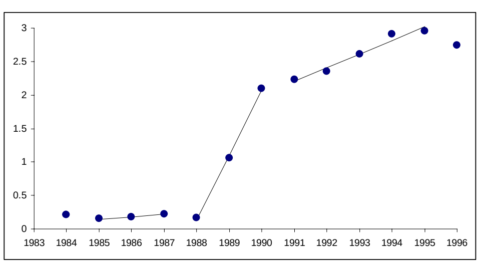



**Figure 32 Sample Means of $\beta_j$'s (ODP)**

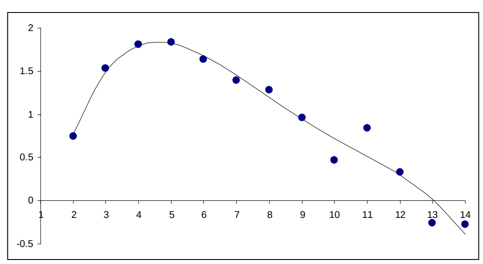


**Figure 33 Sample Means of $\alpha_i$'s (Gamma)**

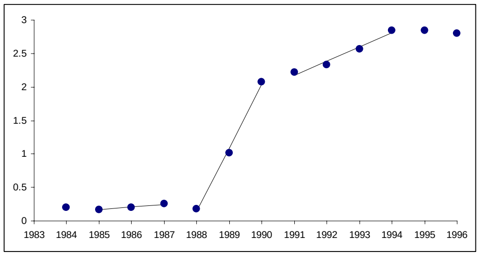
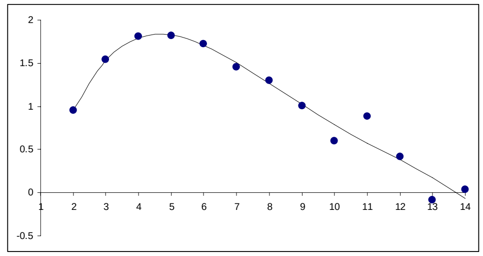



**Figure 34 Sample Means of $\beta_j$'s (Gamma)**




Again, we find that the gamma assumption is more suitable than the ODP assumption here. After weight adjustment, we come up with the following model structure for $1 \le i, j \le 14$:

$X_{i,j} \sim \gamma(\mu_{i,j}, r\, w_j)$ , $w_j = 0.84 + 0.21j$ for $1 \le j \le 5$ ,

and $w_j = 2.8 - 0.19j$ for $6 \le j \le 14$ .

The following graphs of residuals are produced based on the adjusted structure. As shown in Figures 39 and 40, there are no significant trends along the accident year and development year with some mild fluctuations. On the other hand, it can be seen in Figure 41 that there is a rough positive trend of the residuals from calendar year 7 onwards. This trend reflects that there may be some form of superimposed inflation being unexplained or AWOTE may not be a suitable inflation index for CTP insurance. We then try to add extra parameters to allow for this trend but the results are awkward and DIC increases by a large extent. Hence we disregard the trend as an approximation. Finally, Table 3 presents our estimation results.

**Figure 35 Residuals vs Fitted Values**

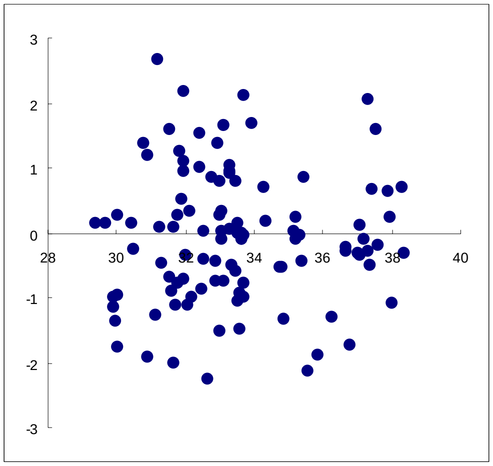



**Figure 36 Absolute Residuals vs Fitted Values**

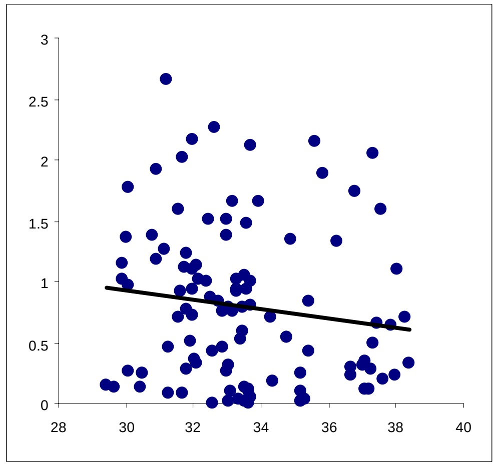



**Figure 37 Q-Q Plot of Residuals**

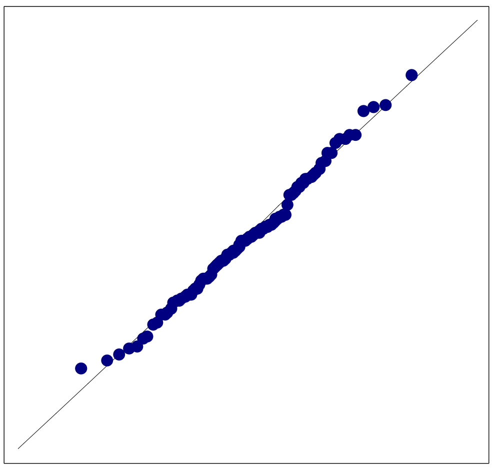


**Figure 38 Densities of Residuals and N(0,1)**

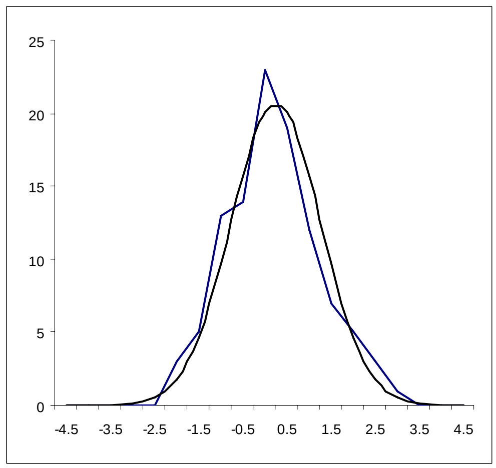



**Figure 39 Residuals vs Accident Year**

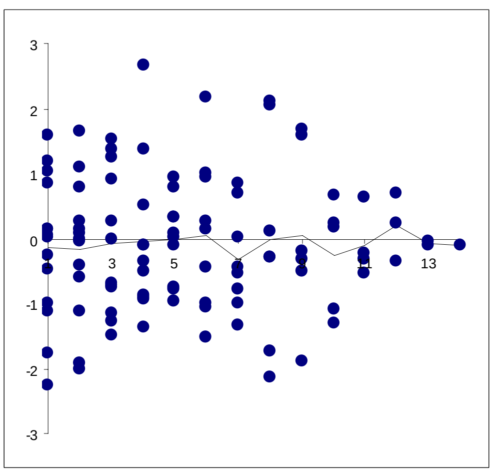



**Figure 40 Residuals vs Development Year**

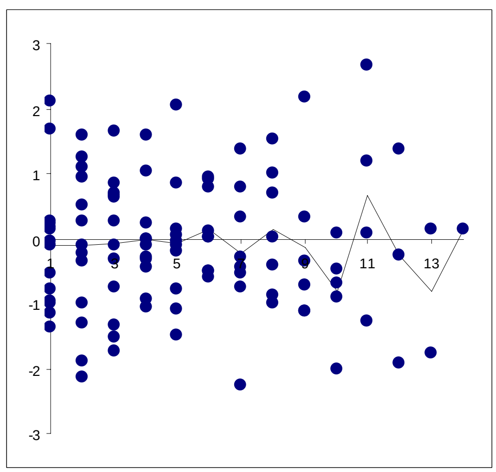


**Figure 41 Residuals vs Calendar Year**

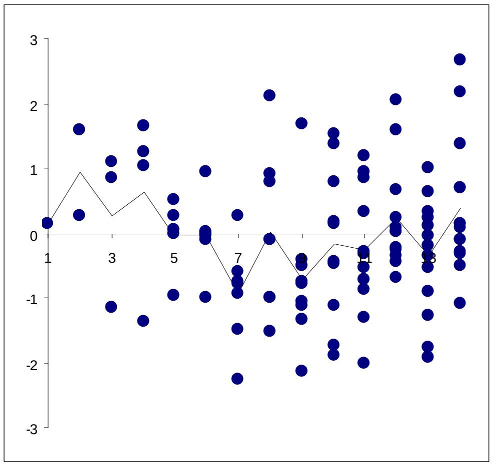
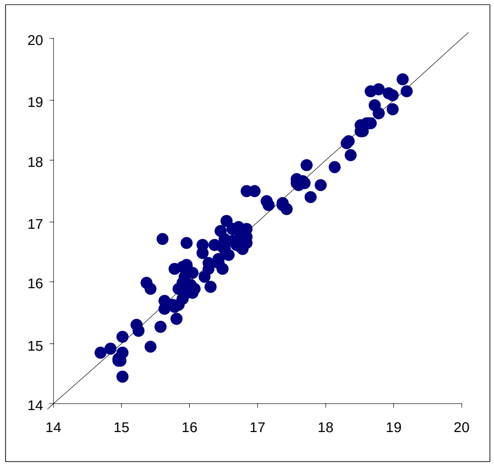



**Figure 42** $\hat{\eta}_{i,j} + \left(X_{i,j} - \hat{X}_{i,j}\right) / \hat{X}_{i,j}$ **vs** $\hat{\eta}_{i,j}$




## Table 3 Mean, Half of Coefficient of Variation, and 75th Percentile Margin of Outstanding Claims Liability of CTP Insurance

| Accident Year | Mean | 0.5 CV | 75th Percentile Margin |
|---|---|---|---|
| **1984** | 3,333,000 | 50.5% | 29.1% |
| **1985** | 7,123,000 | 32.0% | 22.6% |
| **1986** | 12,530,000 | 22.1% | 18.6% |
| **1987** | 19,250,000 | 15.7% | 15.7% |
| **1988** | 24,900,000 | 12.8% | 12.9% |
| **1989** | 86,290,000 | 9.8% | 11.3% |
| **1990** | 296,900,000 | 8.9% | 10.3% |
| **1991** | 445,500,000 | 8.2% | 9.9% |
| **1992** | 707,600,000 | 6.7% | 8.3% |
| **1993** | 1,089,000,000 | 6.5% | 7.9% |
| **1994** | 1,613,000,000 | 7.6% | 9.5% |
| **1995** | 1,971,000,000 | 11.1% | 12.1% |
| **1996** | 2,017,000,000 | 16.4% | 15.2% |
| **Total** | 8,293,426,000 | 5.8% | 6.9% |

# 7. COMPARISON WITH INDUSTRY FIGURES

Table 4 below presents a comparison between the aggregate risk margin estimates (gross of reinsurance) computed in Sections 4 to 6 and some industry figures for the three lines of business under consideration. The industry figures, which are net of reinsurance, are collected from Bateup and Reed (2001), Collings and White (2001), and some APRA statistics (industry average for the year ended June 2002) released in Actuary Australia (July 2003).

**Table 4 Aggregate Risk Margin of Total Outstanding Claims Liability**

| Line of Business | Bayesian MCMC | Bateup & Reed (2001) | Collings & White (2001) | APRA Statistics (2002) |
|---|---|---|---|---|
| **Motor*** | 9% | 6% | 5% | 6% |
| **Public Liability** | 11% | 10% | 7% | 8% |
| **CTP** | 7% | 12% | 7% | 10% |

* For Bateup and Reed (2001) and Collings and White (2001), an average is computed between domestic motor and commercial motor insurance. For the APRA statistics, the figure is only for domestic motor insurance.

For motor insurance and public liability insurance, the MCMC figures are larger than the industry figures. This phenomenon makes intuitive sense as the MCMC figures are gross of reinsurance and the industry figures are net of reinsurance, in which the gross of reinsurance claim amounts are generally perceived to involve higher uncertainty and the net of reinsurance claim amounts have their uncertainty reduced by reinsurance. (As noted in Barnett and Zehnwirth (2005), however, net of reinsurance liabilities have higher uncertainty sometimes when certain non-proportional reinsurance arrangements do not work properly to reduce uncertainty for an insurer.) Moreover, as the MCMC figures computed here are based on the claims data of the whole general insurance private sector, there are additional diversification benefit effects between different insurers in the private sector included in the figures. In this regard, an individual insurer does not have these diversification benefits and so may consider larger aggregate risk margin estimates than the MCMC figures here for its gross of reinsurance total outstanding claims liabilities.

In contrast, the MCMC figure is smaller than the industry figures for CTP insurance. This result appears to arise from certain data insufficiency, in which there are separate shifts of business into the private sector as mentioned in the previous section.

Finally, we have collected the aggregated 'Payments on Claims other than Indirect Claims Settlement Costs' and 'Payments on Indirect Claims Settlement Costs' of the private sector from the 'Selected Statistics on the General Insurance Industry' for payment years 1997 to 2002. These actual figures are modified (refer to Appendix I for details) to provide some rough estimates of the actual gross of reinsurance claim payments of the three lines of business. These 'modified' actual figures are then compared with the corresponding MCMC figures (expected total outstanding claims liability plus aggregate risk margin) generated previously, as shown in Table 5 below.

This practice is an experience analysis to some extent and the modification of the actual figures is similar to the 'hindsight re-estimate' concept stated in Houltram (2005).

**Table 5 Expected Total Outstanding Claims Liability plus Aggregate Risk Margin vs Actual Payments**

| Line of Business | MCMC Estimate – Expected Value | MCMC Estimate – 75% Sufficiency | Actual Outcome – Modified | Actual Outcome / MCMC Estimate |
|---|---|---|---|---|
| **Motor** | 980 | 1,067 | 900 | 84% |
| **Public Liability** | 2,507 | 2,787 | 2,276 | 82% |
| **CTP** | 8,293 | 8,867 | 6,112 | 69% |

For motor insurance and public liability insurance, our MCMC estimates of the expected total outstanding claims liability plus the aggregate risk margin (gross of reinsurance) under GPS 210 cover the 'actual future experience' sufficiently but not excessively. In particular, our estimates of the expected total outstanding claims liability are comparable to the actual outcome in magnitude.

For CTP insurance, it appears that there is a certain degree of overestimation. As aforementioned there are separate shifts of business over time, so the past development patterns may not be suitable for projecting the claims of the most recent accident years, in which the expected outstanding claims liabilities of accident years 1993 to 1996 account for more than 80% of the expected total outstanding claims liability.

# 8. DISCUSSION

In this paper, we demonstrate that Bayesian models with MCMC simulation provide a sophisticated framework for modelling the past claims data. Its application is straightforward via BUGS and it allows different prior and posterior distributions and different parameter arrangements. Examination of residuals is useful for justifying a particular model structure. The underlying trends can be readily identified. Other claims run-off formats such as a trapezium or individual claims data can be modelled properly. Moreover, the software BUGS is user-friendly and offers a convenient platform for coding a variety of Bayesian models.

For motor insurance and public liability insurance, our MCMC estimates of the expected total outstanding claims liability plus the aggregate risk margin (gross of reinsurance) cover the actual claim payments made after the valuation date sufficiently but not excessively. In particular, we find that the BUGS parameter estimates are generally comparable to the corresponding GLM parameter estimates (refer to Appendix III for details). We also show that the BUGS parameters estimates can be smoothed by fitting some lines or curves via regression, which lead to more parsimonious use of parameters.

Nevertheless, there are some limitations behind the application of BUGS. First, the ODP structure adopted is only an approximation. There is no exact way to incorporate over-dispersion into the Poisson assumption with BUGS. Second, a higher power (> 2) of $\mu_{i,j}$ in the variance term (e.g. $\phi\,\mu_{i,j}^3$ for inverse Gaussian) has to be allowed for by using the ‘zeros trick’ (described in the BUGS manual), as BUGS does not define some distributions explicitly. Third, as noted in Scollnik (2001), the prior distributions cannot be too weak. Otherwise, BUGS may have problems in its simulation process. Fourth, the MCMC simulation process so far depends heavily on the BUGS platform. This software has been under wide public testing for around fifteen years and its reliability will be further improved after more testing or being released as a commercial package in the future. One may also develop its own MCMC simulation process with certain programming languages by using the techniques described in Gilks et al (1996) and Brooks (1998).

Last but not the least, there are a few points to note regarding the analysis of the claims data in this paper. First, for the three lines of business concerned, though there are no significant trends of the residuals along the calendar year, the patterns seem to be less random than those along the accident year and the development year. This phenomenon may indicate a degree of insufficiency of using AWOTE to adjust the claim payments for ‘normal’ inflation, or may suggest that there is some extent of superimposed inflation not being modelled properly. Even if the past superimposed inflation may be identified in some cases, it is often difficult to project it into the future in practice. We have tried to model the past calendar period effect with extra parameters, with the result that either the parameter estimates are awkward, there is no convergence, DIC increases by a large extent, or it becomes too arbitrary to project the past trends. Use of individual claims data, if available, may give better results when modelling the inflation effects. Second, the adjustments to the cells in the upper right corner of each run-off triangle are only approximations and these adjustments may have a significant impact on the estimated values of the parameters and on the reasonableness of the trends identified. Determination of the run-off patterns beyond the latest development year in the data, if necessary, is also highly subjective. Third, as noted in Section 2, there are several sources of data insufficiency. One has to be careful in the interpretation of the results here. Fourth, to reduce the autocorrelations between the simulated samples, we select (at least) every $50^{th}$ iteration to contribute to the required statistics. This practice can significantly lengthen the simulation time. Finally, different persons can end up with the same model structure through different routes, or different persons can end up with equally valid but different model structures. The true values of the parameters may also change as future conditions change. The final decision often depends on the experience and preference of the user.

# 9. ACKNOWLEDGEMENT

The author would like to thank Prof. David Dickson and Dr. Greg Taylor for reviewing the contents of this paper. Responsibility for the opinions expressed and the calculations, however, remains with the author. The author would also like to thank the Institute of Actuaries of Australia for their financial support through the A H Pollard Scholarship.

# 10. BIBLIOGRAPHY

Aitkin M., Francis B., and Hinde J. (2005). *Generalized linear models: a framework for statistical modelling with GLIM4*. Oxford University Press.

Australian Bureau of Statistics (ABS 1983-2004). *Average Weekly Ordinary Time Earnings (AWOTE)*. www.abs.gov.au

Australian Prudential Regulation Authority (APRA 1973). *Insurance Act 1973*.

Australian Prudential Regulation Authority (APRA 2002). *Prudential Standard GPS 210 and Guidance Note GGN 210.1*. www.apra.gov.au

Australian Prudential Regulation Authority (APRA 1983-2002). *Selected statistics on the general insurance industry*. www.apra.gov.au

Australian Prudential Regulation Authority (APRA 2006). *Prudential Standard GPS 310 and Guidance Note GGN 310.1 and 310.2*. www.apra.gov.au

Barnett G. and Zehnwirth B. (2005). *Reinsurance of long tail liabilities*. The Institute of Actuaries of Australia Research Conference – New Developments in the Quantitative Modelling of Insurance and Financial Risk.

Bateup R. and Reed I. (2001). *Research and data analysis relevant to the development of standards and guidelines on liability valuation for general insurance*. The Institute of Actuaries of Australia and Tillinghast – Towers Perrin. www.actuaries.asn.au

Brooks S. P. (1998). *Markov chain Monte Carlo method and its application*. The Statistician 47: 69-100.

Collings S. and White G. (2001). *APRA risk margin analysis*. The Institute of Actuaries of Australia XIII$^{th}$ General Insurance Seminar, Trowbridge Consulting. www.actuaries.asn.au

De Alba E. (2002a). *Bayesian estimation of outstanding claims reserves*. North American Actuarial Journal 6(4): 1-20.

De Alba E. (2002b). *Claims reserving when there are negative values in the development triangle*. 37$^{th}$ Actuarial Research Conference, The University of Waterloo.

De Alba E. (2003). *Bayesian claims reserving*. No. 16, Institute of Insurance and Pension Research.

Gilks W. R., Richardson S., and Spiegelhalter D. J. (1996). *Markov chain Monte Carlo in practice*. Chapman and Hall, London, UK.

Hart D. G., Buchanan R. A., and Howe B. A. (1996). *The actuarial practice of general insurance*. The Institute of Actuaries of Australia.

Houltram A. (2005). *An empirical approach to insurance liability prediction error assessment*. The Institute of Actuaries of Australia XV$^{th}$ General Insurance Seminar.

Hyams D. (2001). *CurveExpert 1.3*. http://curveexpert.webhop.biz/

McCullagh P. and Nelder J. A. (1989). *Generalised linear models (3$^{rd}$ edition)*. Chapman and Hall, New York.

Ntzoufras I. and Dellaportas P. (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.

Spiegelhalter D., Thomas A., Best N., and Lunn D. (2003). *WinBUGS user manual*. MRC Biostatistics Unit, Institute of Public Health, Cambridge.

Taylor G. (2000). *Loss reserving – an actuarial perspective*. Kluwer Academic Publishers.

Taylor G. and McGuire G. (2004). *Loss reserving with GLMs: a case study*. Research Paper Number 113, Centre for Actuarial Studies, The University of Melbourne.

The Institute of Actuaries of Australia (IAA July 2003). *Actuary Australia*. www.actuaries.asn.au

The Institute of Actuaries of Australia (IAA 2002). *Professional Standard 300 and Guidance Note GN 353*. www.actuaries.asn.au

Verrall R. J. (1990). *Bayes and empirical Bayes estimation for the chain ladder model*. ASTIN Bulletin 20: 217-243.

Verrall R. J. (1999). *Analytic and bootstrap estimate of prediction errors in claims reserving*. Insurance: Mathematics and Economics 25: 281-293.

Zehnwirth B. (1995). *Outstanding claim liabilities: are they predictable?* Research Paper Number 18, Centre for Actuarial Studies, The University of Melbourne.

# Appendix I

Assume $X_{i,j}$ represents the gross of reinsurance claim payments of accident year $i$ and development year $j$ for public liability insurance. Referring to Section 2, all the claims data available actually have the format of:

$$\left(X_{1983,1}, X_{1984,1}, \ldots, X_{1996,1}\right), \left(X_{1982,2}, X_{1983,2}, \ldots, X_{1995,2}\right), \ldots, \left(X_{1973,11}, X_{1974,11}, \ldots, X_{1986,11}\right),$$

where the last set of data (development year 11) is different to the others as each component of it refers to claims of several accident years, e.g. $X_{1986,11}$ is in fact equal to $\sum_{i \le 1986} X_{i,1997-i}$.

To extract approximately the portions that refer only to accident years 1983 to 1986, we assume the ratios of $\sum_{i=0}^{13} X_{1982+i,2} \Big/ \sum_{i=0}^{13} X_{1983+i,1}$, $\sum_{i=0}^{13} X_{1981+i,3} \Big/ \sum_{i=0}^{13} X_{1982+i,2}$, and so on are stable, particularly for the last few development years, in which the ratio is about 0.7 on average. We then make use of this ratio to spread the last set of data across development years 11 to 14. (In addition, the ratio used similarly for both motor insurance and CTP insurance is 0.6.)

The following table lists the AWOTE rates extracted from the Australian Bureau of Statistics (ABS) publications. These rates are used to convert the claim payments to 31 December 1996 dollar values.

| Period | AWOTE | Period | AWOTE |
|---|---|---|---|
| May-Jul 1983 | 339.80 | May-Jul 1994 | 620.00 |
| May-Jul 1984 | 369.40 | May-Jul 1995 | 653.10 |
| May-Jul 1985 | 388.80 | May-Jul 1996 | 674.60 |
| May-Jul 1986 | 419.80 | May-Jul 1997 | 704.30 |
| May-Jul 1987 | 446.00 | May-Jul 1998 | 735.40 |
| May-Jul 1988 | 470.10 | May-Jul 1999 | 753.00 |
| May-Jul 1989 | 509.70 | May-Jul 2000 | 796.10 |
| May-Jul 1990 | 541.70 | May-Jul 2001 | 838.50 |
| May-Jul 1991 | 567.50 | May-Jul 2002 | 879.40 |
| May-Jul 1992 | 585.70 | May-Jul 2003 | 929.60 |
| May-Jul 1993 | 600.80 | May-Jul 2004 | 962.90 |

Referring to Section 7, for the total claim payments of a calendar year, we compute its average proportion that belongs to a particular development year, based on the claims data of payment years 1983 to 1996. We then apply these proportions to the claim payments of payment years 1997 to 2002 in order to extract the components that arise only from accident year 1996 or before. These components are approximately the actual claim payments made within payment years 1997 to 2002 regarding the outstanding claims liability as at 31 December 1996.

Furthermore, based on the chain ladder method and the corresponding estimates, we modify the actual claim payments above to compute an approximate total amount of the actual claim payments made after 31 December 1996 regarding the outstanding claims liability as at 31 December 1996.

# Appendix II

As an example, the initial gamma model structure for public liability insurance in Section 5 can be implemented with the following BUGS coding (refer to the BUGS manual for details of programming).

**Gamma** (for public liability insurance)

```
model;
{
miu ~ dnorm(0 , 0.001)
alpha[1] <- 0
for( i in 2 : 14 ){
alpha[i] ~ dnorm(0 , 0.01)
}
beta[1] <- 0
for( i in 2 : 14 ){
beta[i] ~ dnorm(0 , 0.01)
}
for( i in 1 : 14 ){
for( j in 1 : 14 ){
u[i , j] <- exp( miu + alpha[i] + beta[j] )
a[i , j] <- r / u[i , j]
x[i , j] ~ dgamma( r , a[i , j] )
```

```
}
}
r ~ dunif(0 , 100)
for( i in 2 : 14 ){
os[i] <- sum( x[i , (16 - i) : 14] )
}
ostotal <- sum( os[2 : 14] )
}
```

For motor insurance in Section 4, some development-year parameter estimates of the initial approximate ODP model structure are awkward compared to those generated from a GLM. This result may be explained, as reflected in the run-off triangle, by the fact that most of the claim payments are settled in the first few development years and so the claim payments in the later development years are relatively too small to lead to meaningful sampling with BUGS. The approximation of the ODP structure may also contribute to such results. Accordingly, we adopt the following coding, in which the fixed parameters of some prior distributions are adjusted based on the corresponding GLM parameter estimates.

**ODP** (for motor insurance)

```
model;
{
```

```
miu ~ dnorm(0 , 0.001)

alpha[1] <- 0

for( i in 2 : 13 ){

alpha[i] ~ dnorm(0 , 0.01)

}

beta[1] <- 0

for( i in 2 : 4 ){

beta[i] ~ dnorm(0 , 0.01)

}

for( i in 5 : 6 ){

beta[i] ~ dunif(-7 , 0)

}

beta[7] ~ dunif(-8 , 0)

for( i in 8 : 13 ){

beta[i] ~ dunif(-9 , 0)

}

for( i in 1 : 13 ){

for( j in 1 : 13 ){

u[i , j] <- exp( miu + alpha[i] + beta[j] )

z[i , j] <- 1 * ipheta / u[i , j]

x[i , j] ~ dnorm( u[i , j] , z[i , j] )

}

}
```

```
ipheta ~ dgamma(0.01 , 0.01)

for( i in 2 : 13 ){

os[i] <- sum( x[i , (15 - i) : 13] )

}

ostotal <- sum( os[2 : 13] )

}
```

For motor insurance, the model structures adjusted with polynomials and weights also require similar modifications on some of the prior distributions.

# Appendix III

The following table shows the BUGS and GLM parameter estimates of using the initial ODP and gamma model structures for the three lines of business under consideration.

| **Parameter** | **Motor - ODP** | | **Public Liability - ODP** | | **Public Liability - Gamma** | | **CTP - ODP** | | **CTP - Gamma** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Estimate** | **BUGS** | **GLM** | **BUGS** | **GLM** | **BUGS** | **GLM** | **BUGS** | **GLM** | **BUGS** | **GLM** |
| $v$ | 20.7700 | 20.7950 | 17.1300 | 17.0485 | 17.0300 | 17.0140 | 14.8100 | 14.8811 | 14.8100 | 14.8047 |
| $\alpha_2$ | 0.1946 | 0.1878 | -0.4468 | -0.2758 | -0.2179 | -0.2150 | 0.2114 | 0.1949 | 0.2026 | 0.2046 |
| $\alpha_3$ | 0.2826 | 0.2709 | -0.2189 | -0.0706 | -0.0029 | -0.0003 | 0.1510 | 0.1306 | 0.1691 | 0.1702 |
| $\alpha_4$ | 0.4442 | 0.4327 | -0.2502 | -0.0742 | -0.0419 | -0.0407 | 0.1773 | 0.1600 | 0.1984 | 0.1982 |
| $\alpha_5$ | 0.4386 | 0.4229 | -0.1889 | -0.0335 | 0.0238 | 0.0265 | 0.2232 | 0.2098 | 0.2521 | 0.2518 |
| $\alpha_6$ | 0.4948 | 0.4753 | -0.1581 | -0.0057 | 0.0380 | 0.0386 | 0.1628 | 0.1398 | 0.1840 | 0.1851 |
| $\alpha_7$ | 0.6853 | 0.6657 | -0.0924 | 0.0581 | 0.0829 | 0.0836 | 1.0590 | 1.0213 | 1.0200 | 1.0208 |
| $\alpha_8$ | 0.6045 | 0.5788 | 0.1083 | 0.2108 | 0.2416 | 0.2408 | 2.1000 | 2.0229 | 2.0730 | 2.0736 |
| $\alpha_9$ | 0.5740 | 0.5481 | -0.0696 | 0.0717 | 0.1567 | 0.1526 | 2.2280 | 2.1661 | 2.2150 | 2.2142 |
| $\alpha_{10}$ | 0.6530 | 0.6247 | 0.1152 | 0.2182 | 0.2418 | 0.2381 | 2.3490 | 2.2983 | 2.3360 | 2.3325 |
| $\alpha_{11}$ | 0.8199 | 0.7924 | 0.1615 | 0.3124 | 0.3617 | 0.3508 | 2.6110 | 2.5620 | 2.5610 | 2.5582 |
| $\alpha_{12}$ | 0.8742 | 0.8496 | 0.3315 | 0.4599 | 0.5178 | 0.5060 | 2.9140 | 2.8715 | 2.8460 | 2.8390 |
| $\alpha_{13}$ | 1.0760 | 1.0484 | 0.0847 | 0.2513 | 0.3154 | 0.2847 | 2.9600 | 2.9132 | 2.8490 | 2.8387 |
| $\alpha_{14}$ | | | 0.2165 | 0.3696 | 0.4675 | 0.4041 | 2.7420 | 2.6963 | 2.8030 | 2.7727 |
| $\beta_2$ | -1.3670 | -1.3745 | 0.5634 | 0.5305 | 0.5179 | 0.5159 | 0.7491 | 0.7303 | 0.9519 | 0.9534 |
| $\beta_3$ | -3.2840 | -3.5474 | 0.3458 | 0.3328 | 0.3166 | 0.3152 | 1.5340 | 1.5108 | 1.5450 | 1.5467 |
| $\beta_4$ | -5.9550 | -5.5982 | 0.3423 | 0.3450 | 0.3372 | 0.3347 | 1.8130 | 1.7972 | 1.8070 | 1.8085 |
| $\beta_5$ | -6.7320 | -6.4599 | 0.2146 | 0.2286 | 0.2214 | 0.2178 | 1.8340 | 1.8125 | 1.8220 | 1.8232 |
| $\beta_6$ | -6.7640 | -6.9206 | 0.5733 | 0.4256 | 0.4233 | 0.4229 | 1.6340 | 1.6332 | 1.7240 | 1.7236 |
| $\beta_7$ | -7.7050 | -7.4250 | -0.1438 | -0.0910 | -0.0803 | -0.0843 | 1.3960 | 1.3957 | 1.4560 | 1.4566 |
| $\beta_8$ | -8.6710 | -8.5583 | -0.4724 | -0.3730 | -0.3701 | -0.3753 | 1.2800 | 1.2808 | 1.3030 | 1.3020 |
| $\beta_9$ | -8.6090 | -8.5141 | -0.7317 | -0.5952 | -0.5914 | -0.5975 | 0.9612 | 0.9651 | 1.0050 | 1.0051 |
| $\beta_{10}$ | -8.5250 | -8.7123 | -1.0690 | -0.8852 | -0.8943 | -0.9043 | 0.4669 | 0.5426 | 0.5957 | 0.5946 |
| $\beta_{11}$ | -8.3650 | -8.7490 | 0.1685 | -0.0358 | -0.0383 | -0.0507 | 0.8419 | 0.8229 | 0.8833 | 0.8790 |
| $\beta_{12}$ | -8.1150 | -8.8214 | 0.0129 | -0.1902 | -0.1020 | -0.1184 | 0.3339 | 0.3490 | 0.4207 | 0.4123 |
| $\beta_{13}$ | -7.5830 | -8.6160 | -0.4620 | -0.3943 | -0.4257 | -0.4493 | -0.2558 | -0.1256 | -0.0862 | -0.0995 |
| $\beta_{14}$ | | | -2.5780 | -2.0074 | -1.9100 | -1.9729 | -0.2722 | -0.0690 | 0.0398 | 0.0074 |

As shown above, the BUGS and GLM parameter estimates are comparable in magnitude.