

Stochastic Loss Reserving with Dependence: A Flexible Multivariate Tweedie Approach

Benjamin Avanzi^{a,b}, Greg Taylor^a, Phuong Anh Vu^{a,b,*}, Bernard Wong^a

^a*School of Risk and Actuarial Studies, UNSW Australia Business School
UNSW Sydney NSW 2052, Australia*

^b*Département de Mathématiques et de Statistique, Université de Montréal
Montréal QC H3T 1J4, Canada*

Abstract

Stochastic loss reserving with dependence has received increased attention in the last decade. A number of parametric multivariate approaches have been developed to capture dependence between lines of business within an insurer's portfolio. Motivated by the richness of the Tweedie family of distributions, we propose a multivariate Tweedie approach to capture cell-wise dependence in loss reserving. This approach provides a transparent introduction of dependence through a common shock structure. In addition, it also has a number of ideal properties, including marginal flexibility, transparency, and tractability including moments that can be obtained in closed form. Theoretical results are illustrated using both simulated data sets and a real data set from a property-casualty insurer in the US.

Keywords: Stochastic Loss Reserving, Dependence, Multivariate Tweedie Distribution, Common Shock, Bayesian Estimation

JEL code: G22

Accepted for publication in *Insurance: Mathematics and Economics*

1. Introduction

In a non-life insurance company, there is typically a delay between the occurrence of an insured event and the actual payment of its related claims. This delay can be driven by various reasons, including delays in reporting claims, investigation of claim validity, and legal proceedings. In order to sustain financial stability and meet regulatory requirements, it is essential for insurers to have sufficient reserves to respond to the outstanding claims. These loss reserves are one of the largest liabilities on the balance sheet of an insurer (see for example, Alai and Wüthrich, 2009; Shi, 2014; Zhang, Dukic, and Guszcza, 2012), which further emphasises the importance of having an adequate estimation of outstanding liabilities. This is further discussed and motivated in the excellent general reference by Wüthrich and Merz (2008).

When a company has more than one line of business, one approach for loss reserving is to simply add the reserves of each individual business line. However, this approach is only accurate in the perfectly positive dependence case, and as a result, an insurer is unable to enjoy diversification benefits (De Jong, 2012; Shi, Basu, and Meyers, 2012). This subsequently motivates the development of stochastic loss reserving with non-perfect dependence; see for example, Abdallah, Boucher, and Cossette (2015); Merz and Wüthrich (2009a); Merz, Wüthrich, and Hashorva (2013); Zhang, Dukic, and Guszcza (2012).

One of the main streams of the literature on stochastic loss reserving with dependence focuses on cell-wise dependence between business lines (Merz, Wüthrich, and Hashorva, 2013). This refers to the dependence

*Corresponding author.

Email addresses: b.avanzi@unsw.edu.au (Benjamin Avanzi), greg.taylor@unsw.edu.au (Greg Taylor), p.vu@unsw.edu.au (Phuong Anh Vu), bernard.wong@unsw.edu.au (Bernard Wong)

between claims coming from the same accident and development periods in different run-off triangles. Early modelling developments with this type of dependence include multivariate additive models by Hess, Schmidt, and Zocher (2006); Merz and Wüthrich (2009b) and Schmidt (2006). Copulas have also been used in models with cell-wise dependence, for example, Shi and Frees (2011) and Zhang and Dukic (2013). A later stream of the literature has also incorporated an additional source of dependence arising from calendar year effects, see, for example, Bühlmann and Moriconi (2015); De Jong (2012); Salzmänn and Wüthrich (2012) and Shi (2014).

Methodologies used in loss reserving models with dependence can be classified into two main groups, (i) parametric models, and (ii) non-parametric models (Shi, Basu, and Meyers, 2012). Parametric models utilise distributional families, while the later ones do not. The most popular parametric approach in the literature is to use copulas. Shi and Frees (2011) proposed a flexible copula model with marginal generalised linear models (GLM). De Jong (2012) used a Gaussian copula approach to capture the dependence between lines. Zhang and Dukic (2013) developed a Bayesian copula framework with flexible marginal modelling for cell-wise dependence. Abdallah, Boucher, and Cossette (2015) used a hierarchical Archimedean copula structure to capture calendar year dependence between lines. The copula approach has the great benefit of having flexibility because it allows marginal densities and joint dependence to be modelled separately. Besides copulas, there have also been developments utilising a multivariate distribution for a specific marginal density, particularly, log-normal distributions. Examples include the multivariate log-normal frameworks for incremental claims in Shi, Basu, and Meyers (2012) and for log-link ratios in Merz, Wüthrich, and Hashorva (2013).

The search for multivariate modelling approaches drew our attention to a multivariate Tweedie distribution for margins from the Tweedie family of distributions. The Tweedie family is a major subclass of the exponential dispersion family (EDF) consisting of symmetric and non-symmetric, light-tailed and heavy-tailed distributions (Alai, Landsman, and Sherris, 2015; Jørgensen, 1997). This class and its members are frequently used in loss reserving, see for example, Alai and Wüthrich (2009); Boucher and Davidov (2011); England and Verrall (2002); Peters, Shevchenko, and Wüthrich (2009); Renshaw and Verrall (1998); Taylor (2009, 2015); Wüthrich (2003); Zhang, Dukic, and Guszczka (2012). Furthermore, it is a generalisation of the plain vanilla Chain Ladder Poisson model. A recapitulation of some properties of the univariate Tweedie family of distributions is provided in Appendix 8.

In this work, we focus on a multivariate Tweedie distribution developed by Furman and Landsman (2010). The Tweedie family is a broad class of commonly used distributions, and the multivariate distribution is developed through a common shock approach. These allow a multivariate Tweedie framework to have many advantages. Namely, the advantages of our model include:

- Dependence is introduced with the help of a easily identifiable, explicit common shock dependence structure, which is easily generalised to more than two dimensions.
- While all dimensions must have same parameter p , this can be anywhere in $(-\infty, 0] \cap [1, \infty)$ (rather than fixed at 1 for a Poisson dispersion or 2 for a gamma dispersion, for instance). The flexibility of Tweedie dispersions has been established in Alai, Landsman, and Sherris (2015); Alai and Wüthrich (2009); Furman and Landsman (2010); Jørgensen (1997). Additionally, zero data points do not present any issue.
- The distributions of (multivariate) margins belong to the same Tweedie family. Furthermore, moments and cumulants that can be obtained analytically, and cumulants of the sum (of reserves) can be given in closed form.

The paper is organised as follows. In Section 2, we introduce a multivariate Tweedie framework. Appropriate parametrisation is considered and appealing properties of this framework are also discussed. An analysis of moments and cumulants from the framework is performed in Section 3. Bayesian inference is developed in Section 4 for model fitting and forecasting. The choices of prior distributions, estimation procedure and Markov Chain Monte Carlo (MCMC) method are also discussed. Section 5 applies the theoretical framework to simulated data sets to assess the accuracy of the estimation procedure. An illustration on real data is then provided in Section 6. Section 7 concludes the paper with some remarks about the model and its applications.

2. A multivariate Tweedie dependence approach

2.1. Notation

We consider a portfolio of N sub-portfolios that represent N lines of business. Notation $X_{i,j}^{(n)}$ represents the total incremental claim that corresponds to the accident period i , $i \in \{1, \dots, I\}$, and development period j , $j \in \{0, \dots, J\}$, in the n -th business line. These claims are hence made in calendar period $t = i + j$, $t \in \{1, \dots, I\}$.

Instead of directly modelling the losses, one often standardises incremental claims using exposure variables to obtain consistency across different accident years and lines of business. Common exposure variables, denoted by $\omega_{i,j}^{(n)}$, can, for example, be the number of policies, the amount of premiums written, or the total amount insured. Standardised incremental claims are denoted by

$$Y_{i,j}^{(n)} = \frac{X_{i,j}^{(n)}}{\omega_{i,j}^{(n)}}. \quad (2.1)$$

The set of all claims observations up to the estimation date is represented by

$$Y^U = \left\{ Y_{i,j}^{(n)}; 1 \leq i \leq I, 0 \leq j \leq I - i, 1 \leq n \leq N \right\}, \quad (2.2)$$

and the set of all outstanding claims that are to be predicted is denoted by

$$Y^L = \left\{ Y_{i,j}^{(n)}; 1 < i \leq I, I - i + 1 \leq j \leq J, 1 \leq n \leq N \right\}. \quad (2.3)$$

The primary goal of a loss reserving model is to use historical claims information Y^U to predict the amount of outstanding claims Y^L .

2.2. Model construction

In this section we will develop a multivariate Tweedie framework for claims from multiple lines of business. Following the literature stream that models cell-wise dependence between lines of business, a multivariate Tweedie distribution is used to capture dependence between cell-wise claims. Standardised cell-wise claims from the i -th accident period and j -th development period across all lines of business are first collected into a vector

$$\mathbf{Y}_{i,j} = \begin{pmatrix} Y_{i,j}^{(1)} \\ Y_{i,j}^{(2)} \\ \vdots \\ Y_{i,j}^{(N)} \end{pmatrix}. \quad (2.4)$$

Each element of the above vector is assumed to be a sum of two components

$$Y_{i,j}^{(n)} = \frac{\theta}{\theta_{i,j}^{(n)}} W_{i,j} + Z_{i,j}^{(n)}, \quad (2.5)$$

where $W_{i,j}$ is referred to as the ‘‘common shock’’ and $Z_{i,j}^{(n)}$ is referred to as the ‘‘idiosyncratic effect’’. These two components are assumed to be independent and have additive Tweedie distributions

$$W_{i,j} \sim \text{Tweedie}_p^*(\theta, \lambda), \quad (2.6)$$

$$Z_{i,j}^{(n)} \sim \text{Tweedie}_p^*(\theta_{i,j}^{(n)}, \lambda_{i,j}^{(n)}). \quad (2.7)$$

The definition of additive Tweedie distributions as well as other properties of the Tweedie family of distributions are provided in Appendix 8.

The structure in (2.5) is a typical form of the multivariate reduction construction (Furman, 2008; Karlis, 2003). This construction provides a transparent dependence structure. In particular, all cell-wise claims that are in the same position (i, j) share a common stochastic component $W_{i,j}$. Dependence between all cell-wise claims is then induced from this common shock. As we can also observe from this construction, the effect of the common shock $W_{i,j}$ to each loss cell is scaled by a factor of $\theta/\theta_{i,j}^{(n)}$. This scaling factor aims to adjust effects of the common shock to different lines of business. Unique cell effects are captured in the idiosyncratic component $Z_{i,j}^{(n)}$.

Following Furman and Landsman (2010), the marginal distribution of $Y_{i,j}^{(n)}$ is

$$Y_{i,j}^{(n)} \sim \begin{cases} \text{Tweedie}_1^*(\theta, \lambda + \lambda_{i,j}^{(n)}), & p = 1, \\ \text{Tweedie}_p^*\left(\theta_{i,j}^{(n)}, \lambda \left(\frac{\theta}{\theta_{i,j}^{(n)}}\right)^\tau + \lambda_{i,j}^{(n)}\right), & p \neq 1. \end{cases} \quad (2.8)$$

where $\tau = (p-2)/(p-1)$. The marginal densities in this multivariate model are Tweedie distributions, this model satisfies the marginal closure property from Joe (1997, Chapter 4). In addition, this model also provides a general framework to a variety of familiar models, such as the multivariate Poisson model ($p = 1$), the multivariate gamma model ($p = 2$), and the multivariate inverse Gaussian model ($p = 3$). Importantly, it is also a generalisation of the multivariate Tweedie's compound Poisson framework with $1 < p < 2$, which is able to deal with data sets with zeros in claims observations.

Remark 2.1. *This multivariate Tweedie framework provides a framework for modelling outstanding claims with a flexible selection of power parameter p . This is an improvement over existing multivariate models such as multivariate Poisson and multivariate gamma frameworks with restricted p values. It is also worth noting that a common power parameter p is required for all lines of business in implementing the framework. This allows the multivariate Tweedie framework to have the closure property under the taking of margins mentioned in Joe (1997, Chapter 4). However, this can become a limitation when dealing with a large number of business lines.*

The multivariate density of vector $\mathbf{Y}_{i,j}$ can then be expressed using the convolutions of pairs of independent random variables $W_{i,j}$ and $Z_{i,j}^{(n)}$. The random variable $W_{i,j}$ plays the role of the ‘‘common shock’’, and is common across all these convolutions.

Remark 2.2. *In the construction of the model, the most simple parametrisation is used for the common shock component $W_{i,j}$ with canonical parameter θ and index parameter λ . This is chosen for the sake of simplicity. Note, however, that this parametrisation can be modified for different scenarios. Indeed, these parameters (of the common shock) could vary, for example, across development years, accident years and/or calendar years.*

This simplified assumption can lead to an unsatisfactory performance of the model for unbalanced data that involves different lines of business with different development durations. This limitation can be overcome by specifying column-specific or volume-related parameters for the common shock $W_{i,j}$.

Remark 2.3. *In loss reserving, negative payments can sometimes occur as a result of salvage recoveries, payments from third parties or rejections by the insurer (De Alba, 2006). There exist methods for incremental claims with negative payments in the literature, such as claims translation using three-parameter log-normal distributions in De Alba (2006), and the mixture methods in Kunkler (2006) and Xia and Scollnik (2015). The claims translation method can be used together with the multivariate Tweedie framework to handle negative claims. Namely, all claims in each loss triangle can be shifted by a parameter so that all translated claims are non-negative. The multivariate Tweedie framework can then be applied to loss triangles of translated claims and the parameters used to translate claims are estimated in the model estimation.*

2.3. Model parametrisation

In this section, we provide further guidance on model parametrisation. One of the properties of the Tweedie family and the EDF family in general, as mentioned in Appendix 8, is the availability of two

representations, which are known as the additive form, and the reproductive form, respectively. The general framework is constructed using the additive representation. However, the reproductive representation has a location parameter and a dispersion parameter that specifies the mean and the dispersion of the distribution. This form is more meaningful and convenient to work with in the loss reserving context. Therefore, this representation is used for further parametrisation. In the special case of Poisson ($p = 1$), we use an equivalent Poisson representation of the model. We also provide an extension to the Poisson case to also allow for over-dispersion. A summary of model parametrisation for these two cases is given in Table 2.1.

Case 1: $p \neq 1$

We perform the model parametrisation on the reproductive Tweedie representation for cases with $p \neq 1$. In the equivalent marginal reproductive representation we have

$$W_{i,j} \sim \text{Tweedie}_p(\alpha, \beta), \quad (2.9)$$

$$Z_{i,j}^{(n)} \sim \text{Tweedie}_p(\eta_i^{(n)} \nu_j^{(n)}, \gamma_{i,j}^{(n)}). \quad (2.10)$$

The duality between the additive form and the reproductive form of a Tweedie distribution (see Appendix 8.10) can be used to express the relationships between parameters of the additive and reproductive representations. This is provided in Table 2.1. After performing an appropriate claims standardisation, we further assume that $\gamma_{i,j}^{(n)} = \gamma^{(n)}$ to simplify the analysis. This accounts for variable levels of dispersions between different lines of business in a similar way for all cells. This simplification corresponds to the simplification $\lambda_{i,j}^{(n)} = \lambda^{(n)}$ in the additive form. The parameters $\eta_i^{(n)}, \nu_j^{(n)}$ contain one degree of redundancy for each n . The redundancy is removed by setting $\eta_1^{(n)} = 1, n = 1, \dots, N$.

Using the equivalent reproductive representation, the model becomes

$$Y_{i,j}^{(n)} = \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} W_{i,j} + Z_{i,j}^{(n)}. \quad (2.11)$$

In this expression, the effect of the common shock $W_{i,j}$ is scaled by a product of ratios between dispersion parameters β and $\gamma^{(n)}$, and between mean parameters α and $\eta_i^{(n)} \nu_j^{(n)}$, of the common shock $W_{i,j}$ and the idiosyncratic effect $Z_{i,j}^{(n)}$ respectively.

Using the relationship between the additive and reproductive representations of the Tweedie family in Jørgensen (1997, Chapter 4), the marginal distribution of $Y_{i,j}^{(n)}$ is given by

$$Y_{i,j}^{(n)} \sim \text{Tweedie}_p \left(\eta_i^{(n)} \nu_j^{(n)} \left[\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{2-p} \frac{\gamma^{(n)}}{\beta} + 1 \right], \gamma^{(n)} \left[\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{2-p} \frac{\gamma^{(n)}}{\beta} + 1 \right]^{1-p} \right). \quad (2.12)$$

The multivariate density using reproductive representation is

$$f_{\mathbf{Y}_{i,j}} \left(y_{i,j}^{(1)}, \dots, y_{i,j}^{(N)} \right) = \int_0^{A_{i,j}} f_{W_{i,j}}(w_{i,j}) \prod_{n=1}^N f_{Z_{i,j}^{(n)}} \left(y_{i,j}^{(n)} - \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} w_{i,j} \right) dw_{i,j}, \quad (2.13)$$

where

$$A_{i,j} = \min \left(\left(\frac{\eta_i^{(1)} \nu_j^{(1)}}{\alpha} \right)^{1-p} \frac{\beta}{\gamma^{(1)}} y_{i,j}^{(1)}, \dots, \left(\frac{\eta_i^{(N)} \nu_j^{(N)}}{\alpha} \right)^{1-p} \frac{\beta}{\gamma^{(N)}} y_{i,j}^{(N)} \right), \quad (2.14)$$

and where $f(\cdot)$ is the Tweedie density in reproductive form.

Case 2: $p = 1$ (Over-dispersed Poisson)

In the standard Poisson notation (Jørgensen, 1997, Chapter 4), when $p = 1$, the model construction in (2.11) reduces to

$$W_{i,j} \sim \text{Over-dispersed Poisson}(\alpha, \beta), \quad (2.15)$$

$$Z_{i,j}^{(n)} \sim \text{Over-dispersed Poisson}\left(\eta_i^{(n)} \nu_j^{(n)}, \gamma_{i,j}^{(n)}\right), \quad (2.16)$$

or equivalently,

$$\frac{W_{i,j}}{\beta} \sim \text{Poisson}\left(\frac{\alpha}{\beta}\right), \quad (2.17)$$

$$\frac{Z_{i,j}^{(n)}}{\gamma_{i,j}^{(n)}} \sim \text{Poisson}\left(\frac{\eta_i^{(n)} \nu_j^{(n)}}{\gamma_{i,j}^{(n)}}\right). \quad (2.18)$$

The restriction $\eta_1^{(n)} = 1$, $n = 1, \dots, N$ is imposed to remove parameter redundancy. Parameters β and $\gamma_{i,j}^{(n)}$ are referred to as dispersion parameters. Similar to the general case, we further assume that $\gamma_{i,j}^{(n)} = \gamma^{(n)}$ to simplify the analysis.

Incremental claims are then assumed to be

$$Y_{i,j}^{(n)} = \frac{\gamma^{(n)}}{\beta} W_{i,j} + Z_{i,j}^{(n)}. \quad (2.19)$$

The effect of the common shock $W_{i,j}$ to each loss cell is scaled by a factor which is a ratio between dispersion parameters $\gamma^{(n)}/\beta$. Consequently the marginal density of each loss cell is

$$Y_{i,j}^{(n)} \sim \text{Over-dispersed Poisson}\left(\frac{\gamma^{(n)}}{\beta} \alpha + \eta_i^{(n)} \nu_j^{(n)}, \gamma^{(n)}\right). \quad (2.20)$$

The multivariate mass function of $\mathbf{Y}_{i,j}$ is then given by

$$p_{\mathbf{Y}_{i,j}}\left(y_{i,j}^{(1)}, \dots, y_{i,j}^{(N)}\right) = \sum_{w_{i,j}=0}^{A_{i,j}} p_{W_{i,j}}(w_{i,j}) \prod_{n=1}^N p_{Z_{i,j}^{(n)}}\left(y_{i,j}^{(n)} - \frac{\gamma^{(n)}}{\beta} w_{i,j}\right), \quad (2.21)$$

where

$$A_{i,j} = \min\left(\frac{\beta}{\gamma^{(1)}} y_{i,j}^{(1)}, \dots, \frac{\beta}{\gamma^{(N)}} y_{i,j}^{(N)}\right), \quad (2.22)$$

and where $p(\cdot)$ is the probability mass function.

The general model framework in Section 2.2 produces the original multivariate Poisson framework in Kocherlakota and Kocherlakota (1992). In this original multivariate Poisson framework, unit dispersions are used, i.e. $\beta = \gamma^{(n)} = 1$. The framework introduced previously provides an extension to the original multivariate Poisson framework to also allow for over-dispersion.

p	Additive form	Reproductive form/Poisson notation
$p \neq 1$	$Y_{i,j}^{(n)} = \frac{\theta}{\theta_{i,j}^{(n)}} W_{i,j} + Z_{i,j}^{(n)}$ $Y_{i,j}^{(n)} \sim \text{Tweedie}_p^* \left(\theta_{i,j}^{(n)}, \lambda \left(\frac{\theta}{\theta_{i,j}^{(n)}} \right)^\tau + \lambda_{i,j}^{(n)} \right)$ $W_{i,j} \sim \text{Tweedie}_p^*(\theta, \lambda)$ $Z_{i,j}^{(n)} \sim \text{Tweedie}_p^*(\theta_{i,j}^{(n)}, \lambda_{i,j}^{(n)})$	$Y_{i,j}^{(n)} = \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} W_{i,j} + Z_{i,j}^{(n)}$ $Y_{i,j}^{(n)} \sim \text{Tweedie}_p \left(\eta_i^{(n)} \nu_j^{(n)} \left[\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{2-p} \frac{\gamma^{(n)}}{\beta} + 1 \right], \gamma^{(n)} \left[\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{2-p} \frac{\gamma^{(n)}}{\beta} + 1 \right]^{1-p} \right)$ $W_{i,j} \sim \text{Tweedie}_p(\alpha, \beta)$ $Z_{i,j}^{(n)} \sim \text{Tweedie}_p(\eta_i^{(n)} \nu_j^{(n)}, \gamma^{(n)})$
	$\lambda_{i,j}^{(n)} = \lambda^{(n)}, \quad \alpha = \lambda \left(\frac{\theta}{\tau - 1} \right)^{\tau-1}, \quad \alpha = \lambda \left(\frac{\theta}{\tau - 1} \right)^{\tau-1}, \quad \beta = \lambda^{1-p}, \quad \gamma^{(n)} = (\lambda^{(n)})^{1-p}$	<p>Relationships between parameters:</p>
$p = 1$	$Y_{i,j}^{(n)} = W_{i,j} + Z_{i,j}^{(n)}$ $Y_{i,j}^{(n)} \sim \text{Tweedie}_1^*(\theta, \lambda + \lambda_{i,j}^{(n)})$ $W_{i,j} \sim \text{Tweedie}_1^*(\theta, \lambda)$ $Z_{i,j}^{(n)} \sim \text{Tweedie}_1^*(\theta, \lambda_{i,j}^{(n)})$	$Y_{i,j}^{(n)} = W_{i,j} + Z_{i,j}^{(n)}$ $Y_{i,j}^{(n)} \sim \text{Poisson}(\alpha + \eta_i^{(n)} \nu_j^{(n)})$ $W_{i,j} \sim \text{Poisson}(\alpha)$ $Z_{i,j}^{(n)} \sim \text{Poisson}(\eta_i^{(n)} \nu_j^{(n)})$
		<p>Extension to over-dispersed Poisson:</p> $Y_{i,j}^{(n)} = \frac{\gamma^{(n)}}{\beta} W_{i,j} + Z_{i,j}^{(n)}$ $Y_{i,j}^{(n)} \sim \text{Over-dispersed Poisson} \left(\frac{\gamma^{(n)}}{\beta} \alpha + \eta_i^{(n)} \nu_j^{(n)}, \gamma^{(n)} \right)$ $W_{i,j} \sim \text{Over-dispersed Poisson}(\alpha, \beta)$ $Z_{i,j}^{(n)} \sim \text{Over-dispersed Poisson}(\eta_i^{(n)} \nu_j^{(n)}, \gamma^{(n)})$
		<p>Relationships between parameters:</p> <p>Poisson: $\alpha = \lambda \exp(\theta), \quad \eta_i^{(n)} \nu_j^{(n)} = \lambda_{i,j}^{(n)} \exp(\theta)$</p> <p>Over-dispersed Poisson: $\frac{\alpha}{\beta} = \lambda \exp(\theta), \quad \frac{\eta_i^{(n)} \nu_j^{(n)}}{\gamma^{(n)}} = \lambda_{i,j}^{(n)} \exp(\theta)$</p>

Table 2.1: Summary table of model parametrisation

3. Analysis of moments

This section provides an analysis of moments in the multivariate Tweedie framework. This framework has the advantage of allowing moments of each claim cell $Y_{i,j}^{(n)}$, and consequently, moments of the total sum of outstanding claims to be obtained in closed form.

The generating function of the (over-dispersed) Poisson case with $p = 1$ can be obtained using properties of the Poisson distribution. Moment-generating functions of the common shock $W_{i,j}$, the idiosyncratic component $Z_{i,j}^{(n)}$ and the marginal variable $Y_{i,j}^{(n)}$ for the non-Poisson case with $p \neq 1$ are provided in Furman and Landsman (2010). They are represented using parameters in the reproductive form and provided in Appendix 9. Using these moment-generating functions, moments of $Y_{i,j}^{(n)}$ can be expressed in a general form as

$$E \left[\left(Y_{i,j}^{(n)} \right)^m \right] = \sum_{r=0}^m \binom{m}{r} \left(\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right)^r E [W_{i,j}^r] E \left[\left(Z_{i,j}^{(n)} \right)^{m-r} \right], \quad (3.1)$$

and cumulants of $Y_{i,j}^{(n)}$ are

$$K_{Y_{i,j}^{(n)}}^{(m)} = \begin{cases} \left(\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right)^m \beta^{\frac{1}{1-p}} \kappa_p^{(m)} \left(\frac{\alpha^{1-p}}{\beta(1-p)} \right) + (\gamma^{(n)})^{\frac{1}{1-p}} \kappa_p^{(m)} \left(\frac{\left(\eta_i^{(n)} \nu_j^{(n)} \right)^{1-p}}{\gamma^{(n)}(1-p)} \right), & p \neq 1, \\ \frac{\alpha}{\beta} (\gamma^{(n)})^m + \eta_i^{(n)} \nu_j^{(n)} (\gamma^{(n)})^{m-1}, & p = 1, \end{cases} \quad (3.2)$$

where $\kappa(\cdot)$ is the unit cumulant function of the Tweedie family provided in (8.9).

Utilising these general formulas of moments and cumulants, closed form expressions of any moments and cumulants can be provided. A critical moment in loss reserving is the mean, which can be given by

$$E \left[Y_{i,j}^{(n)} \right] = \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \alpha + \eta_i^{(n)} \nu_j^{(n)}. \quad (3.3)$$

The mean of $Y_{i,j}^{(n)}$ in the case $p \neq 1$ is the location parameter of its marginal distribution. Alternatively, incremental claim cell $Y_{i,j}^{(n)}$ is also a sum of two independent components $W_{i,j}$ and $Z_{i,j}^{(n)}$. As a result, its mean is also computed as the sum of the (scaled) means of its individual components $W_{i,j}$ and $Z_{i,j}^{(n)}$.

The common shock structure provides a convenient interpretation to the effect of dependence on the mean of the marginal claim cells. In comparison to marginal Tweedie models, for example, Alai and Wüthrich (2009); Peters, Shevchenko, and Wüthrich (2009); Wüthrich (2003), and other multivariate models, for example, Abdallah, Boucher, and Cossette (2015); Shi (2014); Zhang, Dukic, and Guszczka (2012), this multivariate Tweedie framework decomposes the mean of $Y_{i,j}^{(n)}$ into two separate parts. As shown in (3.3), the first part comes from the common shock $W_{i,j}$ and the second part comes from the idiosyncratic component $Z_{i,j}^{(n)}$. The mean $\eta_i^{(n)} \nu_j^{(n)}$ of idiosyncratic component $Z_{i,j}^{(n)}$ incorporates the accident period effect $\eta_i^{(n)}$ and the development period effect $\nu_j^{(n)}$. This is the total expected claim in cell (i, j) in the n -th loss triangle assuming business lines are independent. A common shock effect $W_{i,j}$ introduces an expected additional claims level α . However, this effect on each of the business lines is scaled by a scaling factor as given in the model construction in (2.11).

The variance of $Y_{i,j}^{(n)}$ can also be given using the general cumulant formula in (3.2)

$$\text{Var} \left[Y_{i,j}^{(n)} \right] = \left(\left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right)^2 \beta \alpha^p + \gamma^{(n)} \left(\eta_i^{(n)} \nu_j^{(n)} \right)^p. \quad (3.4)$$

The variance can also be obtained using the mean parameter and the dispersion parameter of the marginal density of $Y_{i,j}^{(n)}$ in (2.12). The overdispersion impact of the common shocks is, again, clearly identifiable. Similar to the mean, the variance of $Y_{i,j}^{(n)}$ can also be computed as the sum of variances of two components $W_{i,j}$ and $Z_{i,j}^{(n)}$ as shown in (3.4). This is due to the independence between these components in the model construction as described in Section 2.2. This suggests that dependence due to a stochastic factor $W_{i,j}$ can increase the variance of the outstanding claims $Y_{i,j}^{(n)}$. As shown in (3.4), the variance of common shock $W_{i,j}$, after being scaled for each line of business, is added to the existing variance coming from the unique idiosyncratic component $Z_{i,j}^{(n)}$. One can relate this to practical situations in which different business lines have a positive dependence structure, and factors causing claims volatility to increase in one business line can likely increase claims volatility in other lines.

The covariance between corresponding cells across different lines of business can also be obtained by

$$Cov [Y_{i,j}^{(n)}, Y_{i,j}^{(m)}] = \frac{\alpha^{2-p}}{\left(\eta_i^{(n)} \nu_j^{(n)} \eta_i^{(m)} \nu_j^{(m)}\right)^{1-p}} \frac{\gamma^{(n)} \gamma^{(m)}}{\beta}, \quad m \neq n. \quad (3.5)$$

This comes directly from the construction of the model in which the common shock term deduces dependence across lines of business. In particular, the covariance is calculated as the product of the variance of the common shock $W_{i,j}$ and its scaling factors in each individual business line. The common shock $W_{i,j}$ is the dependence generator which is introduced in the model construction. This transparent introduction of dependence allows an explicit expression of the covariance to be obtained. Of course, this covariance is null when $\alpha = 0$.

The common shock structure in the multivariate Tweedie framework also allows us to obtain closed form expressions for cumulants of the sum of claims. Because all cell-wise claims $Y_{i,j}^{(n)}$ in the same position (i, j) share a common shock component $W_{i,j}$, the sum of these claims can be given by

$$S_{i,j} = \left(\sum_{n=1}^N \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right) W_{i,j} + \sum_{n=1}^N Z_{i,j}^{(n)}. \quad (3.6)$$

Due to the independence between the common shock $W_{i,j}$ and the idiosyncratic components $Z_{i,j}^{(n)}$, any cumulant of $S_{i,j}$ (if it is defined), can be calculated as the sum of corresponding cumulants of $W_{i,j}$ and $Z_{i,j}^{(n)}$ with appropriate scaling factors. Two special cumulants, the mean and variance, of the sum $S_{i,j}$ (if they are defined) can be obtained in closed form as

$$E [S_{i,j}] = \left(\sum_{n=1}^N \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right) \alpha + \sum_{n=1}^N \eta_i^{(n)} \nu_j^{(n)}, \quad (3.7)$$

$$Var [S_{i,j}] = \left(\sum_{n=1}^N \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \right)^2 \beta \alpha^p + \sum_{n=1}^N \gamma^{(n)} \left(\eta_i^{(n)} \nu_j^{(n)} \right)^p. \quad (3.8)$$

The sum of all claims in all loss triangles of a company can be calculated by aggregating all sums of cell-wise claims $S_{i,j}$ for all i and j . Using the independence between claims from different positions within a loss triangle, any cumulant of the sum of all claims in the portfolio can be obtained as the sum of corresponding cumulants of $S_{i,j}$ for all i and j .

4. Model estimation

In the current literature, a number of estimation approaches have been considered for some special cases of the multivariate Tweedie distribution. For example, Karlis (2003) developed an Expectation-Maximisation

(EM) approach for the multivariate Poisson model, Tsonas (2004) developed a Bayesian framework with a Gibbs algorithm for the multivariate gamma model. The only estimation method that has been used for the general multivariate Tweedie distribution in the literature is the method of moments, see for example, Alai, Landsman, and Sherris (2015) and Furman and Landsman (2010). However, due to the small sample size often encountered in loss reserving, this method is not suitable to estimate parameters in the multivariate Tweedie framework.

Besides bootstrapping, Bayesian analysis is an appropriate choice for a small data set in the loss reserving context due to its ability to account for parameter uncertainty when providing a prediction for outstanding losses (Meyers and Shi, 2011). In the stochastic loss reserving literature, there has also been an increasing interest in Bayesian methods; see, for example, De Alba (2002); England and Verrall (2006); Meyers (2009); Shi, Basu, and Meyers (2012); Zhang and Dukic (2013). This is due to rapid developments in Bayesian methodologies as a result of high-speed computers and Monte Carlo methods (Verrall, Hössjer, and Björkwall, 2012). Profound developments in Monte Carlo methods, see for example, Brooks, Gelman, Jones, and Meng (2011); Kroese, Taimre, and Botev (2011), allow the Bayesian inference to be more appropriate for complex models. In addition, a Bayesian setup allows one to incorporate expert opinions using prior distributions (Shi, Basu, and Meyers, 2012). The so-called Bayesian forecasting also incorporates all currently available information into the prediction of outstanding claims (Salzmann, Wütrich, and Merz, 2012; Zhang, Dukic, and Guszczka, 2012).

In this section, we develop a Bayesian framework for the multivariate Tweedie model. This Bayesian inference is used to estimate all unknown parameters in the multivariate Tweedie framework except for the power parameter p . We propose the use of a Tweedie log-likelihood profile approach to estimate the power parameter p and using this estimate in the Bayesian inference. This can significantly improve the stability of the algorithm, especially when the parameter vector is large. The conventional Bayesian inference with only one single step of estimation using the multivariate Tweedie densities likelihood has significant instability due to a large number of parameters to be estimated at once. In addition, the multivariate Tweedie density for each claim cell position (i, j) , as given in (2.13), involves an integration over the common shock component that can take time to compute. This results in a rather inefficient Bayesian inference for estimation.

To improve the efficiency of estimation, we use a two stage estimation procedure for our model. This is a non-conventional Bayesian procedure. The idea comes from the closure under the taking of margins property of the multivariate Tweedie density, as mentioned in Remark 2.1. Even though we have pair-wise claim cells following a multivariate Tweedie density, each claim cell $Y_{i,j}^{(n)}$ still has its own marginal Tweedie density with specified location and dispersion parameters given in (2.12). In the marginal estimation stage, we utilise this property and assume that the likelihood is the product of all marginal densities of $Y_{i,j}^{(n)}$ for all i, j and n . In the marginal specification of all $Y_{i,j}^{(n)}$, even though all cells share the same parameters α and β , these two parameters always appear in a ratio, which is denoted by δ such that

$$\delta = \frac{\alpha^{2-p}}{\beta}. \quad (4.1)$$

As a result, there is parameter redundancy if these parameters are included as separate estimands in the marginal estimation stage. Therefore, instead of estimating both α and β , only the newly defined parameter δ is estimated in the first stage. The second stage is the multivariate estimation stage, which is when we estimate α and β separately using parameters estimated in the first stage.

In the construction of a Bayesian framework, one needs to specify prior distributions and the likelihood function of the framework in use. Details of these are provided in Section 4.1 and Section 4.2. A MCMC algorithm is then specified for Bayesian inference. A Metropolis-Hastings algorithm is chosen due to its ability to work with posterior densities that are in unrecognisable forms. Details of this algorithm are provided in Section 4.3. Section 4.4 then provides a description of the procedure used to obtain the predictive distribution of outstanding claims as well as other quantities of interest. A variety of Bayesian computational packages can be readily applied for special cases of the multivariate Tweedie framework including the multivariate (over-dispersed) Poisson framework (i.e. $p = 1$) and the multivariate gamma framework (i.e. $p = 2$). These tools are developed in a number of computational programs, for example, WinBUGS and

R. However, there are no existing computational tools that are ready for use for the general multivariate Tweedie framework. Hence the focus is placed on estimation of the general multivariate Tweedie framework with $p \neq 1$. Of course, the framework developed for $p \neq 1$ can be adapted to the case $p = 1$.

4.1. Prior distributions

Prior distributions can be chosen informatively where prior knowledge can be utilised to specify prior distributions (Koop, 2003). They can also be chosen to be uninformative prior distributions that assign equal possibilities to all values in the feasible set of parameter values. However, it is worth noting that the more informative prior distributions are, the faster the Monte Carlo chain converges (Congdon, 2010, Chapter 1). It is also mentioned in Brooks, Gelman, Jones, and Meng (2011, Chapter 23) that in highly parametrised models, somewhat informative prior distributions are necessary. Alternatively, expert knowledge can be used to specify informative prior distributions. A preliminary analysis can be performed on each line of business to get an idea on the possible ranges of the parameters.

Recall from the construction of the model that each claim cell $Y_{i,j}^{(n)}$ is the sum of a common shock component $W_{i,j}$ and an idiosyncratic component $Z_{i,j}^{(n)}$. Each claim cell $Y_{i,j}^{(n)}$ also follows a Tweedie marginal distribution. A naive analysis can be performed, assuming that $Y_{i,j}^{(n)}$ has a marginal Tweedie($H_i^{(n)}U_j^{(n)}, P^{(n)}$) distribution with location parameter $H_i^{(n)}U_j^{(n)}$ and dispersion parameter $P^{(n)}$. This corresponds to the univariate Tweedie model that has been used in univariate loss reserving in the literature; see, for example, Alai and Wüthrich (2009); Peters, Shevchenko, and Wüthrich (2009) and Wüthrich (2003). Maximum likelihood estimates for these parameters can be obtained using existing results in the literature. The estimates of $H_i^{(n)}$, $U_j^{(n)}$ and $P^{(n)}$ can give an idea of possible ranges for parameters $\eta_i^{(n)}$, $\nu_j^{(n)}$ in the location expression and dispersion parameter $\gamma^{(n)}$ of idiosyncratic component $Z_{i,j}^{(n)}$. The range of dispersion parameter β of the common shock $W_{i,j}$ can be chosen to be consistent with ranges of $\gamma^{(n)}$. We can analyse the level of dependence between lines of business to choose a reasonable range for location parameter α of the common shock $W_{i,j}$. Analyses are done heuristically for β and α , hence relatively uninformative prior distributions are recommended for these parameters.

4.2. Likelihood functions and posterior distributions

We estimate the parameters of the multivariate Tweedie model using a two step approach. The first step of the estimation process is the marginal estimation step. In this step, we can estimate all parameters, except for α and β . However, this estimation step allows us to obtain an estimate of δ , which is a function of α and β as defined in (4.1). The posterior distribution of the parameter vector in this marginal estimation step utilises marginal densities of $Y_{i,j}^{(n)}$ and is given by

$$f_{\Omega|Y^U}(\Omega|Y^U) \propto \left(\prod_{n=1}^N \prod_{i=1}^I \prod_{j=0}^{I-i} f_{Y_{i,j}^{(n)}}(y_{i,j}^{(n)}|\Omega) \right) f_{\delta}(\delta) \prod_{i=1}^I f_{\eta_i}(\eta_i) \prod_{j=0}^J f_{\nu_j}(\nu_j) f_{\gamma}(\gamma), \quad (4.2)$$

where

$$\Omega = \begin{pmatrix} \delta \\ \eta_i \\ \nu_j \\ \gamma \end{pmatrix}, \quad \eta_i = \begin{pmatrix} \eta_i^{(1)} \\ \eta_i^{(2)} \\ \vdots \\ \eta_i^{(N)} \end{pmatrix}, \quad \nu_j = \begin{pmatrix} \nu_j^{(1)} \\ \nu_j^{(2)} \\ \vdots \\ \nu_j^{(N)} \end{pmatrix}, \quad \text{and where } \gamma = \begin{pmatrix} \gamma^{(1)} \\ \gamma^{(2)} \\ \vdots \\ \gamma^{(N)} \end{pmatrix}. \quad (4.3)$$

In the second step of the estimation process, α and β are estimated. The multivariate density needs to be used because α and β plays separate roles in the multivariate density. To avoid instability in the MCMC chain when dealing with the multivariate density, all other parameters are held fixed at their marginal estimates. A restriction is also implied on α and β using δ . In particular, only α is estimated using the multivariate

density, and β is obtained such that δ remains at its marginal estimate. The posterior distribution of the parameter vector is given by

$$f_{\alpha|\mathbf{Y}^U, \Omega}(\alpha|\mathbf{Y}^U, \Omega) \propto \left(\prod_{i=1}^I \prod_{j=0}^{I-i} f_{\mathbf{Y}_{i,j}}(\mathbf{y}_{i,j}|\alpha, \Omega) \right) f_{\alpha}(\alpha). \quad (4.4)$$

Both posterior densities in (4.2) and (4.4) are not in recognisable forms. Hence Metropolis-Hastings algorithms are used to simulate from these densities. Summary statistics of these densities including the medians and standard deviations are then obtained using simulation draws from these densities.

4.3. Metropolis-Hastings algorithm

Metropolis-Hastings is a popular class of MCMC algorithms used in Bayesian inference to create posterior simulators for a wide range of models, especially when the posterior distribution is not in a recognisable form (Koop, 2003, Chapter 5). There have been a number of applications using the Metropolis-Hastings algorithms in the loss reserving literature such as Meyers (2009) and Peters, Shevchenko, and Wüthrich (2009). Because the posterior densities in both the marginal estimation and the multivariate estimation steps are not in closed form, Metropolis-Hastings is an appropriate choice.

Random walk Metropolis-Hastings algorithms are used for marginal estimation and multivariate estimation. As a rule of thumb, the optimal acceptance probability is 0.234 for a vector of parameters and is 0.44 for a single parameter (Brooks, Gelman, Jones, and Meng, 2011, Chapter 4). However, a rate that is neither too low nor too high can generally be accepted (Brooks, Gelman, Jones, and Meng, 2011). To achieve a desired acceptance rate, we choose appropriate values for the variance of the proposal densities. This process is referred to as tuning. This can be performed manually using trial runs or automatically using an adaptive Metropolis-Hastings algorithm with a coerced acceptance rate (see, for instance, Vihola, 2012).

From this algorithm, we obtain posterior draws of parameters in the model. A burn-in period might be required to remove the unstable part of the chains. We also need to thin the chains by keeping every d -th draw to break the dependence between draws (Kruschke, 2011, Chapter 23). The thinned draws are then used to compute the posterior median, standard deviation and credibility interval for parameters in the model.

4.4. Predictive distribution of outstanding claims

Given the simulated parameters from the posterior distributions, we can make predictions for outstanding claims in the lower loss triangles. To generate the m -th sample of outstanding claims, we use the m -th draw from the thinned simulation draws of parameters. This includes the m -th draw of Ω in the marginal estimation and m -th draw of α and β in the multivariate estimation. Outstanding claims in the lower triangles are then calculated using (2.11) with (2.9)-(2.10) for $1 < i \leq I$, $I - i + 1 \leq j \leq J$, $1 \leq n \leq N$.

Predictive distributions of outstanding claims can then be obtained using simulated draws of outstanding claims. The distribution of outstanding claims can also be obtained for each accident year, each loss triangle, and in total. Summary statistics of interest including the mean, variance and quantiles, for example 75% VaR and 95% VaR can then be calculated using the simulated sample.

5. Simulation illustrations

Two simulated data sets are used to assess the effectiveness of the fitting procedure. The first data set is generated from a multivariate Tweedie distribution with parameters chosen to replicate an empirical data set. It contains two standardised triangles of loss ratios with ten accident years and ten development years. This data set is presented in Table 10.10 and Table 10.11 in Appendix 10.1.

A subset of the Tweedie family of distributions, namely Tweedie's compound Poisson distributions with $1 < p < 2$, is an useful subset of distributions in loss reserving with the ability to accommodate masses at zeros; see, for example, Alai and Wüthrich (2009); Boucher and Davidov (2011); Wüthrich (2003). The

second simulation illustration is performed on a simulated data set with a 0 observation to assess the performance of the fitting procedure in presence of zeros. This data set is presented in Table 11.14 and Table 11.15 in Appendix 11.1.

5.1. Estimation of p

To find the power parameter p of the model, we set up a univariate Tweedie GLM log-likelihood profile for the combined data set of two loss triangles in which the log-likelihood function is written as a function of power parameter p . The mean structure of incremental claims $Y_{i,j}^{(n)}$ in this log-likelihood profile is given by

$$d + k_i^{(1)} I_{n=1} + u_j^{(1)} I_{n=1} + k_i^{(2)} I_{n=2} + u_j^{(2)} I_{n=2}, \quad (5.1)$$

where d , $k_i^{(n)}$ and $u_j^{(n)}$ are coefficients in the GLM regression and I is the indicator function. The power parameter p is found numerically by testing a range of values for p on the Tweedie GLM log-likelihood profile of the data set. When there exists at least one 0 observation in the data set, the numerical test range used to determine power parameter p is restricted to $(1, 2)$. The value of p that provides the highest likelihood is selected. The 95% CI of the estimate is also obtained using a χ_1^2 distribution approximation to the likelihood. The estimation results for simulated data set 1 and simulated data set 2 are provided in Table 5.2.

Simulated data set	True value	Estimate	95% CI
1	1.32	1.36	(1.21, 1.56)
2	1.32	1.24	(1.14, 1.37)

Table 5.2: Estimate of power parameter p and its 95% confidence interval in both simulated data sets

This analysis is based on the assumption of independence between lines. However, it can provide a reasonable estimate for parameter p as shown in Table 5.2. The estimate of p obtained from this analysis is then assumed to be the true p , when estimating other parameters in the Bayesian framework. Fixing the power parameter p in the Bayesian framework can significantly improve the stability of algorithm.

5.2. Marginal estimation

The marginal estimation step allows us to obtain estimates of all parameters except for α and β . We apply a log transformation on the parameters and use log-scaled parameters in the fitting. Prior distributions in the inference are in the form $\text{Uniform}(a_i, b_i)$. Lower and upper bounds for the prior distributions are chosen informatively using information obtained from the marginal maximum likelihood estimation as described in Section 4.1. Standard deviations of the proposal densities in the random walk Metropolis-Hastings are chosen so that the acceptance rate is reasonably close to 0.234. For both simulated data sets, we run 150,000 simulations and discard the first 50,000 iterations as the burn-in period. After this burn-in period, the MCMC chain approaches a stationary distribution. We provide the MCMC chain sample paths for simulated data set 1 in Figure 5.1 for illustration. We then use every 5th iteration to thin the sample. Estimation results for simulated data set 1 and simulated data set 2 are given in Table 10.12 in Appendix 10.2 and Table 11.16 in Appendix 11.2, respectively. The results represented include the posterior medians, standard deviations and 90% confidence intervals based on the 5th and the 95th quantiles of the posterior densities of parameters. Parameter estimates are obtained using the posterior medians due to the positive skewness of posterior densities. The posterior medians are also closer to the true parameter values than the posterior means. The results show that true values of parameters are reasonably close to the estimates and they all lie within the corresponding confidence intervals.

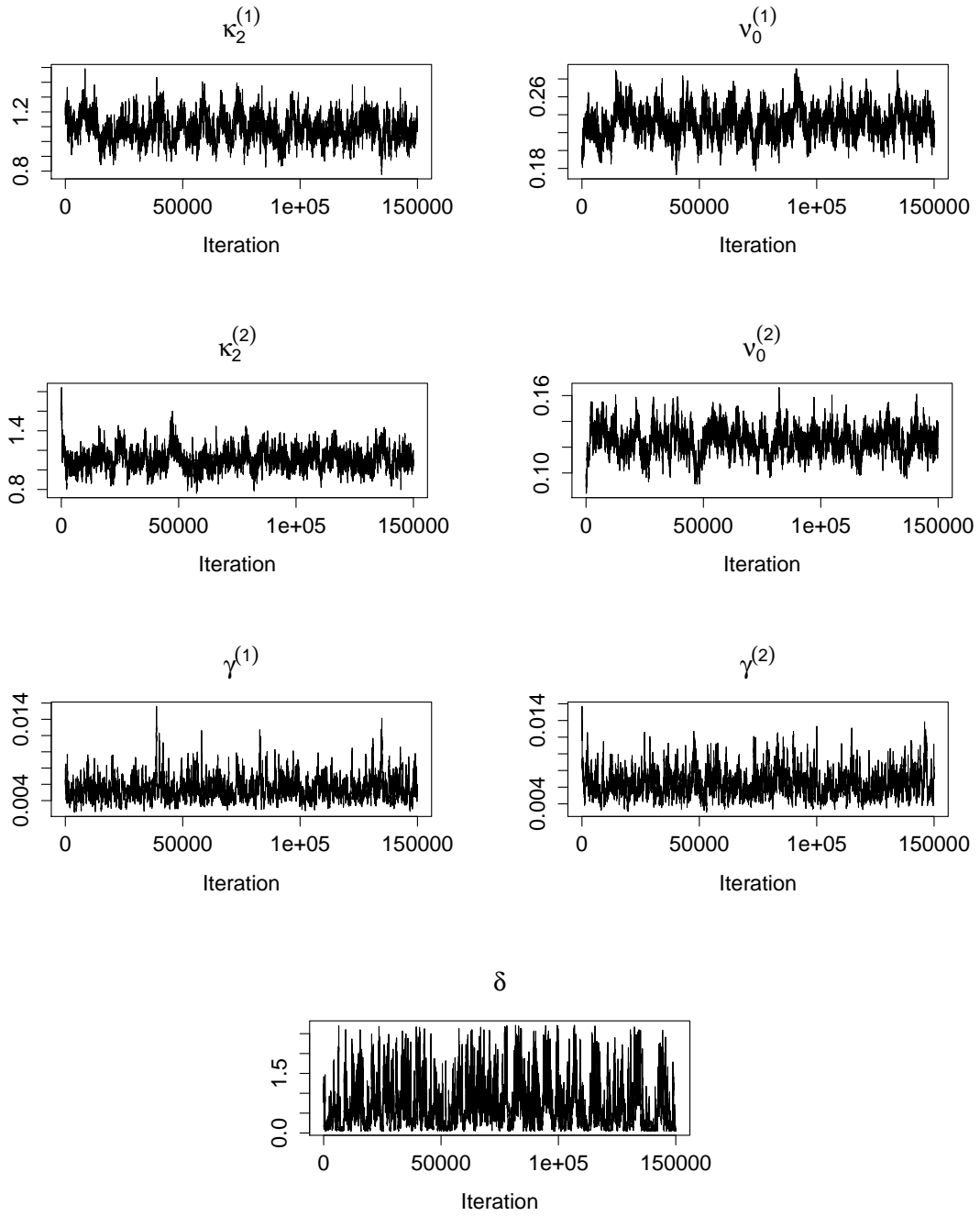


Figure 5.1: MCMC sample paths of $\eta_2^{(1)}, \nu_0^{(1)}, \eta_2^{(2)}, \nu_0^{(2)}, \gamma^{(1)}, \gamma^{(2)}, \delta$

5.3. Multivariate estimation

This step is used to estimate α and β , conditioning on estimated values of all other parameters. An estimate of δ is used to imply a restriction on α and β . In particular, only α is estimated and β is computed as a function of α and δ . Similarly to the marginal estimation, we also use a uniform prior distribution for α . For both simulated data sets, we run 30,000 simulations, discard the first 10,000 iterations as a burn-in period and use every 5th iteration of the rest. Estimates of α and β are provided in Table 10.13 in Appendix 10.2 for simulated data set 1 and Table 11.17 in Appendix 11.2 for simulated data set 2. True values of parameters are also reasonably close to their estimates and they all fall within the corresponding 90% confidence interval.

Both simulation illustrations show that the fitting procedure provides reasonably accurate estimates of parameters. This procedure will then be applied to a real data set in the following section.

6. Illustration with real data

The real data used for illustration consists of two business lines: the personal auto line and the commercial auto line. It belongs to the Pennsylvania National Insurance Group (Schedule P), and was used in Zhang and Dukic (2013). For the paper to be self contained, we also provide these loss triangles in Appendix 12.

6.1. Preliminary data analysis

A preliminary analysis is conducted on this data set. The first step of the analysis is to standardise incremental claims using the corresponding premium amount. This results in two triangles of loss ratios. Incremental loss ratios are converted into cumulative loss ratios and plots of cumulative loss ratios are provided in Figure 6.2. It can be observed that development patterns are quite similar for all years for the personal auto line. However, we observe some volatility in claims development patterns from the commercial auto line. In both lines of business, claims development tends to reach maturity at the end of the 10 year period, suggesting that these are both short-tailed lines.

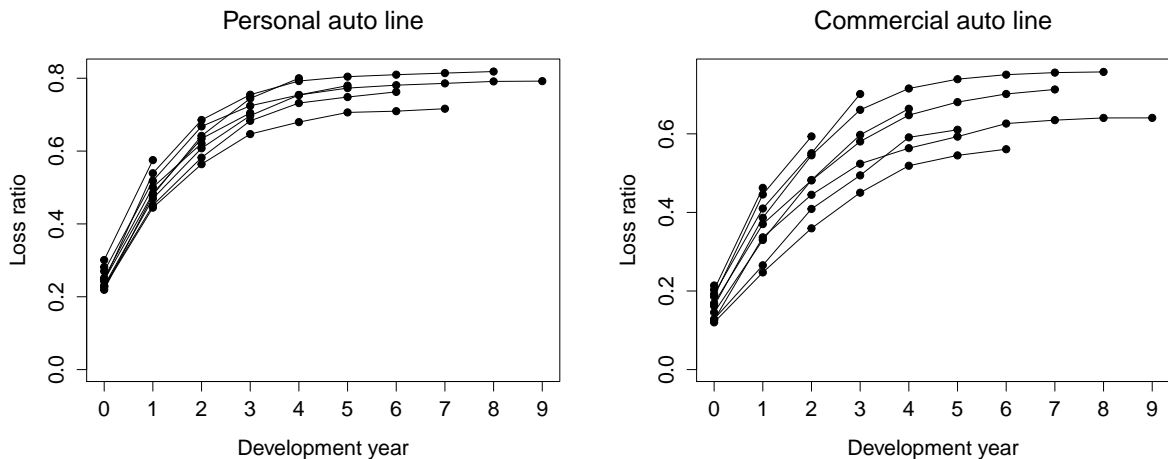


Figure 6.2: Loss ratios of personal auto line and commercial auto line

The power parameter p is found numerically by testing a range of values for p in order to choose a value that provides the highest likelihood. This analysis is performed for each line of business, as well as for the combined portfolio of both lines. The estimate and confidence interval are both provided in Table 6.3.

Line	\hat{p}	95% CI
Personal auto	1.15	(1.07, 1.40)
Commercial auto	1.39	(1.24, 1.63)
All lines	1.32	(1.21, 1.47)

Table 6.3: Estimate of power parameter p and its 95% confidence interval

As a requirement of the multivariate Tweedie distribution, all lines of business must have the same power parameter p . It can be observed from Table 6.3 that the estimate of p that provides the best fit to the combined portfolio is 1.32. This value is within the confidence interval of the estimates from the two individual lines. Hence it will be used in the fitting of the model to the data set.

We assess the dependence between lines of business by analysing the residuals after removing accident year and development year trends. This is performed by applying a GLM framework to both lines independently. To get the most accurate results, the GLM framework fitted to each line is the Tweedie dispersion with the power parameter p that provides the best fit to that line. In particular, a Tweedie distribution with power parameter $p = 1.15$ and a Tweedie distribution with power parameter $p = 1.39$ are each fitted to the personal auto line and the commercial auto line respectively. Residuals are computed. Cell-wise correlations and their p -values are then calculated on the residuals. Hypothesis tests of the Pearson's correlation and the Spearman's correlation are performed using asymptotic t distributions. The p -value of the Kendall's correlation is calculated using a normal approximation. Results are provided in Table 6.4.

Pearson	Spearman	Kendall
0.3879 (0.0034)	0.3752 (0.0050)	0.2538 (0.0062)

Table 6.4: Correlation between cell-wise residuals and the corresponding p -values

It can be observed that there are moderate but significant positive cell-wise correlations between the personal auto line and commercial auto line. This can also be compared with Kendall's correlation coefficient of 0.271 using log-normal margins in Zhang and Dukic (2013).

As pointed out in Avanzi, Taylor, and Wong (2016), careful modelling is needed before committing to any measure of correlation. To further study the dependence between lines of business in this data set, an examination of heat maps of the residuals from the above GLM Tweedie frameworks has been performed. This analysis reveals some systematic variation in residuals that affects earlier accident years. There also appears to be sympathetic variation that affects both loss triangles simultaneously. Consequently, one could make a decision whether these systematic deviations represent a signal that needs to be modelled or whether these are simply correlated noise. If these deviations were to be treated as a signal, an augmented model allowing these deviations as fixed effects could be considered. As a result, these additional features would be allowed for in the forecast of outstanding claims. By allowing more fixed effects in the model, the correlation coefficients between lines of business could be significantly reduced. However, there is no obvious physical interpretation of these systematic deviations, hence it is equally possible that they are simply correlated noise. For illustration purposes, we consider all variations as correlated noise and apply the multivariate Tweedie framework to this data set. However, it is worth noting that there can be other possible modelling approaches to capture features represented in this data set.

6.2. Bayesian inference and estimation results

For the estimation of margins, we run 300,000 iterations and discard the first 150,000 iterations as a burn-in sample. We then thin the sample to reduce the sequential dependence between each iteration in use. The posterior statistics of parameters obtained from these draws are given in Table 6.5.

We then use the multivariate estimation procedure to estimate α and β . Estimates of α and β are provided in Table 6.6.

	Median	SD	($Q_{0.05}; Q_{0.95}$)		Median	SD	($Q_{0.05}; Q_{0.95}$)
$\eta_2^{(1)}$	1.0203	0.0928	(0.8727; 1.1816)	$\eta_2^{(2)}$	1.1468	0.1200	(0.9694; 1.3567)
$\eta_3^{(1)}$	0.9316	0.0844	(0.8023; 1.0796)	$\eta_3^{(2)}$	1.1306	0.1176	(0.9581; 1.3390)
$\eta_4^{(1)}$	1.0220	0.0947	(0.8685; 1.1841)	$\eta_4^{(2)}$	0.8999	0.0966	(0.7598; 1.0727)
$\eta_5^{(1)}$	1.0479	0.0979	(0.8966; 1.2208)	$\eta_5^{(2)}$	1.0157	0.1066	(0.8564; 1.2036)
$\eta_6^{(1)}$	1.1024	0.1021	(0.9458; 1.2837)	$\eta_6^{(2)}$	1.1413	0.1262	(0.9546; 1.3704)
$\eta_7^{(1)}$	1.0089	0.0979	(0.8578; 1.1802)	$\eta_7^{(2)}$	1.3633	0.1456	(1.1518; 1.6222)
$\eta_8^{(1)}$	1.0028	0.1009	(0.8515; 1.1808)	$\eta_8^{(2)}$	1.4040	0.1634	(1.1638; 1.6974)
$\eta_9^{(1)}$	1.2071	0.1323	(1.0050; 1.4402)	$\eta_9^{(2)}$	1.5176	0.1921	(1.2371; 1.8635)
$\eta_{10}^{(1)}$	1.1863	0.1679	(0.9340; 1.4898)	$\eta_{10}^{(2)}$	1.5957	0.3030	(1.2073; 2.1689)
$\nu_0^{(1)}$	0.2334	0.0178	(0.2064; 0.2652)	$\nu_0^{(2)}$	0.1311	0.0116	(0.1123; 0.1503)
$\nu_1^{(1)}$	0.2369	0.0178	(0.2098; 0.2680)	$\nu_1^{(2)}$	0.1680	0.0142	(0.1448; 0.1914)
$\nu_2^{(1)}$	0.1343	0.0109	(0.1172; 0.1533)	$\nu_2^{(2)}$	0.1153	0.0109	(0.0975; 0.1335)
$\nu_3^{(1)}$	0.0779	0.0074	(0.0664; 0.0907)	$\nu_3^{(2)}$	0.0922	0.0092	(0.0776; 0.1077)
$\nu_4^{(1)}$	0.0405	0.0047	(0.0334; 0.0487)	$\nu_4^{(2)}$	0.0593	0.0071	(0.0485; 0.0720)
$\nu_5^{(1)}$	0.0186	0.0030	(0.0141; 0.0239)	$\nu_5^{(2)}$	0.0232	0.0040	(0.0171; 0.0303)
$\nu_6^{(1)}$	0.0066	0.0017	(0.0042; 0.0096)	$\nu_6^{(2)}$	0.0176	0.0034	(0.0124; 0.0236)
$\nu_7^{(1)}$	0.0043	0.0015	(0.0023; 0.0071)	$\nu_7^{(2)}$	0.0065	0.0021	(0.0036; 0.0105)
$\nu_8^{(1)}$	0.0039	0.0016	(0.0020; 0.0072)	$\nu_8^{(2)}$	0.0027	0.0013	(0.0014; 0.0054)
$\nu_9^{(1)}$	0.0005	0.0004	(0.0002; 0.0016)	$\nu_9^{(2)}$	0.0001	0.0001	(0.0001; 0.0003)
$\gamma^{(1)}$	0.0058	0.0014	(0.0040; 0.0085)	$\gamma^{(2)}$	0.0076	0.0019	(0.0054; 0.0114)
δ	0.8127	0.6487	(0.0986; 2.1341)				

Table 6.5: Posterior statistics of $\eta_i^{(1)}$, $\nu_j^{(1)}$, $\eta_i^{(2)}$, $\nu_j^{(2)}$, $\gamma^{(1)}$, $\gamma^{(2)}$, δ

	Median	SD	($Q_{0.05}; Q_{0.95}$)
α	0.0041	0.0047	(0.0011; 0.0157)
β	0.0293	0.0197	(0.0117; 0.0731)

Table 6.6: Posterior statistics of α and β

QQ plots of Pearson residuals from the fitting are used to assess the goodness of fit for the model. These are provided in Figure 6.3. These plots suggest that the multivariate Tweedie framework with $p = 1.32$ provides a good fit to the real data set used.

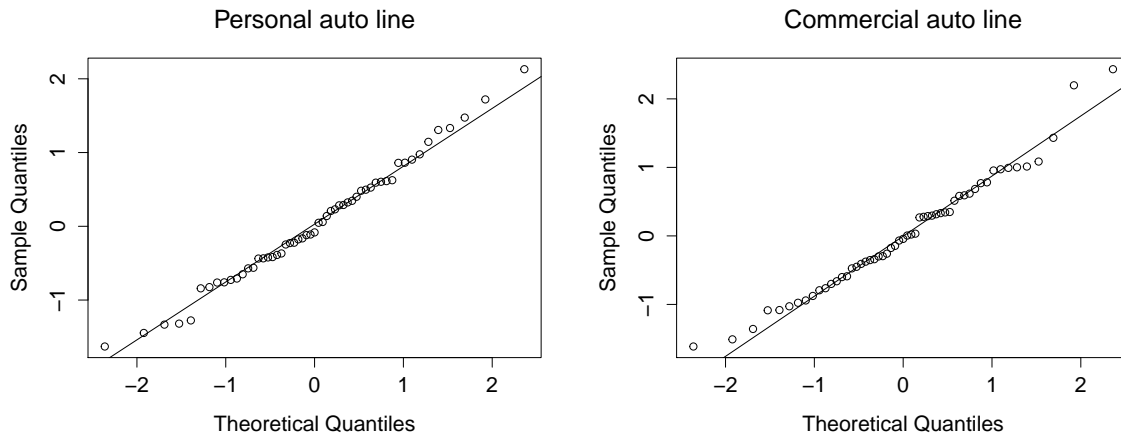


Figure 6.3: QQ plots of residuals from multivariate Tweedie ($p = 1.32$) fitting

6.3. Loss reserve predictions

Using the simulated draws of parameters from the Bayesian inference, we can obtain a predictive distribution of outstanding claims. We simulate 30,000 sets of lower loss triangles using 30,000 posterior draws of parameters. In particular, each set of lower loss triangle is simulated using a set of posterior parameter draw. Table 6.7 summarises claims forecasts by accident year for each line of business and the total portfolio. Results represented include the mean and average standard deviation of total claims from each accident year.

Year	Personal auto		Commercial auto		Total	
	Mean	SD	Mean	SD	Mean	SD
2	67.44	59.64	20.21	21.30	87.65	70.72
3	361.23	186.29	209.11	127.79	570.35	244.59
4	747.07	288.92	472.89	202.02	1,219.96	386.17
5	1,417.86	430.06	1,441.71	429.71	2,859.57	662.39
6	3,555.91	821.73	3,301.36	774.04	6,857.27	1,224.53
7	8,306.58	1,536.38	9,095.14	1,644.76	17,401.72	2,375.35
8	16,000.76	2,489.37	16,304.02	2,509.34	32,304.78	3,678.74
9	27,541.14	3,714.87	23,708.28	3,463.91	51,249.42	5,190.42
10	45,677.21	6,623.84	34,340.32	6,547.81	80,017.53	9,562.36

Table 6.7: Outstanding claims reserves by accident year (numbers are in \$1,000's)

From this simulation, we can also obtain predictive distributions of total outstanding claims for each line of business and the total portfolio. These densities are represented in Figure 6.4. Table 6.8 summarises estimation statistics of the total outstanding claims forecast. Results represented include the mean, variance, 75% and 95% VaR calculated directly from the MCMC simulation. We can observe that the sum of the standard deviations of total claims from each line is larger than the standard deviation of the total claims portfolio, suggesting a diversification benefit due to the lack of perfectly positive dependence between the two lines. This can be supported by the Pearson's correlation of 0.1212 between the total sum of claims of the two lines, obtained from pairs of total outstanding claims from each line in the set of 30,000 simulated lower loss triangles.

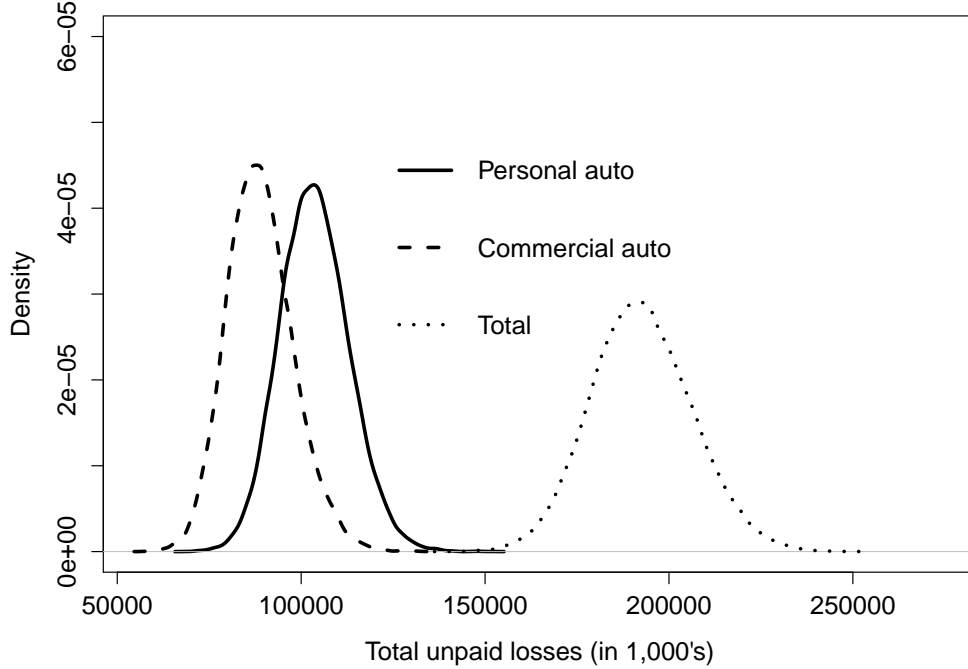


Figure 6.4: Kernel density of predictive distribution of total outstanding claims for each line of business and for the total portfolio (in \$1,000's)

	Personal auto	Commercial auto	Total
Mean	103,675.21 (54.11)	88,893.05 (52.13)	192,568.26 (79.56)
SD	9,372.74 (38.26)	9,028.83 (36.86)	13,779.84 (56.26)
75% VaR	109,765.72 (69.16)	94,542.25 (81.40)	201,636.91 (118.77)
95% VaR	119,584.51 (128.47)	104,657.06 (133.78)	215,961.11 (195.23)

Table 6.8: Summary of the loss reserves prediction with standard errors in brackets (numbers are in \$1,000's)

When assessing diversification benefits, it is also useful to look at diversification benefits obtained for risk margins. Risk margins are often held by insurers as a regulatory requirement to provide protection against unexpected variabilities from central outstanding claims estimates. Risk margin can be calculated as

$$\text{Risk margin}_\alpha = \alpha\% \text{ VaR} - \text{Mean}. \quad (6.1)$$

Using results from Table 6.8, risk margins calculated using 75% VaR and 95% VaR are provided in Table 6.9. Significant diversification benefits of approximately 29% and 35% can be observed for Risk margin₇₅ and Risk margin₉₅, respectively.

	Personal auto	Commercial auto	Total
Risk margin ₇₅	6,090.51	5,649.20	9,068.65
Risk margin ₉₅	15,909.30	15,764.01	23,392.85

Table 6.9: Risk margin statistics using 75% VaR and 95% VaR

Remark 6.1. *The bell-shaped curves in Figure 6.4 suggest that a Gaussian distribution would fit the aggregate outstanding liabilities relatively well. However, an appropriate choice of marginal distribution at the cell level is still of importance, as it defines which distribution is maximised in order to get parameters. This changes the prediction results, including the means, standard deviations and quantiles of outstanding liabilities.*

Such a nicely shaped distribution as a final result is, in fact, not surprising. This is even in the case of heavily skewed individual cells. A similar example can be found in Taylor (2000, Chapter 11). One might very well be able to estimate (moderate) quantiles of the aggregate by means of a Gaussian approximation. However, this cannot be determined before all calculations are made and Figure 6.4 is produced.

Remark 6.2. *The Monte Carlo simulation used produces very low standard errors and uses minimal computational time. However, in cases where the MCMC simulation cannot provide adequate accuracy or efficiency, one can further utilise the benefit of having closed form cumulants to obtain prediction statistics including the mean and the standard deviation of the sum of outstanding claims. It is worth noting that when there is a consideration of parameter uncertainty in a Bayesian framework, the mean and the variance of the sum of all claims given in Section 3 are random variables conditional on parameters. These conditional mean and variance can be used to compute unconditional mean and variance using the law of total expectation and the law of total variance. Being able to obtain the conditional mean and variance of total outstanding claims in closed form also allows one to compute quantiles of interest using the control variate method. This method is well-known for producing estimates with lower standard error; see for example, Givens and Hoeting (2005); Glasserman (2003); Kroese, Taimre, and Botev (2011).*

7. Conclusion

In this paper, we introduced a multivariate Tweedie approach with a common shock structure in the multivariate loss reserving field. This approach has a variety of desirable advantages, including

- Dependence is introduced with the help of a easily identifiable, explicit common shock dependence structure, which is easily generalised to more than two dimensions.
- While all dimensions must have same parameter p , this can be anywhere in $(-\infty, 0] \cap [1, \infty)$ (rather than fixed at 1 for a Poisson dispersion or 2 for a gamma dispersion, for instance). The flexibility of Tweedie dispersions has been established in Alai, Landsman, and Sherris (2015); Alai and Wüthrich (2009); Furman and Landsman (2010); Jørgensen (1997). Additionally, zero data points do not present any issue.
- The distributions of (multivariate) margins belong to the same Tweedie family. Furthermore, moments and cumulants that can be obtained analytically, and cumulants of the sum (of reserves) can be given in closed form.

A Bayesian framework with a two step approach is used to develop a more efficient estimation procedure when $p \neq 1$ (a case that is not available when using common software). The performance of model estimation was assessed using simulated data sets, and illustrated using a real data set from Schedule P. In the illustration with real data, it can be shown that the marginal flexibility property allows the multivariate Tweedie model to provide a very good fit to the data set. The results are obtained efficiently using MCMC, reducing the need to use closed form cumulants of the sum of outstanding claims. However, one can also take the advantage of having tractable cumulants of outstanding claims to efficiently compute the mean and the variance of total outstanding losses.

The multivariate Tweedie approach introduced in this paper is a general approach that provides great flexibility with potential developments. However, a limitation of the proposed approach comes from the requirement of using the same power parameter p for all lines of business as well as the same common shock for all business lines. In our model development, for the sake of simplicity, we have performed the most simple parametrisation for the common shock variable. To be more realistic, one might consider modifying this parametrisation such that the relative effect of common shocks is not identical across all loss cells in loss triangles. Further research can also consider utilising this approach to capture other sources of dependence, such as development year dependence and calendar year dependence.

Acknowledgments

The authors are very grateful for comments from two anonymous reviewers that helped significantly improve the paper. Results in this paper were presented at the Actuarial and Financial Mathematics Seminar of the Quantact Lab in Montreal (Canada). The authors are grateful for constructive comments received from colleagues who attended this seminar.

This research was supported under Australian Research Council's Linkage Projects funding scheme (project number LP130100723). Furthermore, Phuong Anh Vu acknowledges financial support from a University International Postgraduate Award and supplementary scholarships provided by the UNSW Business School, as well as travel funds from a grant of the Natural Science and Engineering Research Council of Canada. The views expressed herein are those of the authors and are not necessarily those of the supporting organisations.

Appendices

A. The Tweedie family of distributions

A random variable Z belonging to the EDF has two representations known as additive and reproductive forms (Jørgensen, 1997, Chapter 3). The additive form has two parameters: a canonical parameter θ and an index parameter λ . The canonical parameter θ belongs to the domain

$$\Theta = \{\theta \in \mathbb{R} : \kappa(\theta) < \infty\}, \quad (\text{A.1})$$

where κ is the unit cumulant function (the cumulant function with a unit dispersion parameter). It is defined mathematically by

$$\kappa(\theta) = \log \int \exp(\theta z) dP(z), \quad (\text{A.2})$$

where $dP(z)$ is a probability measure. The index parameter λ belongs to the index set Λ such that

$$\lambda \kappa(\theta) = \log \int \exp(\theta z) dP_\lambda(z), \quad (\text{A.3})$$

for some probability measure $dP_\lambda(z)$. The index parameter λ plays an intuitive role of allowing non-unit dispersion in the distribution.

Let us now introduce the additive exponential dispersion distribution $\text{ED}^*(\theta, \lambda)$, which has probability measure

$$\exp(\theta z - \lambda \kappa(\theta)) dP_\lambda(z). \quad (\text{A.4})$$

The moment generating function $M_Z(t)$ corresponding to the probability measure in (8.4) is

$$M_Z(t) = \exp(\lambda[\kappa(\theta + t) - \kappa(\theta)]). \quad (\text{A.5})$$

According to Jørgensen (1997, Chapter 3), the family of distributions for $Z' = Z/\lambda$ is called the reproductive exponential dispersion distribution $\text{ED}(\mu, \sigma^2)$ generated by the measure

$$\exp\{\lambda[\theta z' - \kappa(\theta)]\} d\bar{P}_\lambda(z'), \quad (\text{A.6})$$

where $d\bar{P}_\lambda$ is transformed from dP_λ using the change of variable $Z' = Z/\lambda$. Parameter $\mu = \kappa'(\theta)$ is called the mean value parameter and $\sigma^2 = 1/\lambda$ is called the dispersion parameter.

We now introduce the Tweedie family, a subclass of the EDF. This family also has two representations and general properties of the EDF. We denote by $\text{Tweedie}^*(.,.)$ the additive Tweedie distribution and $\text{Tweedie}(.,.)$ the reproductive Tweedie distribution. The unique property that distinguishes the Tweedie

family of distributions from other distributions of the EDF is the special relationship between the variance function and the mean function

$$V(\mu) = \mu^p, \quad (\text{A.7})$$

where p is the power parameter. The canonical parameter θ of the additive Tweedie form belongs to the set Θ_p that satisfies

$$\Theta_p = \begin{cases} \mathbb{R}, & p = 0, 1, \\ [0, \infty), & p < 0, \\ (-\infty, 0), & 1 < p \leq 2, \\ (-\infty, 0], & 2 < p < \infty. \end{cases} \quad (\text{A.8})$$

The unit cumulant function of the Tweedie family is denoted by $\kappa_p(\theta)$ with the subscript p added to emphasise that the Tweedie family is discussed. As defined in Jørgensen (1997, Chapter 4) the unit cumulant function is expressed by

$$\kappa_p(\theta) = \begin{cases} \exp(\theta), & p = 1, \\ -\log(-\theta), & p = 2, \\ \frac{\tau - 1}{\tau} \left(\frac{\theta}{\tau - 1} \right)^\tau, & p \notin (0, 1] \cup [2], \end{cases} \quad (\text{A.9})$$

where $\tau = (p - 2)/(p - 1)$.

As also provided in Jørgensen (1997)[Chapter 4], the relationship between the additive form and the reproductive form of a Tweedie distribution is

$$\text{Tweedie}_p^*(\theta, \lambda) = \text{Tweedie}(\lambda \kappa_p'(\theta), \lambda^{1-p}). \quad (\text{A.10})$$

This relationship can be used to convert one Tweedie representation to another and specify relationships between parameters of these two forms.

In the Tweedie family, the value of the power parameter p identifies the corresponding distribution. For example, $p = 0$ corresponds to the normal distribution, $p = 1$ corresponds to the Poisson distribution, $p = 2$ corresponds to the gamma distribution, $1 < p < 2$ corresponds to the compound Poisson-gamma distribution (i.e. a Poisson sum of gamma random variables), and $p = 3$ corresponds to the inverse Gaussian distribution. These members of the Tweedie family are very commonly used distributions of the EDF.

B. Moment generating functions

Using parameters in the reproductive representation as represented in Section 2.2, moment generating function of the incremental claims $Y_{i,j}^{(n)}$ is given by

$$M_{W_{i,j}}(t) = \begin{cases} \exp\left(\beta^{\frac{1}{1-p}} \left[\kappa_p\left(\frac{\alpha^{1-p}}{\beta(1-p)} + t\right) - \kappa_p\left(\frac{\alpha^{1-p}}{\beta(1-p)}\right) \right]\right), & p \neq 1, \\ \exp\left(\frac{\alpha}{\beta}(\exp(\beta t) - 1)\right), & p = 1, \end{cases} \quad (\text{B.1})$$

$$M_{Z_{i,j}^{(n)}}(t) = \begin{cases} \exp\left((\gamma^{(n)})^{\frac{1}{1-p}} \left[\kappa_p\left(\frac{(\eta_i^{(n)} \nu_j^{(n)})^{1-p}}{\gamma^{(n)}(1-p)} + t\right) - \kappa_p\left(\frac{(\eta_i^{(n)} \nu_j^{(n)})^{1-p}}{\gamma^{(n)}(1-p)}\right) \right]\right), & p \neq 1, \\ \exp\left(\frac{\eta_i^{(n)} \nu_j^{(n)}}{\gamma^{(n)}}(\exp(\gamma^{(n)} t) - 1)\right), & p = 1, \end{cases} \quad (\text{B.2})$$

The m -th cumulants of $W_{i,j}$ and $Z_{i,j}^{(n)}$, denoted by $K_{W_{i,j}}^{(m)}$ and $K_{Z_{i,j}^{(n)}}^{(m)}$, can be obtained from the moment generating functions in (9.1) and (9.2) by

$$K_{W_{i,j}}^{(m)} = \frac{\partial^m \log M_{W_{i,j}}(t)}{\partial t^m} \Big|_{t=0} = \begin{cases} \beta^{\frac{1}{1-p}} \kappa_p^{(m)} \left(\frac{\alpha^{1-p}}{\beta(1-p)} \right), & p \neq 1, \\ \beta^{m-1} \alpha, & p = 1, \end{cases} \quad (\text{B.3})$$

$$K_{Z_{i,j}^{(n)}}^{(m)} = \frac{\partial^m \log M_{Z_{i,j}^{(n)}}(t)}{\partial t^m} \Big|_{t=0} = \begin{cases} (\gamma^{(n)})^{\frac{1}{1-p}} \kappa_p^{(m)} \left(\frac{(\eta_i^{(n)} \nu_j^{(n)})^{1-p}}{\gamma^{(n)}(1-p)} \right), & p \neq 1, \\ \eta_i^{(n)} \nu_j^{(n)} (\gamma^{(n)})^{m-1}, & p = 1, \end{cases} \quad (\text{B.4})$$

where $\kappa_p^{(m)}(\cdot)$ is the m -th derivative of the unit cumulant function $\kappa(\cdot)$ in (8.9).

As a result of the independence between these two components, the moment-generating function of the incremental claims $Y_{i,j}^{(n)}$ is given by

$$M_{Y_{i,j}^{(n)}}(t) = \begin{cases} \exp \left(\beta^{\frac{1}{1-p}} \left[\kappa_p \left(\frac{\alpha^{1-p}}{\beta(1-p)} + \left(\frac{\alpha}{\eta_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} t \right) - \kappa_p \left(\frac{\alpha^{1-p}}{\beta(1-p)} \right) \right] \right) \\ \times \exp \left((\gamma^{(n)})^{\frac{1}{1-p}} \left[\kappa_p \left(\frac{(\eta_i^{(n)} \nu_j^{(n)})^{1-p}}{\gamma^{(n)}(1-p)} + t \right) - \kappa_p \left(\frac{(\eta_i^{(n)} \nu_j^{(n)})^{1-p}}{\gamma^{(n)}(1-p)} \right) \right] \right), & p \neq 1, \\ \exp \left(\frac{\alpha}{\beta} (\exp(\gamma^{(n)} t) - 1) + \frac{\eta_i^{(n)} \nu_j^{(n)}}{\gamma^{(n)}} (\exp(\gamma^{(n)} t) - 1) \right), & p = 1. \end{cases} \quad (\text{B.5})$$

C. Simulated data set 1

C.1. Simulated loss triangles

Year	0	1	2	3	4	5	6	7	8	9
1	0.240151	0.274318	0.094119	0.085875	0.031807	0.013806	0.007190	0.009306	0.001866	0.002008
2	0.253992	0.251274	0.138049	0.068166	0.054149	0.018325	0.009745	0.006340	0.002740	
3	0.234966	0.283883	0.125763	0.069204	0.033658	0.014891	0.004963	0.004028		
4	0.225352	0.244446	0.107632	0.079135	0.045705	0.020370	0.003769			
5	0.218054	0.223552	0.154835	0.087825	0.039442	0.014322				
6	0.302805	0.275805	0.143042	0.085948	0.043431					
7	0.239969	0.296280	0.129593	0.090236						
8	0.200759	0.221279	0.157823							
9	0.291852	0.243345								
10	0.244305									

Table 3.10: Simulated loss triangle 1 (data set 1)

Year	0	1	2	3	4	5	6	7	8	9
1	0.147347	0.196444	0.103671	0.092007	0.047120	0.020597	0.016839	0.005074	0.003454	0.000225
2	0.138343	0.200295	0.112917	0.094464	0.075802	0.019854	0.022214	0.004711	0.008808	
3	0.138050	0.184805	0.145138	0.100257	0.053106	0.028627	0.043037	0.002824		
4	0.099135	0.119242	0.104156	0.108336	0.032670	0.028404	0.025760			
5	0.124336	0.217746	0.137297	0.091600	0.064107	0.028242				
6	0.148453	0.155605	0.131527	0.134106	0.060104					
7	0.187909	0.246437	0.149696	0.116582						
8	0.194526	0.233601	0.172806							
9	0.211970	0.280598								
10	0.213647									

Table 3.11: Simulated loss triangle 2 (data set 1)

C.2. Estimation results

	True value	Median	SD	$(Q_{0.05}; Q_{0.95})$		True value	Median	SD	$(Q_{0.05}; Q_{0.95})$
$\eta_2^{(1)}$	1.0362	1.0814	0.0875	(0.9460; 1.2351)	$\eta_2^{(2)}$	1.1957	1.0996	0.1004	(0.9463; 1.2765)
$\eta_3^{(1)}$	0.9446	1.0024	0.0857	(0.8653; 1.1473)	$\eta_3^{(2)}$	1.1738	1.1498	0.1020	(0.9955; 1.3311)
$\eta_4^{(1)}$	1.0263	0.9828	0.0833	(0.8508; 1.1201)	$\eta_4^{(2)}$	0.9506	0.8706	0.0833	(0.7416; 1.0146)
$\eta_5^{(1)}$	1.0580	1.0111	0.0828	(0.8803; 1.1504)	$\eta_5^{(2)}$	1.0549	1.1370	0.1077	(0.9764; 1.3330)
$\eta_6^{(1)}$	1.1140	1.1671	0.0974	(1.0164; 1.3385)	$\eta_6^{(2)}$	1.1873	1.1495	0.1099	(0.9755; 1.3402)
$\eta_7^{(1)}$	1.0215	1.1116	0.0961	(0.9650; 1.2783)	$\eta_7^{(2)}$	1.4094	1.3904	0.1309	(1.1910; 1.6181)
$\eta_8^{(1)}$	1.0198	0.9772	0.0895	(0.8441; 1.1343)	$\eta_8^{(2)}$	1.4505	1.4697	0.1492	(1.2543; 1.7436)
$\eta_9^{(1)}$	1.2193	1.1174	0.1190	(0.9329; 1.3257)	$\eta_9^{(2)}$	1.5616	1.6773	0.1783	(1.4039; 1.9882)
$\eta_{10}^{(1)}$	1.1992	1.0530	0.1428	(0.8416; 1.3095)	$\eta_{10}^{(2)}$	1.6646	1.6936	0.2721	(1.3143; 2.1981)
$\nu_0^{(1)}$	0.2343	0.2308	0.0158	(0.2071; 0.2584)	$\nu_0^{(2)}$	0.1295	0.1250	0.0095	(0.1096; 0.1407)
$\nu_1^{(1)}$	0.2376	0.2426	0.0165	(0.2178; 0.2706)	$\nu_1^{(2)}$	0.1651	0.1658	0.0122	(0.1464; 0.1871)
$\nu_2^{(1)}$	0.1349	0.1245	0.0094	(0.1108; 0.1412)	$\nu_2^{(2)}$	0.1133	0.1123	0.0090	(0.0980; 0.1276)
$\nu_3^{(1)}$	0.0788	0.0757	0.0063	(0.0662; 0.0871)	$\nu_3^{(2)}$	0.0910	0.0934	0.0079	(0.0803; 0.1064)
$\nu_4^{(1)}$	0.0412	0.0388	0.0040	(0.0326; 0.0457)	$\nu_4^{(2)}$	0.0589	0.0504	0.0053	(0.0422; 0.0594)
$\nu_5^{(1)}$	0.0190	0.0152	0.0022	(0.0120; 0.0192)	$\nu_5^{(2)}$	0.0234	0.0227	0.0032	(0.0181; 0.0287)
$\nu_6^{(1)}$	0.0072	0.0058	0.0013	(0.0039; 0.0081)	$\nu_6^{(2)}$	0.0195	0.0246	0.0030	(0.0192; 0.0289)
$\nu_7^{(1)}$	0.0051	0.0057	0.0014	(0.0036; 0.0082)	$\nu_7^{(2)}$	0.0052	0.0035	0.0009	(0.0026; 0.0055)
$\nu_8^{(1)}$	0.0034	0.0022	0.0007	(0.0016; 0.0037)	$\nu_8^{(2)}$	0.0032	0.0052	0.0016	(0.0030; 0.0084)
$\nu_9^{(1)}$	0.0009	0.0016	0.0010	(0.0007; 0.0039)	$\nu_9^{(2)}$	0.0001	0.0002	0.0001	(0.0001; 0.0004)
$\gamma^{(1)}$	0.0049	0.0052	0.0014	(0.0037; 0.0080)	$\gamma^{(2)}$	0.0062	0.0061	0.0015	(0.0042; 0.0092)
δ	0.5700	0.5170	0.5891	(0.0719; 1.9067)					

Table 3.12: Posterior statistics of $\eta_i^{(1)}$, $\nu_j^{(1)}$, $\eta_i^{(2)}$, $\nu_j^{(2)}$, $\gamma^{(1)}$, $\gamma^{(2)}$, δ

	True value	Median	SD	$(Q_{0.05}; Q_{0.95})$
α	0.0060	0.0042	0.0047	(0.0011; 0.0160)
β	0.0664	0.0578	0.0361	(0.0242; 0.1372)

Table 3.13: Posterior statistics of α and β

D. Simulated data set 2

D.1. Simulated loss triangles

Year	0	1	2	3	4	5	6	7	8	9
1	0.283732	0.226835	0.164292	0.089160	0.029220	0.015849	0.013507	0.011700	0.003694	0.007537
2	0.357937	0.251767	0.177964	0.078834	0.051595	0.029831	0.001342	0.009320	0.009119	
3	0.203430	0.172260	0.109408	0.084336	0.027601	0.008246	0.014475	0.007355		
4	0.223365	0.227374	0.153072	0.054501	0.049119	0.038225	0			
5	0.279627	0.201581	0.191577	0.113466	0.059034	0.023083				
6	0.196836	0.323561	0.199684	0.113159	0.065494					
7	0.214060	0.110836	0.177897	0.098204						
8	0.253372	0.184663	0.149160							
9	0.213116	0.320763								
10	0.210801									

Table 4.14: Simulated loss triangle 1 (data set 2)

Year	0	1	2	3	4	5	6	7	8	9
1	0.104007	0.227287	0.129769	0.108286	0.069063	0.039293	0.025876	0.010197	0.003272	0.001153
2	0.157652	0.220918	0.174850	0.104639	0.092709	0.029221	0.020151	0.003365	0.004267	
3	0.144728	0.199293	0.139864	0.118673	0.084007	0.033882	0.039586	0.007387		
4	0.113809	0.138278	0.080297	0.080579	0.086192	0.029228	0.023066			
5	0.110246	0.191452	0.116688	0.145530	0.064009	0.032577				
6	0.126303	0.178031	0.146857	0.145787	0.071391					
7	0.247466	0.268015	0.195683	0.204113						
8	0.208114	0.202835	0.167232							
9	0.131089	0.309236								
10	0.306811									

Table 4.15: Simulated loss triangle 2 (data set 2)

D.2. Estimation results

	True value	Median	SD	$(Q_{0.05}; Q_{0.95})$		True value	Median	SD	$(Q_{0.05}; Q_{0.95})$
$\eta_2^{(1)}$	1.0362	1.1948	0.2014	(0.9195; 1.5760)	$\eta_2^{(2)}$	1.1956	1.1364	0.1249	(0.9482; 1.3529)
$\eta_3^{(1)}$	0.9445	0.7882	0.1294	(0.6380; 1.0514)	$\eta_3^{(2)}$	1.1737	1.1188	0.1237	(0.9299; 1.3358)
$\eta_4^{(1)}$	1.0262	0.9545	0.1602	(0.7222; 1.2449)	$\eta_4^{(2)}$	0.9506	0.8099	0.0901	(0.6969; 0.9858)
$\eta_5^{(1)}$	1.0579	1.1306	0.1930	(0.8649; 1.4904)	$\eta_5^{(2)}$	1.0549	1.0010	0.1230	(0.8183; 1.2216)
$\eta_6^{(1)}$	1.1140	1.2226	0.2040	(0.9405; 1.6071)	$\eta_6^{(2)}$	1.1873	1.0788	0.1296	(0.8713; 1.3015)
$\eta_7^{(1)}$	1.0214	0.8522	0.1570	(0.6605; 1.1706)	$\eta_7^{(2)}$	1.4093	1.6650	0.1628	(1.3958; 1.9331)
$\eta_8^{(1)}$	1.0197	0.9316	0.1761	(0.6890; 1.2700)	$\eta_8^{(2)}$	1.4504	1.3178	0.1778	(1.0435; 1.6355)
$\eta_9^{(1)}$	1.2193	1.1846	0.2366	(0.8652; 1.6185)	$\eta_9^{(2)}$	1.5616	1.3745	0.2127	(1.0434; 1.7579)
$\eta_{10}^{(1)}$	1.1991	0.9273	0.2526	(0.6534; 1.4578)	$\eta_{10}^{(2)}$	1.6645	1.8487	0.1497	(1.5154; 1.9944)
$\nu_0^{(1)}$	0.2342	0.2238	0.0280	(0.1818; 0.2753)	$\nu_0^{(2)}$	0.1295	0.1282	0.0100	(0.1111; 0.1440)
$\nu_1^{(1)}$	0.2375	0.2051	0.0268	(0.1620; 0.2501)	$\nu_1^{(2)}$	0.1651	0.1782	0.0155	(0.1545; 0.2059)
$\nu_2^{(1)}$	0.1348	0.1520	0.0219	(0.1177; 0.1896)	$\nu_2^{(2)}$	0.1132	0.1204	0.0127	(0.1015; 0.1431)
$\nu_3^{(1)}$	0.0787	0.0788	0.0141	(0.0583; 0.1050)	$\nu_3^{(2)}$	0.0909	0.1035	0.0061	(0.0903; 0.1098)
$\nu_4^{(1)}$	0.0412	0.0370	0.0092	(0.0238; 0.0539)	$\nu_4^{(2)}$	0.0588	0.0711	0.0094	(0.0573; 0.0881)
$\nu_5^{(1)}$	0.0189	0.0165	0.0061	(0.0089; 0.0285)	$\nu_5^{(2)}$	0.0234	0.0279	0.0059	(0.0195; 0.0389)
$\nu_6^{(1)}$	0.0071	0.0045	0.0024	(0.0027; 0.0103)	$\nu_6^{(2)}$	0.0195	0.0223	0.0058	(0.0144; 0.0336)
$\nu_7^{(1)}$	0.0050	0.0049	0.0030	(0.0027; 0.0123)	$\nu_7^{(2)}$	0.0051	0.0044	0.0020	(0.0027; 0.0087)
$\nu_8^{(1)}$	0.0034	0.0041	0.0028	(0.0022; 0.0111)	$\nu_8^{(2)}$	0.0032	0.0026	0.0014	(0.0016; 0.0060)
$\nu_9^{(1)}$	0.0009	0.0023	0.0038	(0.0004; 0.0122)	$\nu_9^{(2)}$	0.0002	0.0004	0.0007	(0.0001; 0.0023)
$\gamma^{(1)}$	0.0200	0.0201	0.0022	(0.0185; 0.0250)	$\gamma^{(2)}$	0.0120	0.0110	0.0012	(0.0101; 0.0139)
δ	0.7994	0.8264	0.4250	(0.1224; 1.5379)					

Table 4.16: Posterior statistics of $\eta_i^{(1)}$, $\nu_j^{(1)}$, $\eta_i^{(2)}$, $\nu_j^{(2)}$, $\gamma^{(1)}$, $\gamma^{(2)}$, δ

	True value	Median	SD	$(Q_{0.05}; Q_{0.95})$
α	0.0360	0.0794	0.2445	(0.0087; 0.7758)
β	0.1000	0.1764	0.3080	(0.0327; 0.9978)

Table 4.17: Posterior statistics of α, β

E. Empirical data set

This data set is drawn from Zhang and Dukic (2013).

Year	Premium	0	1	2	3	4	5	6	7	8	9
1	62,467	16,864	15,508	9,341	3,537	1,853	1,184	500	308	338	50
2	59,821	14,528	17,727	8,747	4,149	2,252	715	325	261	255	
3	62,968	14,241	13,763	7,512	5,207	2,068	1,674	219	421		
4	64,453	14,765	14,323	8,426	6,513	3,144	1,067	913			
5	71,185	16,395	17,038	9,826	6,381	4,037	1,839				
6	82,793	18,136	21,582	13,415	8,519	4,583					
7	100,826	24,727	24,037	15,181	7,105						
8	98,358	24,749	24,501	11,830							
9	76,653	23,063	21,035								
10	71,326	20,083									

Table 5.18: Personal auto line (in 1,000s)

Year	Premium	0	1	2	3	4	5	6	7	8	9
1	42,847	5,407	9,015	4,641	3,384	1,695	1,262	1,425	373	241	6
2	38,829	6,279	8,725	6,172	4,494	2,110	919	447	202	69	
3	43,001	7,256	8,667	4,778	4,262	2,884	1,427	889	493		
4	41,840	5,028	5,317	4,697	3,795	2,871	1,100	657			
5	44,525	5,721	6,097	6,389	3,802	4,306	862				
6	50,923	7,413	9,385	7,772	5,850	3,383					
7	56,601	10,868	12,337	7,966	8,531						
8	54,609	10,143	14,193	8,070							
9	47,204	9,596	12,235								
10	42,412	9,076									

Table 5.19: Commercial auto line (in 1,000s)

References

- Abdallah, A., Boucher, J.-P., Cossette, H., 2015. Modeling dependence between loss triangles with hierarchical Archimedean copulas. *ASTIN Bulletin* 45, 577–599.
- Alai, D. H., Landsman, Z., Sherris, M., 2015. Multivariate Tweedie lifetimes: The impact of dependence. *Scandinavian Actuarial Journal* In press.
- Alai, D. H., Wüthrich, M. V., 2009. Taylor approximations for model uncertainty within the Tweedie exponential dispersion family. *ASTIN Bulletin* 39 (2), 453–477.
- Avanzi, B., Taylor, G. C., Wong, B., 2016. Correlations between insurance lines of business: An illusion or a real phenomenon? Some methodological considerations. *ASTIN Bulletin* 46 (2), 225–263.
- Boucher, J.-P., Davidov, D., 2011. On the importance of dispersion modeling for claims reserving: An application with the Tweedie distribution. *Variance* 5 (2), 158–172.
- Brooks, S., Gelman, A., Jones, G. L., Meng, X.-L. (Eds.), 2011. *Handbook of Markov Chain Monte Carlo*. Boca Raton: Chapman & Hall.
- Bühlmann, H., Moriconi, F., 2015. Credibility claims reserving with stochastic diagonal effects. *ASTIN Bulletin* 45 (2), 309–353.
- Congdon, P. D., 2010. *Applied Bayesian hierarchical methods*. Boca Raton: Chapman & Hall.
- De Alba, E., 2002. Bayesian estimation of outstanding claim reserves. *North American Actuarial Journal* 6 (4), 1–20.
- De Alba, E., 2006. Claims reserving when there are negative values in the runoff triangle: Bayesian analysis using the three-parameter log-normal distribution. *North American Actuarial Journal* 10 (3), 45–59.
- De Jong, P., 2012. Modeling dependence between loss triangles. *North American Actuarial Journal* 16 (1), 74–86.
- England, P. D., Verrall, R. J., 2002. Stochastic claims reserving in general insurance. *British Actuarial Journal* 8 (3), 443–518.
- England, P. D., Verrall, R. J., 2006. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science* 1 (02), 221–270.
- Furman, E., 2008. On a multivariate gamma distribution. *Statistics and Probability Letters* 78 (15), 2353–2360.

- Furman, E., Landsman, Z., 2010. Multivariate Tweedie distributions and some related capital-at-risk analyses. *Insurance: Mathematics and Economics* 46 (2), 351–361.
- Givens, G. H., Hoeting, J. A., 2005. *Computational statistics*. New Jersey: John Wiley & Sons.
- Glasserman, P., 2003. *Monte Carlo methods in financial engineering*. New York: Springer.
- Hess, K. T., Schmidt, K. D., Zocher, M., 2006. Multivariate loss prediction in the multivariate additive model. *Insurance: Mathematics and Economics* 39 (2), 185–191.
- Joe, H., 1997. *Multivariate models and dependence concepts*. London: Chapman & Hall.
- Jørgensen, B., 1997. *The theory of dispersion models*. London: Chapman & Hall.
- Karlis, D., 2003. An EM algorithm for multivariate Poisson distribution and related models. *Journal of Applied Statistics* 30 (1), 63–77.
- Kocherlakota, S., Kocherlakota, K., 1992. *Bivariate discrete distributions*. New York: Marcel Dekker.
- Koop, G., 2003. *Bayesian econometrics*. Chichester: John Wiley & Sons.
- Kroese, D. P., Taimre, T., Botev, Z. I., 2011. *Handbook of Monte Carlo Methods*. New Jersey: John Wiley & Sons.
- Kruschke, J. K., 2011. *Doing Bayesian data analysis: A tutorial introduction with R*. Burlington: Academic Press.
- Kunkler, M., 2006. Modelling negatives in stochastic reserving models. *Insurance: Mathematics and Economics* 38 (3), 540–555.
- Merz, M., Wüthrich, M. V., 2009a. Combining chain-ladder and additive loss reserving method for dependent lines of business. *Variance* 3 (2), 270–291.
- Merz, M., Wüthrich, M. V., 2009b. Prediction error of the multivariate additive loss reserving method for dependent lines of business. *Variance* 3 (1), 131–151.
- Merz, M., Wüthrich, M. V., Hashorva, E., 2013. Dependence modelling in multivariate claims run-off triangles. *Annals of Actuarial Science* 7 (1), 3–25.
- Meyers, G., 2009. Bayesian analysis with the Metropolis-Hastings algorithm. *The Actuarial Review* 36 (4), 32–33.
- Meyers, G. G., Shi, P., 2011. The retrospective testing of stochastic loss reserve models. *Casualty Actuarial Society E-Forum* (Summer), 1–37.
- Peters, G. W., Shevchenko, P., Wüthrich, M. V., 2009. Model uncertainty in claims reserving within Tweedie’s compound Poisson models. *ASTIN Bulletin* 39 (1), 1–33.
- Renshaw, A. E., Verrall, R. J., 1998. A stochastic model underlying the chain-ladder technique. *British Actuarial Journal* 4 (4), 903–923.
- Salzmann, R., Wüthrich, M. V., 2012. Modeling accounting year dependence in runoff triangles. *European Actuarial Journal* 2 (2), 227–242.
- Salzmann, R., Wüthrich, M. V., Merz, M., 2012. Higher moments of the claims development result in general insurance. *ASTIN Bulletin* 42 (1), 355–384.
- Schmidt, K. D., 2006. Optimal and additive loss reserving for dependent lines of business. *Casualty Actuarial Society Forum* (Fall), 319–351.
- Shi, P., 2014. A copula regression for modeling multivariate loss triangles and quantifying reserving variability. *ASTIN Bulletin* 44 (1), 85–102.
- Shi, P., Basu, S., Meyers, G. G., 2012. A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal* 16 (1), 29–51.
- Shi, P., Frees, E. W., 2011. Dependent loss reserving using copulas. *ASTIN Bulletin* 41 (2), 449–486.
- Taylor, G., 2000. *Loss reserving: An actuarial perspective*. Boston: Kluwer Academic Publishers.
- Taylor, G., 2009. The chain ladder and Tweedie distributed claims data. *Variance* 3 (1), 96–104.
- Taylor, G., 2015. Bayesian chain ladder models. *ASTIN Bulletin* 45 (1), 75–99.
- Tsionas, E. G., 2004. Bayesian inference for multivariate gamma distributions. *Statistics and Computing* 14 (3), 223–233.
- Verrall, R., Hössjer, O., Björkwall, S., 2012. Modelling claims run-off with reversible jump Markov Chain Monte Carlo methods. *ASTIN Bulletin* 42 (1), 35–58.
- Vihola, M., 2012. Robust adaptive Metropolis algorithm with coerced acceptance rate. *Statistics and Computing* 22 (5), 997–1008.
- Wüthrich, M. V., 2003. Claim reserving using Tweedie’s compound Poisson model. *ASTIN Bulletin* 33 (2), 331–346.
- Wüthrich, M. V., Merz, M., 2008. *Stochastic claims reserving methods in insurance*. Chichester: John Wiley & Sons.
- Xia, M., Scollnik, D. P., 2015. A Bayesian mixture model accounting for zeros and negatives in the loss triangle. *International Journal of Statistics and Probability* 4 (2), 10.
- Zhang, Y., Dukic, V., 2013. Predicting multivariate insurance loss payments under the Bayesian copula framework. *The Journal of Risk and Insurance* 80 (4), 891–919.
- Zhang, Y., Dukic, V., Guszczka, J., 2012. A Bayesian non-linear model for forecasting insurance loss payments. *Journal of the Royal Statistical Society* 175 (2), 637–656.