REVIEW

On unbalanced data and common shock models in stochastic loss reserving

Benjamin Avanzi[1], Greg Taylor[2], Phuong Anh Vu[3*] and Bernard Wong[2]

[1]Centre for Actuarial Studies, Department of Economics, University of Melbourne, Melbourne, VIC 3010, Australia;
[2]School of Risk and Actuarial Studies, UNSW Business School, UNSW Sydney, NSW 2052, Australia and [3]Taylor Fry, Level 22, 45 Clarence St, Sydney, NSW 2000, Australia
*Corresponding author. E-mail: anh.vu@taylorfry.com.au

Abstract

Introducing common shocks is a popular dependence modelling approach, with some recent applications in loss reserving. The main advantage of this approach is the ability to capture structural dependence coming from known relationships. In addition, it helps with the parsimonious construction of correlation matrices of large dimensions. However, complications arise in the presence of "unbalanced data", that is, when (expected) magnitude of observations over a single triangle, or between triangles, can vary substantially. Specifically, if a single common shock is applied to all of these cells, it can contribute insignificantly to the larger values and/or swamp the smaller ones, unless careful adjustments are made. This problem is further complicated in applications involving negative claim amounts. In this paper, we address this problem in the loss reserving context using a common shock Tweedie approach for unbalanced data. We show that the solution not only provides a much better balance of the common shock proportions relative to the unbalanced data, but it is also parsimonious. Finally, the common shock Tweedie model also provides distributional tractability.

1. Introduction

Outstanding claims reserves are typically some of the most critical components in the financial statement of a non-life insurer (Abdallah *et al*., 2015; Heberle & Thomas, 2016; Saluz & Gisler, 2014). When estimating reserves, the insurer often has to provide the central estimate as well as a risk margin to accommodate for the stochastic nature of outstanding claims. The estimation of reserving variability is also required by many regulators (Gismondi *et al*., 2012). For example, the Australian Prudential Regulation Authority (APRA) requires insurers to provide a risk margin calculated as the larger of a half of one standard deviation, and the difference between the 75th percentile and the expected value of the total outstanding claims distribution. The 99.5th percentile of the distribution of total outstanding claims is also an input in the calculation of risk-based capital for solvency purposes in many regulatory frameworks, for example, Solvency II in Europe and APRA's Prudential Standards in Australia. This was one of the motivations for the development of stochastic reserving methodologies since the early 1980s. For general references on reserving, one can refer to Taylor (2000) and Wüthrich & Merz (2008). A recent strand of the literature focuses on the modelling of individual claims (see, for example, Avanzi *et al*., 2016;

Pinheiro *et al.*, 2003; Wüthrich, 2018; Zhao *et al.*, 2009. However, the focus of this paper is on the modelling of traditional aggregate data in the form of loss triangles.

Non-life insurers typically operate in multiple lines or segments and are required by regulators to estimate loss reserves and risk capital on an aggregate level. Different business lines within an insurer's operation often lack a comonotonic dependence structure. This allows the insurer to enjoy diversification benefits in the calculation of loss reserves and risk capital for their consolidated operation (Avanzi *et al.* 2016*a*). It is hence essential to develop an accurate approach to the modelling of outstanding losses while allowing for dependencies (Côté *et al.*, 2016; Shi *et al.*, 2012). This allows the insurer not only to accurately assess their performance but also to hold an appropriate amount of reserves and capital to optimise its internal use, while satisfying regulatory requirements (Ajne, 1994; Avanzi *et al.*, 2018).

Various multivariate approaches have been developed for stochastic loss reserving, which take into account the dependency across business lines or segments. Some well-known non-parametric approaches include multivariate chain ladder frameworks in Braun (2004), Schmidt (2006), Merz & Wüthrich (2007), Zhang (2010) and the multivariate additive loss reserving framework in Hess *et al.* (2006), Merz & Wüthrich (2009). These approaches are non-parametric and do not utilise any distributional assumptions. They also focus on specific cellwise dependence (i.e. the dependence between cells that are in the same position) across loss triangles. Alternatively, parametric approaches utilising distributional assumptions can be used, see, for example, Shi & Frees (2011), Zhang & Dukic (2013), De Jong (2012), Abdallah *et al.* (2015), Shi (2014).

In this paper, we focus on the common shock approach to dependence modelling. Common shock approaches use common random factors to capture drivers of dependence across related variables. As a result, these drivers can be identified, as well as monitored if needed. The transparent dependence structures in common shock models can then be interpreted more easily. This is indeed one of the four desirable properties of multivariate distributions considered in Joe (1997), Chapter 4 which include

- interpretability;
- closure under the taking of marginals, meaning that the multivariate marginals belong to the same family (this is important if, in modelling, we need to first choose appropriate univariate marginals, then bivariate and sequentially to higher-order marginals);
- flexible and wide range of dependence;
- density and cumulative distribution function in closed-form (if not, they are computationally feasible to work with).

Furthermore, the construction of correlation matrices can be facilitated. Correlation matrices are tools extensively used by practitioners to specify dependence in the aggregation of outstanding claims liabilities or risk-based capital. Explicit dependence structures captured using common shock approaches allow correlation matrices to be specified in a more disciplined and parsimonious manner (see, e.g., Avanzi, Taylor & Wong, 2018).

Common shock approaches have been used to good effect. They are typically used to capture structural dependence, that is, "structural co-movements that are due to known relationships which can be accounted for in a modelling exercise" (International Actuarial Association, 2004). De Jong (2006) introduced three different models to capture dependence across development periods, accident periods and calendar periods, respectively. Calendar period dependence is captured using common shock variables in the multivariate log-normal model of Shi *et al.* (2012). A common shock Tweedie framework was developed in Avanzi *et al.* (2016*b*) to capture cellwise dependence across business lines. It is worth noting that these models are static models which assume a single development pattern for all accident years through the use of fixed effects. A recent use of common shock approach in evolutionary reserving models which allow claims development pattern to evolve can be found in Avanzi *et al.* (2020). There are also various applications of common shock models outside of the reserving literature, including mortality modelling (Alai

et al., 2013, 2016), capital modelling (Furman & Landsman, 2010) and claim counts modelling (Meyers, 2007).

Despite the benefits mentioned above, complications can arise in the application of common shock approaches to loss triangle data. This is due to the "unbalanced" feature of data where expected magnitude of observations within a loss triangle as well as across triangles varies substantially. This feature represents the typical claim experience where the level of claim activity reaches a peak in early years and then dies out as the development lag increases. The "unbalanced-ness" can also be observed in loss data that consist of multiple business lines. In particular, the speed of claims development can vary across business lines where some lines are longer tailed than others. As a result, the magnitude of claim observations in the same accident year and development year can vary across loss triangles. Because of this feature, we say that loss reserving data are an example of "unbalanced data". If a single common shock is applied to these observations that are of different magnitudes, it can contribute insignificantly to the larger ones and/or swamp the smaller ones, unless careful adjustments are made. It is the aim of this paper to study and address this problem.

While this paper aims to examine the challenges for common shock models and propose a solution to address these challenges, a focus of the solution is placed on the Tweedie family of distributions. This is motivated by its popularity. This family is a major subclass of the Exponential Dispersion Family (EDF) consisting of symmetric and non-symmetric, light-tailed and heavy-tailed distributions (Alai *et al.*, 2015; Jørgensen, 1997). Various members of it have been frequently used in the loss reserving literature, see, for example, Alai & Wüthrich (2009), Boucher & Davidov (2011), England & Verrall (2002), Peters *et al.* (2009), Renshaw & Verrall (1998), Taylor (2009), Taylor (2015), Wüthrich (2003), Zhang *et al.* (2012). Avanzi *et al.* (2016a) developed a common shock Tweedie framework for reserving to allow for dependence across business line, while utilising the flexibility of this family of distribution. The solution proposed in this paper will be illustrated using this framework.

Another feature that is occasionally observed in loss triangles is negative claim amounts. These are due to various reasons, for example, salvage recoveries, or payment from third parties. Many commonly used distributions such as gamma distributions and log-normal distributions are unable to handle this feature due to their lack of support for negative values. A remarkably small area of literature has been devoted for the treatment of negative payments in a single business line. The existing methods include a three-parameter-log-normal model in De Alba (2006) and a mixture model in Kunkler (2006). In the development of the new approach for unbalanced data, we will also consider a treatment for negative claims.

The organisation of this paper is as follows: section 2 investigates the issue of unbalanced data for common shock models. A common shock Tweedie approach to unbalanced data is introduced in section 3. Simulation illustrations are provided in section 4, including an illustration using a portfolio of triangles with different tail lengths, and a comparison of the performances of the original common shock Tweedie approach and the modified Tweedie approach with treatment for unbalanced data. An illustration using real data is provided in section 5 and section 6 concludes the paper.

2. Unbalanced Feature of Reserving Data and Its Challenges to Common Shock Models

In this section, we examine the unbalanced feature of loss reserving data in detail. The general common shock framework developed in Avanzi *et al.* (2018) is then described. Challenges that arise in applying common shock models to reserving data due to its unbalanced feature are then discussed.

2.1 Unbalanced feature of reserving data

As described in section 1, loss reserving data typically exhibit unbalanced nature. We consider for illustration a real data set from a Canadian insurer collected from 2003 to 2012 (denoted by years
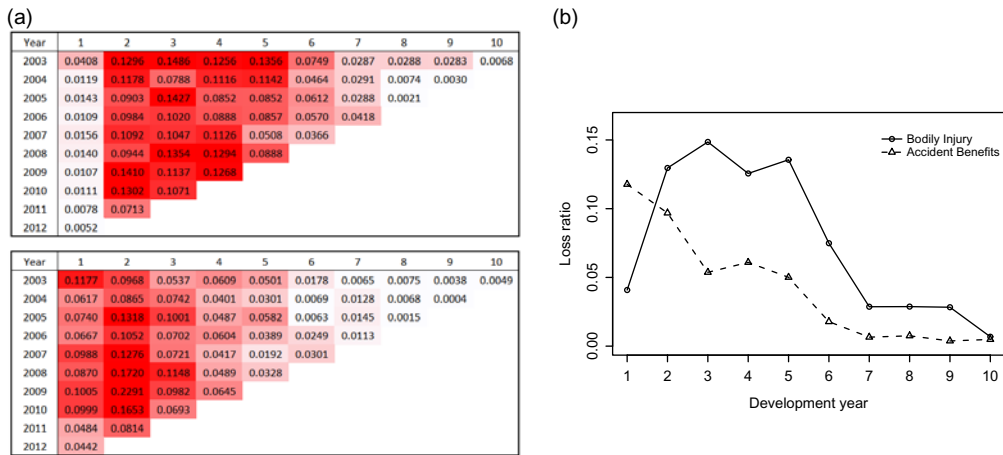
(a)

Year	1	2	3	4	5	6	7	8	9	10
2003	0.0408	0.1296	0.1486	0.1250	0.1356	0.0749	0.0287	0.0288	0.0283	0.0068
2004	0.0119	0.1178	0.0788	0.1116	0.1142	0.0464	0.0291	0.0074	0.0030	
2005	0.0143	0.0903	0.1427	0.0852	0.0852	0.0612	0.0288	0.0021		
2006	0.0109	0.0984	0.1020	0.0888	0.0857	0.0570	0.0418			
2007	0.0156	0.1092	0.1047	0.1126	0.0508	0.0366				
2008	0.0140	0.0944	0.1354	0.1294	0.0888					
2009	0.0107	0.1410	0.1137	0.1268						
2010	0.0111	0.1302	0.1071							
2011	0.0078	0.0713								
2012	0.0052									

Year	1	2	3	4	5	6	7	8	9	10
2003	0.1177	0.0968	0.0537	0.0609	0.0501	0.0178	0.0065	0.0075	0.0038	0.0049
2004	0.0617	0.0865	0.0742	0.0401	0.0301	0.0069	0.0128	0.0068	0.0004	
2005	0.0740	0.1318	0.1001	0.0487	0.0582	0.0063	0.0145	0.0015		
2006	0.0667	0.1052	0.0702	0.0604	0.0389	0.0249	0.0113			
2007	0.0988	0.1276	0.0721	0.0417	0.0192	0.0301				
2008	0.0870	0.1720	0.1148	0.0489	0.0328					
2009	0.1005	0.2291	0.0982	0.0645						
2010	0.0999	0.1653	0.0693							
2011	0.0484	0.0814								
2012	0.0442									

(b)

Figure 1. Loss ratios from Bodily Injury line and Accident Benefits line (from a Canadian insurer). (a) Heat maps of Bodily Injury line (top) and Accident Benefit line (bottom). (b) Plot of loss ratios for accident year 2003.

1–10). This data set is used for illustration in Côté *et al.* (2016) and is provided in Tables C.1 and C.2 in Appendix C. The two lines of business (LOBs) used for illustration are Bodily Injury line and Accident Benefits (excluding Disability Income).

Figure 1 provides heat maps of incremental loss ratios on the left and a plot of incremental loss ratios for accident year 2003 from the two LOBs on the right. For a given accident year, the loss ratio increment for development year j is defined as the ratio of incremental claim payments in that development year to the earned premium for the accident year. Within a single loss triangle, one can observe a quite significant variation in claim observations across development years for any particular accident year. As shown in the heat map in Figure 1, the claim activity for the Bodily Injury line is low in development year 0, then peaks in the next few years and dies out after the peak. This typical pattern is also shown in the plot of loss ratio on the right-hand side of Figure 1 for accident year 2003. For the Accident Benefits line, the claim activity is the highest in development year 0 or 1, then drops quickly as we approach later development year. This pattern is also shown in the plot of loss ratios for accident year 2003 of the Accident Benefit line. The plot of ratios on the right-hand side of Figure 1 also indicates the difference in development patterns for two different business lines. We can say that the Accident Benefits line is shorter-tailed than the Bodily Injury line. A variation can be observed across claim observations that come from the same accident year and the same development year, simply due to different claim development patterns across these business lines. This is in addition to the variation between loss values in different development lags and from different loss triangles, such as cells in development year 1 from the Bodily Injury line and cells in development year 10 from the Accident Benefit line.

Overall, Figure 1 shows a large variation across claim observations in a loss reserving data set. Within a single loss triangle, there is variation across development years due to the development pattern of claims over time. Different claim development patterns can also result in variation between observations across loss triangles. Typically, one often does not expect dependence between lines that have different tail lengths, for example, an Auto Property Damage line is often independent of a Workers Compensation line. However, lines with different tail lengths can still have some association. One of such examples is a portfolio of Accident Benefits line and the Bodily Injury line in the above illustration.

With the variations between claim observations within and across triangles, we refer to loss reserving data as unbalanced data. This data feature creates a number of challenges in applying common shock models to reserving data, which will be discussed in the remainder of this section.

For generality and completeness, the focus is placed on the unbalanced feature of data consisting of multiple LOBs.

2.2 General common shock framework

Consider N loss triangles of claim cells $Y_{i,j}^{(n)}$. The notation $Y_{i,j}^{(n)}$ can represent incremental claim payments or counts. We have the indices i ($i = 1, ..., I$) representing the accident period, j ($j = 0, ..., J$) representing the development period and n ($n = 1, ..., N$) representing the business line. It also follows that the claims $Y_{i,j}^{(n)}$ belong to the calendar period $t = i + j - 1$, ($t = 1, ..., T$).

Let $\mathcal{S}^{(n)} = \{\mathcal{S}_s^{(n)}; s = 1, ..., S\}$ be a partition of the set of all claims $Y_{i,j}^{(n)}$ from business line n. Also assume that all partitions are the same for different lines n for simplicity. Denote by $\boldsymbol{\pi}_{(i,j)} = s$ a unique mapping of claim $Y_{i,j}^{(n)}$ to a set $\mathcal{S}_s^{(n)}$ in the partition. For example, the partition $\mathcal{S}^{(n)} = \{\mathcal{S}_s^{(n)}; s = 1, ..., I\}$, where $\mathcal{S}_s^{(n)} = \{Y_{s,j}^{(n)}; j = 1, ..., J\}$ represents a partition of claims by accident year. The selection of the partition $\mathcal{S}^{(n)}$ is very flexible and can be specified for different types of dependence.

Many multivariate models with different types of dependence can be generalised by the common shock framework in Avanzi *et al.* (2018) with

$$Y_{i,j}^{(n)} = \kappa_{i,j}^{(n)} W_{\boldsymbol{\pi}_{(i,j)}} + \lambda_{i,j}^{(n)} U_{\boldsymbol{\pi}_{(i,j)}}^{(n)} + Z_{i,j}^{(n)} \tag{1}$$

where $\boldsymbol{\pi}_{(i,j)} = s$ denote the unique mapping of the claim $Y_{i,j}^{(n)}$ to the corresponding subset $\mathcal{S}_s^{(n)}$ in the partition $\mathcal{S}^{(n)}$, and $W_{\boldsymbol{\pi}_{(i,j)}}$, $U_{\boldsymbol{\pi}_{(i,j)}}^{(n)}$, and $Z_{i,j}^{(n)}$ are independent stochastic variates. The common shock $W_{\boldsymbol{\pi}_{(i,j)}}$ introduces dependence across all business lines $n =, 1..., N$ on claims that belong to the subsets $\mathcal{S}_s^{(n)}$. For example, accident year dependence across lines can be captured using the partition where $\mathcal{S}_s^{(n)} = \{Y_{s,j}^{(n)}; j = 1, ..., J\}$. The other common shock $U_{\boldsymbol{\pi}_{(i,j)}}^{(n)}$ introduces dependence across claims within the set $\mathcal{S}_s^{(n)}$ of business line n only, such as development year dependence with the partition set specification $\mathcal{S}_s^{(n)} = \{Y_{i,s}^{(n)}; i = 1, ..., I\}$. Overall, the flexibility of choice of the subsets $\mathcal{S}_s^{(n)}$ allows different dependence structures to be captured. Lastly, the idiosyncratic component, which is unique to the claim $Y_{i,j}^{(n)}$, is denoted by $Z_{i,j}^{(n)}$. Scaling factors, denoted by $\kappa_{i,j}^{(n)}$, $\lambda_{i,j}^{(n)}$, control the extent to which the set-wide common shock contributes to individual members of the set. In this section, we have wished to preserve the link to the general notation of Avanzi *et al.* (2018) through the use of the notation $\boldsymbol{\pi}_{(i,j)}$. This notation will be simplified in section 2.3 for specific examples.

Remak 2.1. There can be situations where variables $\{Y_{i,j}^{(n)}; \forall i, j; n > 2\}$ are pairwise dependent (i.e. the dependency between each pair of variables is driven by a different source). For example, there can be a portfolio of three LOBs where there are three independent common shocks that drive the dependence between each of the following three pairs, LoB 1 and LoB 2, LoB 2 and LoB 3 and LoB 3 and LoB 1, respectively. In such cases, one can consider having additional common shock variables $W_{\boldsymbol{\pi}_{(i,j)}}$ that capture dependence across lines, for example, $W_{\boldsymbol{\pi}_{(i,j)}}^{(1,2)}$, $W_{\boldsymbol{\pi}_{(i,j)}}^{(2,3)}$, and $W_{\boldsymbol{\pi}_{(i,j)}}^{(3,1)}$ for the above scenario of three LoBs. However, it is worth noting that these will result in more parameters required for the framework.

2.3 Balancing common shock proportions in loss reserving data

As a result of the unbalanced feature of reserving data, a common shock model can create problems in the absence of careful modelling.

Consider a special case of equation (1) for dependence within a business line (i.e. $W_{\pi_{(i,j)}} = 0$). Further specify accident period dependence (i.e. $\pi_{(i,j)}^{(n)} = p$ for the mapping of subsets in the partition where $\mathcal{S}_s^{(n)} = \{Y_{s,j}^{(n)}; j = 1, ..., J\}$). This allows us to simplify $U_{\pi_{(i,j)}}^{(n)} = X_i^{(n)}$. Hence, the general framework is reduced to

$$Y_{i,j}^{(n)} = \lambda_{i,j}^{(n)} X_i^{(n)} + Z_{i,j}^{(n)} \tag{2}$$

Consequently, the proportionate contribution of the common shock to the expected value of the total observation is

$$\frac{\lambda_{i,j}^{(n)} \mathrm{E}\left[X_i^{(n)}\right]}{\lambda_{i,j}^{(n)} \mathrm{E}\left[X_i^{(n)}\right] + \mathrm{E}\left[Z_{i,j}^{(n)}\right]} \tag{3}$$

If the scaling factor is removed, that is, $\lambda_{i,j}^{(n)} = 1$, this proportion has an inverse relationship with the mean of the idiosyncratic component $\mathrm{E}\left[Z_{i,j}^{(n)}\right]$. As a result, in a set of loss cells in a triangle that are dependent and share a common shock, the cells with large values have a smaller proportion of common shock contribution and vice versa. This is because claims within the same accident period, or within the same calendar period belong to different development periods. As explained in section 2.1, their values can vary significantly due to the variation in claim activity across development periods. This issue can also be observed in the case of calendar period dependence (i.e. $\pi_{(i,j)}^{(n)} = s$ for the mapping of subsets in the partition where $\mathcal{S}_s^{(n)} = \{Y_{i,s-i+1}^{(n)}; i = 1, ..., J\}$).

A similar issue is encountered for a portfolio of dependent business lines with differing tail lengths, such as the two business lines Bodily Injury and Accident Benefits in the illustration in Section 2.1. We consider a special case of equation (1) that allows for dependence between business lines only (i.e. $U_{\pi_{i,j}}^{(n)} = 0$). Further specify cellwise dependence (i.e. partition mapping where $\mathcal{S}_{i,j}^{(n)} = \{Y_{i,j}^{(n)}\}$). This allows us to simplify $W_{\pi_{(i,j)}} = V_{i,j}$. The contribution of the common shock to the total expected observation is then given by

$$\frac{\kappa_{i,j}^{(n)} \mathrm{E}\left[V_{i,j}\right]}{\kappa_{i,j}^{(n)} \mathrm{E}\left[V_{i,j}\right] + \mathrm{E}\left[Z_{i,j}^{(n)}\right]} \tag{4}$$

If the scaling factor is removed, that is, $\kappa_{i,j}^{(n)}$, this proportion also has an inverse relationship with the mean of the idiosyncratic component $\mathrm{E}\left[Z_{i,j}^{(n)}\right]$. As explained in section 2.1, values of claims in a portfolio of multiple triangles can vary in two main ways: across development years within a loss triangle, and across loss triangles. As a result, the proportion of common shock varies within and across loss triangles, wherein loss cells with larger values have smaller common shock contributions. In the case of pairwise dependence considered above, the disproportion is typically a result of varying tail lengths across business lines. However, it is worth noting that unbalanced common shock proportions can also be typically observed for accident year dependence, or calendar dependence across business lines from the same cause.

We consider the case of accident year dependence across the two triangles illustrated in section 2.1 (i.e. partition mapping $\mathcal{S}_s = \{Y_{s,j}^{(n)}; j = 1, ..., J; n = 1, ..., N\}$, and we can simplify $W_{\pi_{(i,j)}} = V_i$). For illustration, the mean of the common shock $E[V_i]$ is set to 5% of the loss ratios in the first development year of each accident year in the Bodily Injury line. The contributions of common shock are shown in Figure 2 assuming no scaling terms. With accident year dependence across business lines, claims within the same accident year share the same common shock. These include claims from different development years within and across loss triangles. Because their values vary
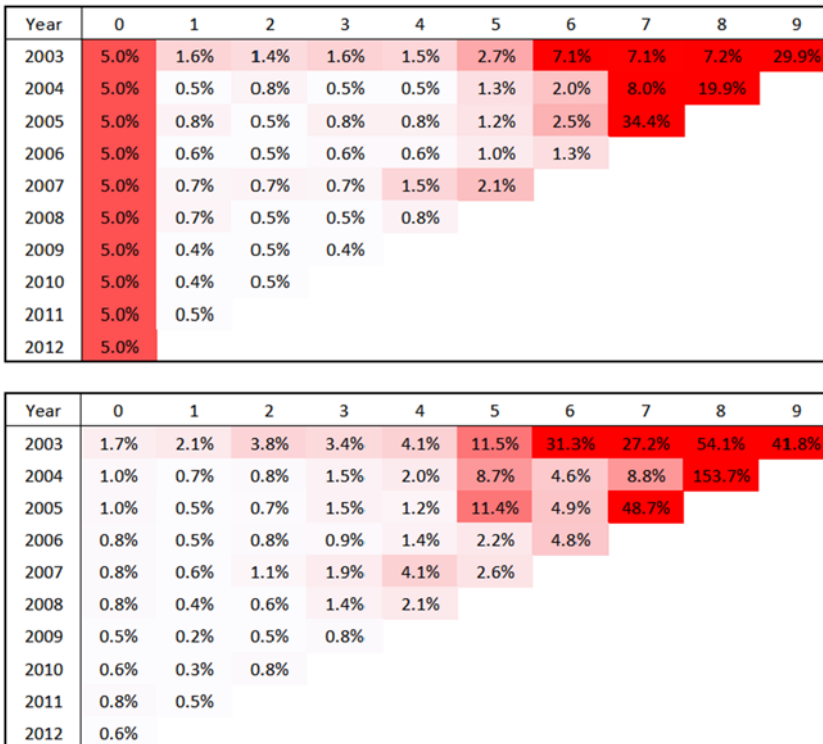
Year	0	1	2	3	4	5	6	7	8	9
2003	5.0%	1.6%	1.4%	1.6%	1.5%	2.7%	7.1%	7.1%	7.2%	29.9%
2004	5.0%	0.5%	0.8%	0.5%	0.5%	1.3%	2.0%	8.0%	19.9%	
2005	5.0%	0.8%	0.5%	0.8%	0.8%	1.2%	2.5%	34.4%		
2006	5.0%	0.6%	0.5%	0.6%	0.6%	1.0%	1.3%			
2007	5.0%	0.7%	0.7%	0.7%	1.5%	2.1%				
2008	5.0%	0.7%	0.5%	0.5%	0.8%					
2009	5.0%	0.4%	0.5%	0.4%						
2010	5.0%	0.4%	0.5%							
2011	5.0%	0.5%								
2012	5.0%									

Year	0	1	2	3	4	5	6	7	8	9
2003	1.7%	2.1%	3.8%	3.4%	4.1%	11.5%	31.3%	27.2%	54.1%	41.8%
2004	1.0%	0.7%	0.8%	1.5%	2.0%	8.7%	4.6%	8.8%	153.7%	
2005	1.0%	0.5%	0.7%	1.5%	1.2%	11.4%	4.9%	48.7%		
2006	0.8%	0.5%	0.8%	0.9%	1.4%	2.2%	4.8%			
2007	0.8%	0.6%	1.1%	1.9%	4.1%	2.6%				
2008	0.8%	0.4%	0.6%	1.4%	2.1%					
2009	0.5%	0.2%	0.5%	0.8%						
2010	0.6%	0.3%	0.8%							
2011	0.8%	0.5%								
2012	0.6%									

Figure 2. Heat maps of common shock contributions in Bodily Injury line (top) and Accident Benefit line (bottom) without using scaling terms.

due to different claim activities within and across lines, their common shock proportions also vary. Specifically, common shock proportions are significantly smaller in areas with high claim activity, and larger in areas with low claim activity, as shown in Figure 2.

In general, quite significant variations in common shock proportions can be observed within and across segments in the absence of careful modelling as a result of the unbalanced nature of loss reserving data. One may wish to confine the relation of the common shock to total observations over the entire range of the triangles.

2.4 Maintaining model parsimony

The most straightforward solution to the balancing common shock proportions within and across triangles is to have cell-specific scaling factors $\kappa_{i,j}^{(n)}$, $\lambda_{i,j}^{(n)}$ to adjust the common shock effects for each total observation $Y_{i,j}^{(n)}$. However, this implies that $2IJN$ new parameters are required for the entire range of triangles of observed data and outstanding claims to be predicted. Given that the variation in claim observations typically occurs across development periods, one may simplify the scaling factors to be column-specific $\kappa_{i,j}^{(n)} = \kappa_j^{(n)}$, $\lambda_{i,j}^{(n)} = \lambda_j^{(n)}$. However, this still results in $2JN$ new parameters.

Loss triangle data typically have a small sample size. While the presence of scaling factors can mitigate the impact of the unbalanced nature of reserving data, it also adds many more parameters to the model. If scaling factors are not chosen carefully, it may result in overfitting and the number of parameters to be estimated can even exceed the number of observations.

2.5 Maintaining distributional tractability

On some occasions, parameters $\lambda_{i,j}^{(n)}$ and $\kappa_{i,j}^{(n)}$ need to be specified such that the total observation $Y_{i,j}^{(n)}$ follows a specific distribution (Avanzi *et al.*, 2018). This is referred to as distributional tractability, or closure under the taking of marginals, which is considered in Joe (1997, Chapter 4) to be one of the four desirable properties of a multivariate model (see also section 1).

Consider as an example the common shock Tweedie framework in Avanzi *et al.* (2016b). This framework is developed for cellwise dependence across business lines (i.e. $\mathcal{S}_{i,j}^{(n)} = \{Y_{i,j}^{(n)}\}$, $U_{i,j}^{(n)} = 0$). Fitting this into the general common shock structure in equation (1) and simplifying $W_{\boldsymbol{\pi}_{(i,j)}} = V_{i,j}$, we have

$$Y_{i,j}^{(n)} = \kappa_{i,j}^{(n)} V_{i,j} + Z_{i,j}^{(n)} \tag{5}$$

where the two components $V_{i,j}$, $Z_{i,j}^{(n)}$ are assumed to be independent and have Tweedie distributions

$$V_{i,j} \sim \text{Tweedie}_p(\alpha, \beta) \tag{6}$$

$$Z_{i,j}^{(n)} \sim \text{Tweedie}_p(\eta_i^{(n)} v_j^{(n)}, \gamma^{(n)}) \tag{7}$$

Parameter p is the power parameter which specifies a member of the Tweedie family, for example, $p = 1$ corresponds to a Poisson distribution. The representation of Tweedie distributions used is the reproductive representation (Jørgensen, 1997, Chapter 4). This representation specifies a Tweedie random variable using a location (or mean) parameter and a dispersion parameter. In the above model specification, parameters α and $\eta_i^{(n)} v_j^{(n)}$ are the location parameters, and parameters β and $\gamma^{(n)}$ are the dispersion parameters. The reproductive representation has a distinctive property, wherein the weighted average of independent Tweedie variables with the same power parameter p and the same location parameter is also a Tweedie variable with the same power and location parameters. The weighting factors are determined using dispersion parameters of the component variables in the weighted average.

It then follows that the mean and variance of the two components $V_{i,j}$, $Z_{i,j}^{(n)}$ are

$$\text{E}\left[V_{i,j}\right] = \alpha, \qquad\qquad \text{Var}\left[V_{i,j}\right] = \beta\alpha^p \tag{8}$$

$$\text{E}\left[Z_{i,j}^{(n)}\right] = \eta_i^{(n)} v_j^{(n)}, \qquad\qquad \text{Var}\left[Z_{i,j}^{(n)}\right] = \gamma^{(n)}(\eta_i^{(n)} v_j^{(n)})^p \tag{9}$$

As stated in Remark 2.2 of Avanzi *et al.* (2016b), the most simple parametrisation is used for the common shock component $V_{i,j}$ with parameters α and β.

As mentioned earlier in this section, it can be desirable to maintain distributional tractability, or closure under the taking of marginals for ease of interpretation. It follows from the form of closure under addition of the Tweedie family of distributions, as proven in Jørgensen (1997, Chapter 3), that a specific choice of $\kappa_{i,j}^{(n)}$ is required to ensure that $Y_{i,j}^{(n)}$ also has a Tweedie distribution. This choice is

$$\kappa_{i,j}^{(n)} = \left(\frac{\alpha}{\eta_i^{(n)} v_j^{(n)}}\right)^{1-p} \frac{\gamma^{(n)}}{\beta} \tag{10}$$

The mean expression is given by

$$\text{E}\left[Y_{i,j}^{(n)}\right] = \left(\frac{\alpha}{\eta_i^{(n)} v_j^{(n)}}\right)^{2-p} \frac{\gamma^{(n)}}{\beta} \eta_i^{(n)} v_j^{(n)} + \eta_i^{(n)} v_j^{(n)} \tag{11}$$

where the first term in the summation is the contribution from the common shock and the second term is the contribution from the idiosyncratic component. The expected contribution of the common shock to the total expected observation is

$$
\frac{\left(\dfrac{\alpha}{\eta_i^{(n)} v_j^{(n)}}\right)^{2-p} \dfrac{\gamma^{(n)}}{\beta}}{\left(\dfrac{\alpha}{\eta_i^{(n)} v_j^{(n)}}\right)^{2-p} \dfrac{\gamma^{(n)}}{\beta} + 1} \tag{12}
$$

The following observation can be made on the effect of the power parameter p:

- If $p < 2$: The above ratio increases as $v_j^{(n)}$ decreases. As a result, the proportion of common shock is understated in early development periods and overstated in late development periods (Avanzi *et al.*, 2018). In a portfolio of segments with varying tail lengths, the larger the discrepancy between the tail lengths (i.e. between $v_j^{(n)}$ and $v_j^{(m)}$), the larger the variation in the common shock contributions. The behaviour of the above ratio has been examined with respect to development factor $v_j^{(n)}$ in particular because variation within and across lines of business is mainly driven by the development pattern of claims as explained in section 2.1. As a result, one would expect the development factors to vary the most.
- If $p > 2$: The opposite observation is made for the relationship between the above ratio and $v_j^{(n)}$ (i.e. the above ratio decreases as $v_j^{(n)}$ decreases).
- If $p = 2$: In this special case, the common shock contribution is simplified to

$$
\frac{\dfrac{\gamma^{(n)}}{\beta}}{\dfrac{\gamma^{(n)}}{\beta} + 1} \tag{13}
$$

which is now independent of accident and development periods. Consequently, the common shock contributes proportionately to the total observations over the entire range of the triangles. It is also worth emphasising that specifying $p = 2$ gives the multivariate gamma case of the multivariate Tweedie framework.

The above analyses and examples show that the choices of scaling factors $\kappa_{i,j}^{(n)}$ and $\lambda_{i,j}^{(n)}$ are subject to many constraints. To accurately capture the dependence structure, these parameters are required to balance the common shock proportions within all claim observations over the entire range of the triangles. However, this can result in overfitting, which can be a critical issue in loss reserving due to small sample size data. Furthermore, the specification of these parameters may need to be restricted in some cases for the purpose of preserving distributional tractability. It is then the aim of this paper to find a solution that compromises between these conflicting issues with a specific application on the common shock Tweedie approach in Avanzi *et al.* (2016*b*).

3. A Common Shock Tweedie Approach to Unbalanced Data

In this section, we propose a solution that compromises between conflicting challenges encountered by common shock models when they are applied to reserving data due to the unbalanced feature of the data. The focus of this development is on a common shock Tweedie approach to unbalanced data. The estimation method for this approach is also given.

The multivariate Tweedie framework described in section 2.5 is a typical example of an application of the common shock approach in stochastic loss reserving. It is of particular interest due to its various advantages. Developed on the Tweedie family of distributions, it offers flexible choices of marginal density that also include Tweedie's compound Poisson density with the ability to deal with zero data points. The framework can also be generalised to more than two dimensions. In addition, the explicit common shock structure allows the correlation matrix to be obtained in closed form. Moment- and cumulant-generating-functions can also be obtained analytically, enhancing the tractability of the model. Similar to other common shock models, this framework also encounters the issue of unbalanced data. As explained in section 2, the selection of scaling coefficients for the common shock term in this framework is constrained by the need to balance common shock proportions, while maintaining model parsimony and distributional tractability.

3.1 Theoretical framework

Claims are first standardised using a common unit of exposure such as the number of claims, or the total amount of premium collected, to ensure consistency across accident periods and business lines. Recall the specification of the common shock Tweedie model in Avanzi *et al.* (2016*b*) described in section 2.5,

$$Y_{i,j}^{(n)} = \kappa_{i,j}^{(n)} V_{i,j} + Z_{i,j}^{(n)} \tag{14}$$

where

$$V_{i,j} \sim \text{Tweedie}_p(\alpha, \beta) \tag{15}$$

$$Z_{i,j}^{(n)} \sim \text{Tweedie}_p(\eta_i^{(n)} v_j^{(n)}, \gamma^{(n)}) \tag{16}$$

$$\kappa_{i,j}^{(n)} = \left(\frac{\alpha}{\eta_i^{(n)} v_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} \tag{17}$$

Recall that α and $\eta_i^{(n)} v_j^{(n)}$ are location (mean) parameters, and β and $\gamma^{(n)}$ are dispersion parameters of $V_{i,j}$ and $Z_{i,j}^{(n)}$, respectively.

As shown in equation (10), the common shock scaling factor has to be specified in the above form that involves parameters of the common shock $V_{i,j}$ and the idiosyncratic component $Z_{i,j}^{(n)}$. However, due to the unbalanced feature of reserving data with $v_j^{(n)}$ varying across development lag j and business line n, the common shock contributes disproportionately to the total observation $Y_{i,j}^{(n)}$. It is also desirable to maintain model parsimony.

Given the above considerations, we can replace the non-cell-specific parameter α in the scaling factor with column-specific parameter

$$\alpha_j = \tilde{c} \left(\prod_n \text{E} \left[Z_{i,j}^{(n)} \right] \right)^{\frac{1}{N}} = \tilde{c} \left(\prod_n \eta_i^{(n)} v_j^{(n)} \right)^{\frac{1}{N}} \tag{18}$$

$$\approx c \sqrt[N]{v_j^{(1)} ... v_j^{(N)}} \tag{19}$$

The parameter α_j is also the location parameter of the common shock $V_{i,j}$. As a result, we approximately have

$$V_{i,j} \sim \text{Tweedie}_p(\alpha_j, \beta) = \text{Tweedie}_p \left(c \sqrt[N]{v_j^{(1)} ... v_j^{(N)}}, \beta \right) \tag{20}$$

Essentially, the common shock parameter α_j is proportional to the geometric average of idiosyncratic components of claims which share the same common shock component. In this case, these are claims in the same accident period and development period as the framework is used to capture cellwise dependence. This geometric average can then be simplified by removing accident period factors because we can reasonably expect limited variation across accident periods as a result of claims standardisation, assuming no significant changes occur across accident periods.

The above specification of scaling factor aims to balance the impact of unbalanced feature in loss reserving data which are mainly introduced by variations in development factors $v_j^{(n)}$. Using this specification, the common shock proportion is given by

$$
\frac{\left(\dfrac{c\sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)}v_j^{(n)}}\right)^{2-p}\dfrac{\gamma^{(n)}}{\beta}}{\left(\dfrac{c\sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)}v_j^{(n)}}\right)^{2-p}\dfrac{\gamma^{(n)}}{\beta}+1}
\tag{21}
$$

This does not provide a complete balance of common shock proportions because the effect of $v_j^{(n)}$ is reduced by a factor $\sqrt[N]{v_j^{(n)}}$. However, it still provides quite a significant improvement over the original framework. This will be demonstrated in the simulation illustration in section 4. This specification can also preserve distributional tractability of the framework. In addition, model parsimony is retained as the total number of parameters in the framework is unchanged. This can be considered an effective solution given the three constraints discussed in section 2.

In addition to the above treatment for unbalanced data, we also introduce a treatment for negative claims

$$
Y_{i,j}^{(n)}+\xi^{(n)}=\left(\frac{c\sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)}v_j^{(n)}}\right)^{1-p}\frac{\gamma^{(n)}}{\beta}V_{i,j}+Z_{i,j}^{(n)}
\tag{22}
$$

where a translation factor is used and defined such that

$$
\xi^{(n)}=\begin{cases}0 & \text{if }\min\{Y_{i,j}^{(n)},\ \forall i,\ j\}\geq0,\\ \geq-\min\{Y_{i,j}^{(n)}\} & \text{if }\min\{Y_{i,j}^{(n)},\ \forall i,\ j\}<0\end{cases}
\tag{23}
$$

The translation is only needed for a loss triangle if it contains at least one negative value and it must be large enough to offset the smallest negative value. It is worth emphasising that in this case, while its lower bound is deterministic, the actual value of $\xi^{(n)}$ still has to be estimated. The generalisation of this treatment to the general common shock framework in Avanzi *et al.* (2018) is straightforward.

Following from the above specification, the marginal density is then given by

$$
Y_{i,j}^{(n)}+\xi^{(n)}\sim\text{Tweedie}_p\left(\eta_i^{(n)}v_j^{(n)}\left[\left(\frac{c\sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)}v_j^{(n)}}\right)^{2-p}\frac{\gamma^{(n)}}{\beta}+1\right],\right.
$$
$$
\left.\gamma^{(n)}\left[\left(\frac{c\sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)}v_j^{(n)}}\right)^{2-p}\frac{\gamma^{(n)}}{\beta}+1\right]^{1-p}\right)
\tag{24}
$$

where the first parameter is the location parameter and also the mean of $Y_{i,j}^{(n)} + \xi^{(n)}$. The second parameter is the dispersion parameter. It follows that the vector of translated claims in the same position across all triangles

$$
{}_\xi Y_{i,j} = \begin{pmatrix} Y_{i,j}^{(1)} + \xi^{(1)} \\ Y_{i,j}^{(2)} + \xi^{(2)} \\ \vdots \\ Y_{i,j}^{(N)} + \xi^{(N)} \end{pmatrix} \tag{25}
$$

has a multivariate Tweedie distribution with the multivariate density

$$
f_{{}_\xi Y_{i,j}} \left(y_{i,j}^{(1)} + \xi^{(1)}, ..., y_{i,j}^{(N)} + \xi^{(N)} \right) =
$$

$$
\int_0^{A_{i,j}} f_{V_{i,j}}(w_{i,j}) \prod_{n=1}^{N} f_{Z_{i,j}^{(n)}} \left(y_{i,j}^{(n)} + \xi^{(n)} - \left(\frac{c \sqrt[N]{v_j^{(1)}...v_j^{(N)}}}{\eta_i^{(n)} v_j^{(n)}} \right)^{1-p} \frac{\gamma^{(n)}}{\beta} w_{i,j} \right) dw_{i,j} \tag{26}
$$

where

$$
A_{i,j} = \min \left(\left(\frac{\eta_i^{(1)} v_j^{(1)}}{c \sqrt[N]{v_j^{(1)}...v_j^{(N)}}} \right)^{1-p} \frac{\beta}{\gamma^{(1)}} (y_{i,j}^{(1)} + \xi^{(1)}), ..., \left(\frac{\eta_i^{(N)} v_j^{(N)}}{c \sqrt[N]{v_j^{(1)}...v_j^{(N)}}} \right)^{1-p} \frac{\beta}{\gamma^{(N)}} (y_{i,j}^{(N)} + \xi^{(N)}) \right)
$$

$$\tag{27}$$

and where $f(.)$ is the Tweedie density in reproductive form (see also Jørgensen, 1997, Chapter 4).

3.2 Model estimation with Bayesian inference

Bayesian inference is used for model estimation. Bayesian estimation has gained its popularity in the loss reserving literature due to rapid computing advancements and Markov Chain Monte Carlo (MCMC) methods that allow the calculation of intractable posterior densities to be performed significantly faster (Avanzi *et al.* 2016b; Verrall *et al.*, 2012). In addition, the incorporation of prior densities in the calculation of posterior densities is a natural way to allow for parameter error in modelling (Shi *et al.*, 2012; England *et al.*, 2012). Another aim of using a Bayesian setup is to also estimate the power parameter p and translation parameter $\xi^{(n)}$ with allowance of parameter uncertainty. This is to formalise the estimation of these parameters as they are often estimated heuristically in practice. It is worth emphasising that the Bayesian structure is not integral to our model but serves as a device for estimation.

A two-step procedure is used for estimation, similar to that in Avanzi *et al.* (2016b). The first stage is the estimation of all parameters except c and β of the common shock $V_{i,j}$. This stage, however, gives the estimate of a ratio of these parameters denoted as

$$
\delta = \frac{c^{2-p}}{\beta} \tag{28}
$$

as can observed from equation (20). This is followed by the multivariate stage that estimates c and β conditional on estimates of other parameters from the first stage. The motivation for this procedure comes from properties of the common shock Tweedie framework. Claim observations in the same position across triangles in this framework follow a multivariate Tweedie distribution, and each observation itself also has a marginal Tweedie distribution. In addition, the multivariate density has an integral calculation, as shown in equation (22). This can prolong the estimation of the posterior density, making the tuning and convergence of MCMC much more difficult.

A Bayesian setup requires the specification of the likelihood functions, prior densities and, if posterior densities are not in closed form, computational algorithms used to approximate them. The likelihood functions follow from equation (20) for the first stage and equation (22) for the second stage.

Prior densities are then specified. Prior densities can be chosen to be informative or uninformative. Uninformative priors assign equal possibilities to all values in the feasible set of parameter values, whereas informative priors convey some prior preference for certain values of the parameters. However, the use of informative priors can significantly improve the convergence rate, especially when the parameter dimension is large (Congdon, 2010). Parameter estimates from univariate Tweedie model (Alai & Wüthrich, 2009) can assist in the specification of informative prior densities for parameters $\eta_i^{(n)}$, $\nu_j^{(n)}$ and $\gamma^{(n)}$. A preliminary analysis of the dependence structure can help select informative prior densities for the common shock parameters c and β. Regarding the prior densities for p and $\xi^{(n)}$, some constraints need to be taken into account. In particular, p is not defined in (0,1), and $\xi^{(n)}$ has a lower bound as per its specification in equation (19).

Putting together the likelihood and prior specifications, the posterior density in the first stage is given by

$$f_{\boldsymbol{\Omega}|Y^U}(\boldsymbol{\Omega}|Y^U) \propto \left(\prod_{i,j,n} f_{Y_{i,j}^{(n)}+\xi^{(n)}} \left(y_{i,j}^{(n)} + \xi^{(n)}|\boldsymbol{\Omega} \right) \right) f_p(p) f_{\boldsymbol{\xi}}(\boldsymbol{\xi}) f_{\delta}(\delta) f_{\boldsymbol{\eta}}(\boldsymbol{\eta}) f_{\boldsymbol{\nu}}(\boldsymbol{\nu}) f_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}) \qquad (29)$$

where

$$\boldsymbol{\Omega} = \begin{pmatrix} p \\ \boldsymbol{\xi} \\ \delta \\ \boldsymbol{\eta} \\ \boldsymbol{\nu} \\ \boldsymbol{\gamma} \end{pmatrix}, \; \boldsymbol{\xi} = \begin{pmatrix} \xi^{(1)} \\ \xi^{(2)} \\ \vdots \\ \xi^{(N)} \end{pmatrix}, \; \boldsymbol{\eta}_i = \begin{pmatrix} \eta_i^{(1)} \\ \eta_i^{(2)} \\ \vdots \\ \eta_i^{(N)} \end{pmatrix}, \; \boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \\ \vdots \\ \boldsymbol{\eta}_I \end{pmatrix}, \; \boldsymbol{\nu}_j = \begin{pmatrix} \nu_j^{(1)} \\ \nu_j^{(2)} \\ \vdots \\ \nu_j^{(N)} \end{pmatrix}, \; \boldsymbol{\nu} = \begin{pmatrix} \boldsymbol{\nu}_1 \\ \boldsymbol{\nu}_2 \\ \vdots \\ \boldsymbol{\nu}_J \end{pmatrix}, \; \boldsymbol{\gamma} = \begin{pmatrix} \gamma^{(1)} \\ \gamma^{(2)} \\ \vdots \\ \gamma^{(N)} \end{pmatrix}$$

and where Y^U is a vector of claim observations in the upper claim triangles.

From the model structure in equation (18), we have that all claims $Y_{i,j}^{(n)}$ are independent conditional on common shock. Hence, the joint likelihood can be written as a product of two separate parts: a product of the densities of claims conditional on common shock, and the density of the common shock. In the first stage of the estimation procedure, the likelihood obtained is the first part of the joint likelihood. As also mentioned earlier, this stage provides the estimates of mean parameters $\boldsymbol{\nu}$, $\boldsymbol{\eta}$ and dispersion parameters $\boldsymbol{\gamma}$ of the idiosyncratic variables $Z_{i,j}^{(n)}$, translation parameters $\boldsymbol{\xi}$ and power parameter p. This stage also provides the estimate of δ which is a function of parameters c and β of the common shock $V_{i,j}$.

In the second estimation step, we work with the joint likelihood directly since common shock components are not observed. In this step, the estimation of c and β is carried out conditioning on estimates of other parameters in the first step, including δ which is a function of c and β. The multivariate Tweedie density of $Y_{i,j}$ is used to obtain the likelihood in this estimation. The posterior density in this step is given by

$$f_{c|Y^U,\boldsymbol{\Omega}}(c|Y^U,\boldsymbol{\Omega}) \propto \left(\prod_{i,j} f_{\boldsymbol{\xi} Y_{i,j}} \left(\boldsymbol{\xi} y_{i,j}|c, \boldsymbol{\Omega} \right) \right) f_c(c) \qquad (30)$$

The posterior densities in both stages are not in recognisable forms, hence MCMC algorithms are required for the estimation. The MCMC algorithm used is Metropolis-Hastings, which is a popular class of MCMC algorithms when the posterior distribution is not in a recognisable

form. Random walk Metropolis-Hastings algorithms are used for marginal estimation and multivariate estimation. Proposal densities are chosen (tuned) so that acceptance probabilities are within desirable ranges. The tuning process can be done manually using classical Metropolis-Hastings algorithms. Alternatively, it can be done automatically in adaptive Metropolis-Hastings algorithms using coerced acceptance rates (Haario *et al.*, 2001; Vihola, 2012).

4. Simulation Illustrations

Two illustrations are performed on two data sets. The first illustration, provided in section 4.1, is to assess the accuracy of the estimation procedure. Since true parameter values are known in a simulated data, a comparison of their estimates with their true values gives an indication of the appropriateness of the estimation procedure. The second illustration, provided in section 4.2, is to compare the performance of the common shock Tweedie approach with treatment for unbalanced data and the original common shock Tweedie approach in Avanzi *et al.* (2016*b*). This comparison focuses particularly on the contributions of common shock estimated from the two approaches.

4.1 An illustration with unbalanced data and negative claims

A data set consisting of two business lines, one of which has a negative claim observation, is simulated. The two loss triangles are represented in Tables A.1 and A.2 in Appendix A. These two triangles consist of simulated claim observations. Each observation in the triangles is drawn from the multivariate Tweedie model for unbalanced data represented in section 3. For simplicity, these observations are assumed to have been adjusted for changes in exposure across accident years.

The marginal fitting is first performed. Parameters are first transformed using the log transformation, and uniform prior densities are used. 200,000 simulations are run and 100,000 simulations are discarded as the burn-in period. The sample chain is thinned by accepting every 5th iteration to reduce the serial dependence between iterations. MCMC paths of some parameters are given in Figure A.1. A similar procedure is performed for the multivariate estimate. The estimates of c and β are obtained from this step. Parameter estimates are provided in Table A.3 in Appendix A.

To evaluate the Bayesian inference used for estimation, we compare the true parameter values with 90% confidence intervals obtained from the posterior distributions of these parameters. The results the true values always lie within the corresponding 90% confidence intervals. This indicates the accuracy of the estimation procedure.

We have calibrated the model on the same simulated data set using sub-triangles of dimension 5×5 to assess the robustness of the proposed calibration method. The results show that true values also fall within the 90% confidence intervals of the estimates from this calibration. However, confidence intervals are generally larger than those from the calibration that uses full-size triangles. This is expected due to higher uncertainty in the estimates coming from smaller sample size.

To evaluate the bias in the resulting reserve predictions, forecasts of outstanding claims using our model are compared with the true forecasts as well as forecasts from a multivariate chain ladder model. The true forecasts are calculated as the expected value of outstanding claims using true parameter values. The multivariate chain ladder model used is the model in Pröhl & Schmidt (2005). The results are shown in Table 1.

It can be observed from Table 1 that the true forecasts fall within the 90% confidence intervals of the balanced multivariate Tweedie model forecasts. The forecasts from our model are also closer to the true forecasts than those from the multivariate chain ladder model. However, it is worth noting that the simulated data were generated from the multivariate Tweedie model in this illustration.

We further assess bias in the resulting dependence structure by comparing the true cellwise Pearson correlation coefficients for the outstanding claims and the cellwise Pearson correlation coefficients calculated using the parameter estimates. Residual ratios, defined as ratios of estimated

Table 1. Comparison of outstanding claims forecasts

	Balanced multivariate Tweedie				Multivariate chain ladder	
LoB	Mean	Standard error	90% CI	True forecasts	Mean	Standard error
1	159.38	19.74	(129.64;193.88)	157.56	138.97	16.55
2	598.98	76.15	(484.99;734.18)	563.92	531.31	52.17
Total	758.37	82.88	(633.14; 904.17)	721.48	670.28	61.31

Table 2. Residual ratios of estimated Pearson correlation coefficients to true Pearson correlation coefficients

		Development year								
		1	2	3	4	5	6	7	8	9	10
Accident year	1										
	2										1.07
	3									0.99	0.99
	4								1.02	1.04	1.04
	5							1.00	0.98	0.99	1.00
	6						1.03	1.02	1.01	1.02	1.03
	7					1.02	1.03	1.03	1.01	1.03	1.03
	8				1.01	1.02	1.03	1.03	1.01	1.03	1.03
	9			1.04	1.03	1.04	1.05	1.05	1.03	1.05	1.05
	10		1.00	0.99	0.98	0.99	1.00	1.00	0.98	1.00	1.00

Pearson correlation coefficients to true Pearson correlation coefficients, are provided in Table 2. The ratios are close to 1, indicating that the cellwise dependence in the data is well captured.

We acknowledge that the use of the two-step Bayesian inference does not provide the full picture due to the dependence between the estimated parameters and the reserve being a non-linear function in terms of these parameters. However, this calibration approach was selected due to a number of advantages as mentioned in section 3. These include overcoming the difficulties in dealing with Tweedie densities which are not in tractable form and enhancing computational speed. The above analyses indicate that

– the calibration method can capture the dependence structure well;
– the resulting reserve predictions show no apparent bias and they are in line with the chain ladder predictions.

Therefore, even though we may not get the full picture, the above results give us confidence that this would not have a material impact on the performance of the calibration.

4.2 A comparison of performances of multivariate Tweedie models on unbalanced data

A natural question arises regarding the performance of the multivariate Tweedie approach for unbalanced data compared to the original multivariate Tweedie approach introduced in Avanzi *et al.* (2016*b*). To be able to assess their performances more accurately, this comparison is performed on a simulated data set whose underlying model is known. True common shock contributions are also known, and these serve as the benchmark for the comparison.

To not put any particular framework at a disadvantage, the synthetic data used for this illustration are simulated from a mixture of models. We deliberately select a (extreme) data set to which neither of the frameworks is properly adapted. In particular, two loss triangles of 10 development periods and 10 accident periods are generated such that the dependence is strong in the first four
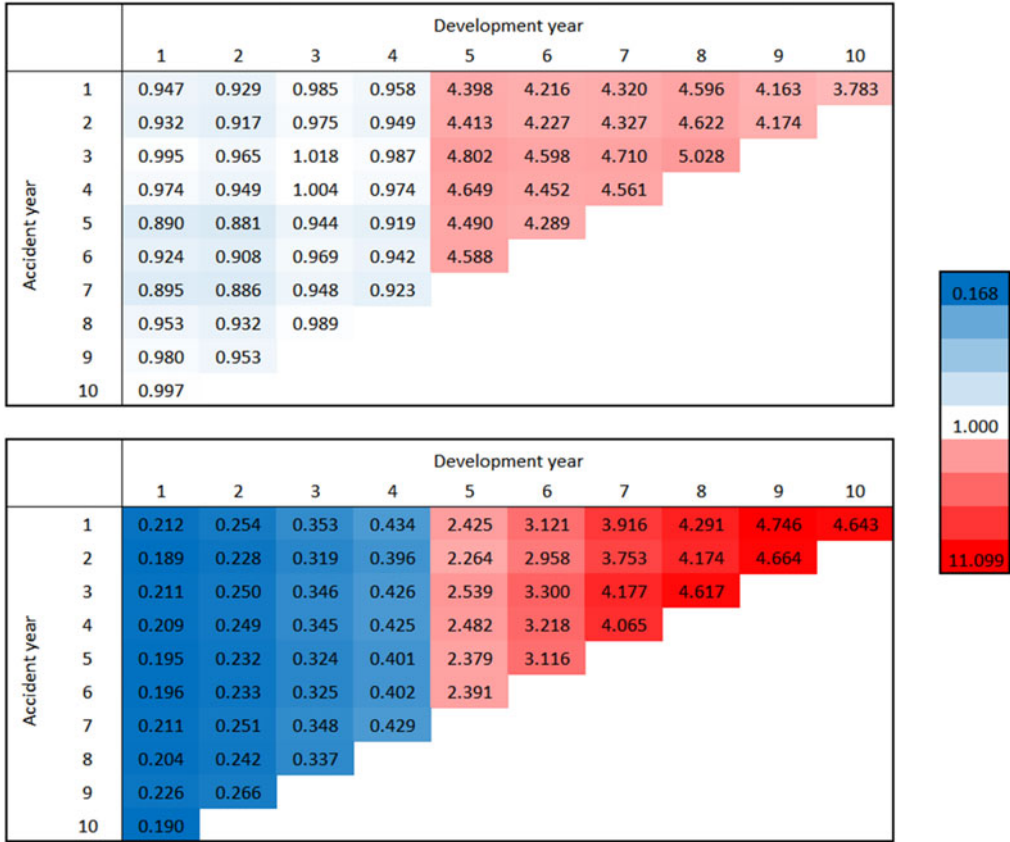
Figure 3. Heat maps of ratios of fitted common shock proportions to true proportions for triangle 1 (top: Tweedie framework modified for unbalanced data, bottom: original common shock Tweedie framework).

development periods, and not as strong in the last six periods. The common shock components are generated with column-specific mean parameters $\alpha_j = c_j \sqrt{v_j^{(1)} v_j^{(2)}}$ with $c_j = 0.5$ for $1 \leq j \leq 4$, and $c_j = 0.02$ for $5 \leq j \leq 10$. The second business line is also simulated to be longer-tailed than the first. Similar to the previous illustration, each observation in the two triangles is drawn from the multivariate Tweedie model for unbalanced data represented in section 3. These observations are assumed to have been standardised for accident year effect for simplicity. The two loss triangles are presented in Tables B.1 and B.2 in Appendix B.

Heat maps of ratios of fitted common shock proportions to true proportions are given in Figure 3 for triangle 1. Fitted values are calculated using posterior median of parameters, and true values are calculated using true parameter values. The modified Tweedie model provides a very good fit for the first four development periods. The goodness of fit is considerably less satisfactory in the later development periods when the true common shock proportion drops. The discrepancy is more significant for the first business line which has shorter tail development. The original common shock Tweedie model provides a poor goodness of fit overall, especially in early development periods. The proportions of common shock are underestimated in early development periods and overestimated in later periods. Even though not reported here, similar results are also observed in heat maps of ratios of fitted common shock proportions to true proportions for triangle 2.

Overall the modified Tweedie framework does not fully eliminate the issues of unbalanced data across development periods. However, there is a reduction. The fitting is quite good in early development periods but is less satisfactory in later periods. The use of the geometric average of column factors across multiple triangles may contribute to this performance as the geometric average may not be close to some individual column factors if the development patterns are too different. However, it is worth emphasising that the example used has quite an extreme variation in common shock proportions across development periods, and one should not expect such radical variation in practice. In addition, the poor performance also arises from the discrepancy between the modified model with a constant scaling term c and the true model generating the data (with column-specific scaling term c_j). With this specification, it is not surprising that the earlier (large) development periods dominate the estimation of c. We do not expect good results because of model misspecification, but we can arrive at two main conclusions: the modified framework outperforms the original framework; and the common shock proportions are mis-estimated in the higher development periods, where amounts are small and do not contribute significantly to total liability. It is also worth noting that the poor estimation of common shock proportion does not affect mean forecasts, only dependency between the triangles, and then only where the magnitudes of the forecasts are small.

5. Illustration with Real Data

The data used for illustration are a set of two triangles from the Bodily Injury line (1) and the Accident Benefit (excluding Disability Income) line (2) from a Canadian insurance company provided in Côté *et al.* (2016). These two triangles have also been used for illustrations in sections 1 and 2 and their details can be found therein.

5.1 Preliminary analysis

A preliminary analysis is performed to assess the suitability of this data set. This includes the assessment of the tails, as well as the dependence structure.

5.1.1 Analysis of the tails

From the plots of loss ratios provided earlier in Figure 1, it can be observed that the Bodily Injury line has longer claims development than the Accident Benefits line. Tail lengths of the two business lines are also assessed using age-to-age development factors

$$f_j^{(n)} = \frac{\sum_{i=1}^{I-j} Y_{i,j+1}^{(n)}}{\sum_{i=1}^{I-j} Y_{i,j}^{(n)}} \tag{31}$$

Results are given in Table 3. It can be observed that the development factors of the Bodily Injury dominate those of the Accident Benefits line for all development periods, except in the final year. However, this blip may be a false signal due to the truncation of data at the last development period and only one single observation is made in this final year. Hence, the Bodily Injury line is convincingly longer-tailed than the Accident Benefits line.

5.1.2 Explanatory dependence analysis

A heuristic dependence analysis is performed by fitting to each line a Tweedie Generalised Linear Model (GLM) with a log-link and the chain ladder mean structure

$$a_i^{(n)} + b_j^{(n)} \tag{32}$$

Table 3. Claims development factors for each development period

j	1	2	3	4	5	6	7	8	9
$f_j^{(1)}$	8.1617	1.8968	1.4521	1.2652	1.1249	1.0624	1.0225	1.0254	1.0092
$f_j^{(2)}$	2.5844	1.3584	1.1708	1.1140	1.0481	1.0305	1.0137	1.0057	1.0118

Table 4. Correlation coefficients between cellwise GLM residuals and their corresponding *p*-values

Pearson	Spearman	Kendall
0.3659 (0.0060)	0.3480 (0.0096)	0.2525 (0.0065)

Table 5. Correlation coefficients between cellwise GLM residuals and their corresponding *p*-values after removing fixed calendar year effects

Pearson	Spearman	Kendall
0.3416 (0.0107)	0.3250 (0.0159)	0.2202 (0.0176)

This is to remove fixed accident period and development period effects. Correlations between GLM Pearson residuals of the two lines are given in Table 4. The dependence between residuals is strong and significant after allowing for fixed accident period and development period effects.

To examine whether this strong correlation comes from calendar year effects that can impact both lines simultaneously, we also perform another GLM analysis with an additional fixed calendar year effect in the mean structure

$$a_i^{(n)} + b_j^{(n)} + h_t^{(n)} \tag{33}$$

Correlations between GLM Pearson residuals of the two lines are then given in Table 5. The correlation coefficients have been reduced, however, not very significantly.

Heat maps of residual ratios are given in Figure C.1 in Appendix C. Residual ratios are defined as ratios of observed values to GLM fitted values with the mean structure specified in equation (29). There are some common cellwise patterns that are quite obvious from the heat maps, for example, low payments in development year 7 compensated by accelerated payments in years 8–9 in the first accident year, payment dips in accident year 4 and development year 2. These show that this data set is suitable for illustration of the model.

Results from the preliminary analysis show that this data set is suitable to be used for illustration of the model.

5.2 Estimation and goodness-of-fit assessment

Bayesian inference is used for estimation. The marginal fitting is first performed. 400,000 simulations are run, and 300,000 simulations are discarded as the burn-in period. The sample chain is thinned by accepting every 5$^{\text{th}}$ iteration to reduce the serial dependence between iterations. The multivariate fitting is then performed with 90,000 simulations, and the first 30,000 are discarded as the burn-in period. The chain is then thinned by selecting every 3$^{\text{th}}$ iteration. Summary statistics are then computed on these posterior samples. The results are given in Tables C.3 and C.4 of Appendix C.

Marginal and multivariate goodness of fits are assessed. Marginal goodness of fit is assessed using QQ plots of residuals in Figure 4. The plot shows that the fit is quite off in the right tail of the Bodily Injury line and slightly off in both tails of the Accident Benefit line. The goodness of fit in other regions, however, is reasonable. This may be a result of the restriction of using the same
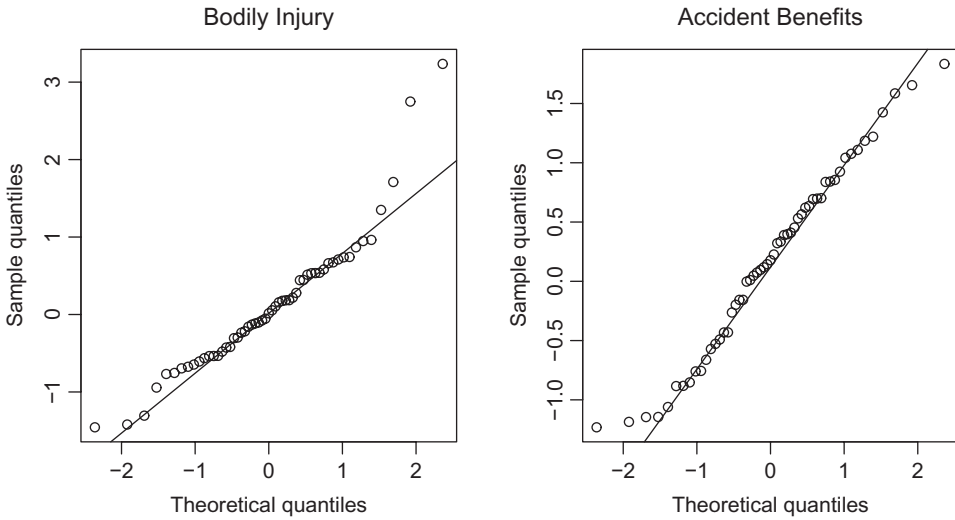
Figure 4. QQ plots of residuals from common shock Tweedie model ($p = 1.829$).
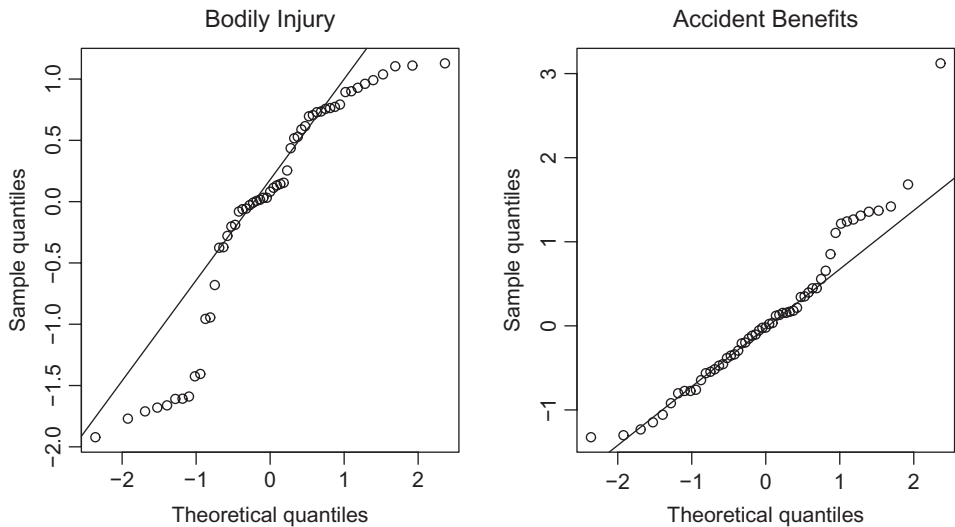

Figure 5. QQ plots of residuals from common shock normal model.

power parameter p for both lines. However, the multivariate Tweedie framework still provides marginal flexibility with flexible choices of p. For comparison, similar QQ plots are performed for a common shock normal model in Figure 5. It can be observed that the Tweedie marginals provide a much better fit than the normal marginals (with power parameter $p = 0$).

Multivariate goodness of fit is assessed by comparing the empirical bivariate marginals of real data observations and of back-fitted values. These are obtained using the empirical cumulative distribution functions of claim observations from each triangle. Because of the use of a Bayesian inference, various sets of back-fitted data can be generated. A path is randomly chosen for illustration. Scatter plots of these empirical bivariate marginals are presented in Figure 6. It can be observed that the model can capture the general positive dependence structure in the data.
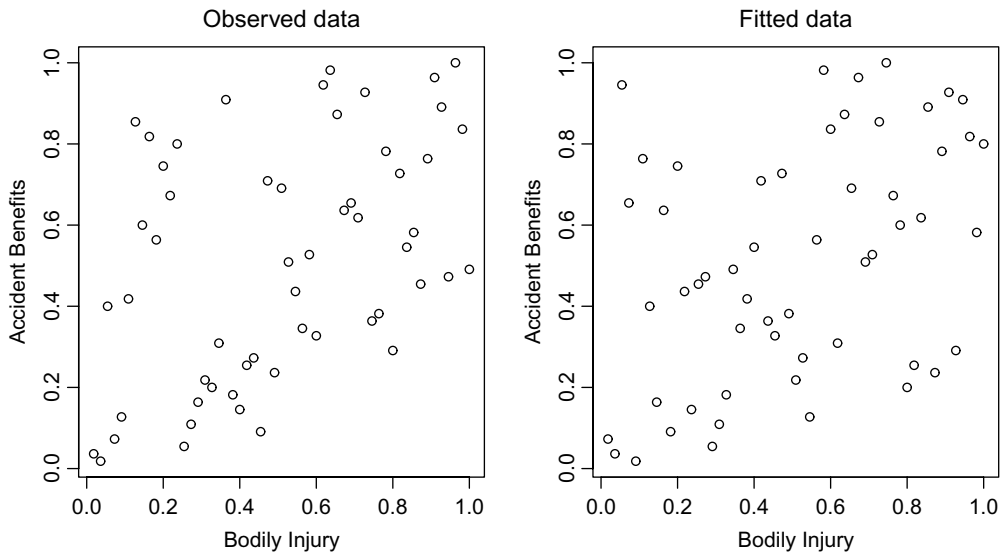
Figure 6. Plots of empirical bivariate marginals for observed values and back-fitted values.

To look for any trace of dependence not captured by the model, we examine the residuals from model fitting. These residuals are obtained as the differences between observations and fitted values, where the latter are calculated using posterior estimates. The Pearson correlation coefficient of these residuals reduces to 0.1204 (*p*-value 0.3812). This is much weaker than the correlation coefficient of 0.3416 of GLM Pearson residuals in section 5.1.2 and is also insignificant. The insignificant correlation indicates that our model has explained away most of the dependence in the data.

5.3 Common shock proportions

Predictive distributions of outstanding claim observations in the lower triangles can be calculated using predictive Bayesian inference. Using parameter estimates, the contributions of common shock within each cell in the two triangles are calculated and given in Tables 6 and 7. It can be observed that there is only a very mild variation in the common shock proportions within and across triangles. We can relate this result to the challenges coming from applying a common shock model to loss reserving data which has an unbalanced nature discussed in section 2. It shows that the proposed approach has provided a balance of common shock proportions across all loss cells within and across loss triangles.

5.4 Outstanding claims forecast

To obtain the distributions of the outstanding claims, posterior samples of parameters from the Bayesian inference are used to project claims in lower triangles. This projection utilises the specification in equations (7), (16) and (18). This gives a set of samples of future claims in the lower triangles. Using this set, summary statistics of the total outstanding claims distributions are given in Table 8 and kernel densities of outstanding claims are given in Figure 7. Summary statistics provided include the posterior mean, standard deviation, $VaR_{75\%}$ and $VaR_{95\%}$ of the distribution of total outstanding claims for each line, as well as for both lines.

The empirical bivariate marginals of total reserves are shown in Figure 8. For illustration purpose, we show the scatter plot of total reserves from 1,000 posterior samples. The plot shows a mild positive dependence structure in the total outstanding claims across two lines. This is accompanied by a Pearson correlation of 0.0855 (*p*-value <2.2e-16). There is no clear evidence

Table 6. Proportions of common shock to the expected total observations calculated using parameter estimates – Bodily Injury

		Development year									
		1	2	3	4	5	6	7	8	9	10
Accident year	1	4.8%	4.2%	4.1%	4.0%	3.9%	3.9%	4.0%	4.1%	3.6%	4.2%
	2	5.2%	4.6%	4.4%	4.3%	4.2%	4.2%	4.3%	4.4%	3.9%	4.5%
	3	4.9%	4.3%	4.2%	4.0%	3.9%	4.0%	4.1%	4.2%	3.7%	4.2%
	4	5.1%	4.5%	4.3%	4.2%	4.1%	4.2%	4.2%	4.3%	3.8%	4.4%
	5	4.9%	4.3%	4.1%	4.0%	3.9%	4.0%	4.0%	4.1%	3.6%	4.2%
	6	5.0%	4.4%	4.2%	4.1%	4.0%	4.1%	4.1%	4.2%	3.7%	4.3%
	7	4.7%	4.1%	4.0%	3.8%	3.7%	3.8%	3.9%	4.0%	3.5%	4.0%
	8	5.0%	4.3%	4.2%	4.1%	4.0%	4.1%	4.1%	4.2%	3.7%	4.3%
	9	5.1%	4.5%	4.3%	4.2%	4.1%	4.2%	4.2%	4.3%	3.8%	4.4%
	10	6.2%	5.4%	5.3%	5.1%	5.0%	5.1%	5.1%	5.2%	4.6%	5.3%

Table 7. Proportions of common shock to the expected total observations calculated using parameter estimates – Accident Benefits

		Development year									
		1	2	3	4	5	6	7	8	9	10
Accident year	1	4.4%	5.0%	5.1%	5.3%	5.4%	5.4%	5.3%	5.2%	5.9%	5.1%
	2	4.6%	5.2%	5.3%	5.5%	5.7%	5.6%	5.5%	5.4%	6.1%	5.3%
	3	4.4%	5.0%	5.1%	5.3%	5.4%	5.3%	5.3%	5.1%	5.9%	5.1%
	4	4.2%	4.7%	4.9%	5.1%	5.2%	5.1%	5.0%	4.9%	5.6%	4.8%
	5	4.4%	5.0%	5.1%	5.3%	5.4%	5.4%	5.3%	5.2%	5.9%	5.1%
	6	4.1%	4.7%	4.8%	5.0%	5.1%	5.1%	5.0%	4.9%	5.5%	4.8%
	7	4.1%	4.7%	4.8%	5.0%	5.1%	5.0%	4.9%	4.8%	5.5%	4.7%
	8	4.2%	4.8%	5.0%	5.1%	5.3%	5.2%	5.1%	5.0%	5.7%	4.9%
	9	4.2%	4.8%	5.0%	5.1%	5.3%	5.2%	5.1%	5.0%	5.7%	4.9%
	10	4.0%	4.6%	4.7%	4.9%	5.0%	4.9%	4.8%	4.7%	5.4%	4.6%

Table 8. Summary statistics of outstanding claims distributions

	Bodily Injury	Accident Benefits	Both lines
Mean	165,185.92	108,465.81	273,651.73
SD	22,720.88	18,554.65	30,538.83
VaR$_{75\%}$	179,057.18	120,100.43	293,061.56
VaR$_{95\%}$	205,752.20	141,426.24	326,177.22

of a concentration in the tail regions of the dependence. There can be diversification across claims within a single loss triangle. Hence, the dependence on the aggregate reserves from each line is mild and tail dependence may not be apparent.

The two business lines do not have a comonotonic dependence structure, and this allows the insurer to gain some diversification benefits when they set their risk margins. Using the specification of a risk margin under APRA's Prudential Standards GPS 340, we have the following definition of risk margin and diversification benefit

Table 9. Risk margin and diversification benefits statistics

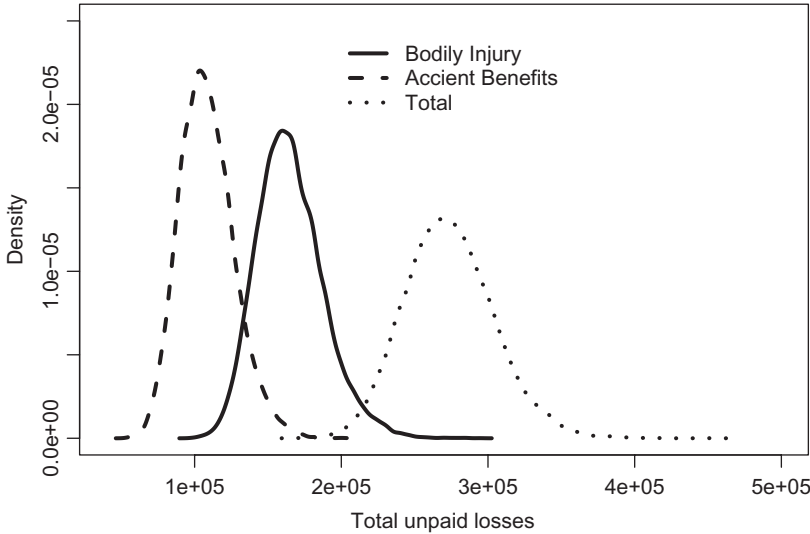|	Bodily injury	Accident benefits	Both lines	Diversification benefits (%)
Risk margin$_{75\%}$	13,871.26	11,634.61	19,409.83	23.9
Risk margin$_{95\%}$	40,566.28	32,960.43	52,525.49	28.6

Figure 7. Kernel densities of predictive distributions of total outstanding claims in each line of business and in the aggregate portfolio.

$$\text{Risk margin}_{\chi\%}[Y] = \max\left\{\text{VaR}_{\chi\%}[Y] - \text{E}[Y]; \frac{1}{2}\text{SD}[Y]\right\} \tag{34}$$

Diversification benefit

$$= \frac{\left(\text{Risk margin}_{\chi\%}[Y_1] + \text{Risk margin}_{\chi\%}[Y_2]\right) - \text{Risk margin}_{\chi\%}[Y_1 + Y_2]}{\text{Risk margin}_{\chi\%}[Y_1] + \text{Risk margin}_{\chi\%}[Y_2]} \times 100\% \tag{35}$$

Risk margin$_{75\%}$ and Risk margin$_{95\%}$, as well as associated diversification benefits, are provided in Table 9. It can then be observed that diversification benefits can be gained as a result of allowing for (non-comonotonic) dependence across business lines.

6. Conclusion

Common shock approaches can provide many benefits in the modelling of outstanding claims. However, they often require very careful parametrisation. This arises from the unbalanced nature of loss reserving data. It is often desirable to use scaling factors to adjust the common shock effects so that they can contribute proportionately to the total observations over the entire range of the triangles. However, an excessive use of scaling factors can result in overparametrisation. In some cases, such as the common shock Tweedie framework developed in Avanzi *et al.* (2016*b*), it is also desirable to select scaling factors such that distributional tractability is preserved. These requirements place conflicting constraints on the specification of scaling factors in common shock models.
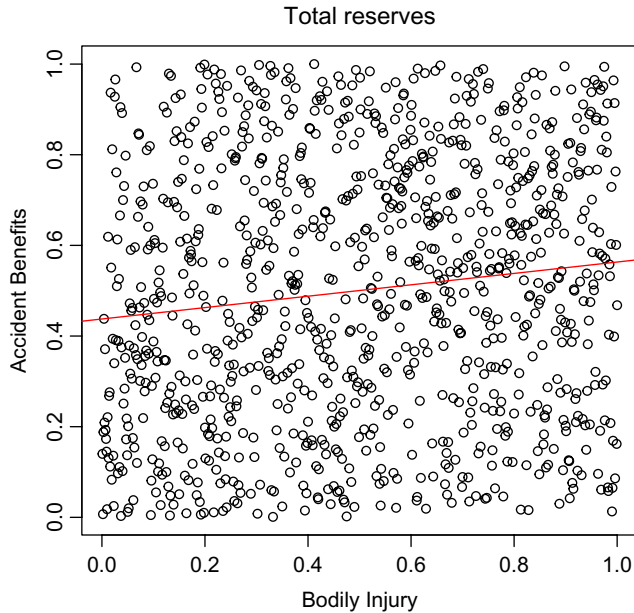
Figure 8. Plot of empirical bivariate marginals of total reserves (using 1,000 posterior samples).

In this paper, we propose an approach that compromises the various constraints mentioned above. This approach involves using careful and parsimonious parameterisation to develop a common shock Tweedie framework modified for unbalanced data. Additional modifications for negative claims are also undertaken under this framework. Illustrations using simulated and with real data are presented. These illustrations show that while the proposed approach cannot fully eliminate the issue of unbalanced common shock proportions, the improvement over the original framework in Avanzi *et al.* (2016*b*) is quite substantial.

We examined a common shock Tweedie approach for cellwise dependence in this paper. Future research could consider applications on other structures of dependence (such as calendar period dependence). This paper raises some potential issues of common shock models when they are applied to reserving data that have an unbalanced nature. These issues, however, might appear whenever common shock models are applied to heterogeneous data. These can include mortality data for different group ages, or capital modelling for different types of risks. The proposed solution could be extended to solve similar problems in other contexts. While this solution can reduce the problems of unbalanced data quite substantially, a complete balance in common shock proportions cannot be achieved. Future research could consider a better solution to this problem. Other multivariate models with explicit dependence structures such as mixture models could also be considered as they might be more applicable to unbalanced data.

References

Abdallah, A., **Boucher, J.-P. & Cossette, H.** (2015). Modeling dependence between loss triangles with hierarchical Archimedean copulas. *ASTIN Bulletin*, **45**, 577–599.

Ajne, B. (1994). Additivity of chain-ladder projections. *ASTIN Bulletin*, **24**(2), 311–318.

Alai, D.H., **Landsman, Z. & Sherris, M.** (2013). Lifetime dependence modelling using a truncated multivariate gamma distribution. *Insurance: Mathematics and Economics*, **52**(3), 542–549.

Alai, D.H., **Landsman, Z. & Sherris, M.** (2015). A multivariate Tweedie lifetime model: censoring and truncation. *Insurance: Mathematics and Economics*, **64**, 203–213.

Alai, D.H., **Landsman, Z. & Sherris, M.** (2016). Multivariate Tweedie lifetimes: the impact of dependence, *Scandinavian Actuarial Journal*, **2016**(8), 692–712.

Alai, D.H. & Wüthrich, M.V. (2009). Taylor approximations for model uncertainty within the Tweedie exponential dispersion family. *ASTIN Bulletin*, **39**(2), 453.

Avanzi, B., **Taylor, G.**, **Vu, P.A. & Wong, B.** (2016*b*). Stochastic loss reserving with dependence: a flexible multivariate Tweedie approach. *Insurance: Mathematics and Economics*, **71**, 63–78.

Avanzi, B., **Taylor, G.**, **Vu, P.A. & Wong, B.** (2020). A multivariate evolutionary generalised linear model framework with adaptive estimation for claims reserving. *Insurance: Mathematics and Economics*, **93**, 50–71.

Avanzi, B., **Taylor, G. & Wong, B.** (2016*a*). Correlations between insurance lines of business: an illusion or a real phenomenon? Some methodological considerations. *ASTIN Bulletin*, **46**(2), 225–263.

Avanzi, B., **Taylor, G. & Wong, B.** (2018). Common shock models for claim arrays. *ASTIN Bulletin*, **48**(3), 1–28.

Avanzi, B., **Wong, B. & Yang, X.** (2016). A micro-level claim count model with overdispersion and reporting delays. *Insurance: Mathematics and Economics*, **71**, 1–14.

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.

Braun, C. (2004). The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin*, **34**, 399–424.

Congdon, P.D. (2010). *Applied Bayesian Hierarchical Methods*. Chapman & Hall, Boca Raton.

Côté, M.-P., **Genest, C. & Abdallah, A.** (2016). Rank-based methods for modeling dependence between loss triangles. *European Actuarial Journal*, **6**(2), 377–408.

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. (2006). Forecasting runoff triangles. *North American Actuarial Journal*, **10**(2), 28–38.

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. & Wüthrich, M.V.** (2012). Bayesian over-dispersed Poisson model and the Bornhuetter–Ferguson claims reserving method. *Annals of Actuarial Science*, **6**(2), 258–283.

Furman, E. & Landsman, Z. (2010). Multivariate Tweedie distributions and some related capital-at-risk analyses. *Insurance: Mathematics and Economics*, **46**(2), 351–361.

Gismondi, F., **Janssen, J. & Manca, R.** (2012). The construction of the claims reserve distribution by means of a semi-Markov backward simulation model. *Annals of Actuarial Science*, **6**(1), 23–64.

Haario, H., **Saksman, E. & Tamminen, J.** (2001). An adaptive Metropolis algorithm. Bernoulli, **7**(2), 223–242.

Heberle, J. & Thomas, A. (2016). The fuzzy Bornhuetter–Ferguson method: an approach with fuzzy numbers. *Annals of Actuarial Science*, **10**(2), 303–321.

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.

International Actuarial Association (2004). A global framework for insurer solvency assessment, Website. Last accessed: 9/12/2018. https://www.actuaries.org/LIBRARY/Papers/Global_Framework_Insurer_Solvency_Assessment-public.pdf

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, New York.

Jørgensen, B. (1997). *The Theory of Dispersion Models*. Chapman & Hall, London.

Kunkler, M. (2006). Modelling negatives in stochastic reserving models. *Insurance: Mathematics and Economics*, **38**(3), 540–555.

Merz, M. & Wüthrich, M.V. (2007). Prediction error of the chain ladder reserving method applied to correlated run-off triangles. *Annals of Actuarial Science*, **2**(1), 25–50.

Merz, M. & Wüthrich, M.V. (2009). Prediction error of the multivariate additive loss reserving method for dependent lines of business. *Variance*, **3**(1), 131–151.

Meyers, G.G. (2007). The common shock model for correlated insurance losses. *Variance*, **1**(1), 40–52.

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.

Pinheiro, P.J.R., Andrade e Silva, J.M. & de Lourdes Centeno, M. (2003). Bootstrap methodology in claim reserving. *Journal of Risk and Insurance*, **70**(4), 701–714.

Pröhl, C. & Schmidt, K.D. (2005). *Multivariate Chain-Ladder*, Technology University, Institute für Mathematische Stochastik.

Renshaw, A.E. & Verrall, R.J. (1998). A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, **4**(04), 903–923.

Saluz, A. & Gisler, A. (2014). Best estimate reserves and the claims development results in consecutive calendar years. *Annals of Actuarial Science*, **8**(2), 351–373.

Schmidt, K.D. (2006). Optimal and additive loss reserving for dependent lines of business. *Casualty Actuarial Society Forum (Fall)*, pp. 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. (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.

Taylor, G.C. (2000). *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers, Boston.

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**(3), 331–346.

Wüthrich, M.V. (2018). Machine learning in individual claims reserving. *Scandinavian Actuarial Journal*, **2018**(6), 465–480.

Wüthrich, M.V. & Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. John Wiley & Sons, Chichester.

Zhang, Y. (2010). A general multivariate chain ladder model. *Insurance: Mathematics and Economics*, **46**(3), 588–599.

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. & Guszcza, J. (2012). A Bayesian non-linear model for forecasting insurance loss payments. *Journal of Royal Statistical Society*, **175**(2), 637–656.

Zhao, X.B., Zhou, X. & Wang, J.L. (2009). Semiparametric model for prediction of individual claim loss reserving. *Insurance: Mathematics and Economics*, **45**(1), 1–8.

Appendix A. Simulated Data Set 1

Table A.1. Simulated triangle 1 (data set 1)

		Development year									
		1	2	3	4	5	6	7	8	9	10
Accident year	1	85.57	43.18	20.58	13.40	4.40	2.34	1.86	0.55	0.28	0.15
	2	78.22	28.65	12.74	5.08	6.97	2.82	1.50	0.07	−0.01	
	3	85.90	36.58	22.21	14.29	2.23	3.31	0.82	1.86		
	4	67.86	36.94	16.01	11.23	5.54	4.68	1.40			
	5	83.45	33.30	21.24	10.80	4.32	3.04				
	6	63.85	39.38	24.71	2.84	7.77					
	7	78.80	31.17	16.96	8.27						
	8	90.32	36.19	13.56							
	9	97.94	35.43								
	10	58.14									

Table A.2. Simulated triangle 2 (data set 1)

		Development year									
		1	2	3	4	5	6	7	8	9	10
Accident year	1	24.12	38.93	45.70	43.19	16.04	8.70	4.78	1.83	1.45	1.66
	2	21.04	40.05	35.83	19.93	15.27	11.21	6.84	2.81	1.12	
	3	23.98	38.59	40.73	47.22	22.01	10.36	3.49	3.53		
	4	26.34	42.48	57.27	29.72	24.03	12.11	1.86			
	5	29.46	33.18	44.63	39.51	25.97	11.60				
	6	23.67	48.70	49.66	20.12	21.34					
	7	29.10	36.51	50.52	43.98						
	8	30.58	53.40	49.21							
	9	31.16	50.48								
	10	31.04									

Table A.3. Posterior statistics of parameters (data set 1)

	True value	Median	SD	90% CI		True value	Median	SD	90% CI
$\eta_2^{(1)}$	1.0300	0.8950	0.0756	(0.8250; 1.0610)	$\eta_2^{(2)}$	1.1900	1.0620	0.1285	(0.8850; 1.3050)
$\eta_3^{(1)}$	1.1900	1.2820	0.1523	(1.0530; 1.5560)	$\eta_3^{(2)}$	1.1700	1.2300	0.1693	(0.9950; 1.5440)
$\eta_4^{(1)}$	1.1200	1.0250	0.1190	(0.8590; 1.2460)	$\eta_4^{(2)}$	1.1500	1.1020	0.1507	(0.8860; 1.3810)
$\eta_5^{(1)}$	1.1500	1.0600	0.1200	(0.8790; 1.2700)	$\eta_5^{(2)}$	1.1500	1.3190	0.1635	(1.0710; 1.6040)
$\eta_6^{(1)}$	1.1600	1.1740	0.1468	(0.9650; 1.4460)	$\eta_6^{(2)}$	1.2000	1.1390	0.1421	(0.9340; 1.3970)
$\eta_7^{(1)}$	1.1200	1.0010	0.1159	(0.8530; 1.2270)	$\eta_7^{(2)}$	1.4000	1.4480	0.1912	(1.1690; 1.7950)
$\eta_8^{(1)}$	1.1400	1.0380	0.1219	(0.8690; 1.2700)	$\eta_8^{(2)}$	1.4500	1.5000	0.2095	(1.1890; 1.8750)
$\eta_9^{(1)}$	1.2100	1.1310	0.1343	(0.9350; 1.3760)	$\eta_9^{(2)}$	1.5600	1.4820	0.1908	(1.2110; 1.8320)
$\eta_{10}^{(1)}$	1.1900	1.0680	0.1879	(0.8530; 1.4470)	$\eta_{10}^{(2)}$	1.6600	1.9840	0.5113	(1.3110; 2.9670)
$\nu_1^{(1)}$	60.0000	56.8610	4.2527	(49.8400; 63.7100)	$\nu_1^{(2)}$	10.0000	10.2800	1.1138	(8.5280; 12.1490)
$\nu_2^{(1)}$	20.0000	21.7210	1.9812	(18.6780; 25.1880)	$\nu_2^{(2)}$	20.0000	22.1790	2.0222	(18.9970; 25.6420)
$\nu_3^{(1)}$	10.0000	10.0160	1.3134	(8.1100; 12.4490)	$\nu_3^{(2)}$	25.0000	27.2130	2.5612	(23.1110; 31.5170)
$\nu_4^{(1)}$	5.0000	4.6610	0.7082	(3.5870; 5.8940)	$\nu_4^{(2)}$	20.0000	21.2330	2.1120	(17.9160; 24.8910)
$\nu_5^{(1)}$	2.5000	2.2920	0.2783	(1.8700; 2.7830)	$\nu_5^{(2)}$	15.0000	13.9600	1.6821	(11.4140; 16.8640)
$\nu_6^{(1)}$	1.2500	1.5100	0.2095	(1.2120; 1.892)	$\nu_6^{(2)}$	8.0000	6.5560	0.9920	(5.0850; 8.33707)
$\nu_7^{(1)}$	0.6000	0.7290	0.1453	(0.5320; 1.0040)	$\nu_7^{(2)}$	3.0000	2.7070	0.4708	(2.0370; 3.5830)
$\nu_8^{(1)}$	0.3000	0.2620	0.0449	(0.2010; 0.3460)	$\nu_8^{(2)}$	2.0000	1.9430	0.3335	(1.4350; 2.5340)
$\nu_9^{(1)}$	0.1500	0.1430	0.0272	(0.1040; 0.1930)	$\nu_9^{(2)}$	1.0000	0.7700	0.1122	(0.6180; 1.0100)
$\nu_{10}^{(1)}$	0.1500	0.2040	0.0445	(0.1440; 0.2860)	$\nu_{10}^{(2)}$	1.0000	0.9480	0.1997	(0.6580; 1.3140)
$\gamma^{(1)}$	0.5000	0.5530	0.0649	(0.4560; 0.6680)	$\gamma^{(2)}$	0.7000	0.6220	0.0936	(0.4880; 0.7930)
p	1.3000	1.3640	0.0507	(1.2850; 1.4510)	$\xi^{(1)}$	0.0100	0.0120	0.0013	(0.0100; 0.0140)
δ	1.0118	1.0420	0.1293	(0.8470; 1.2710)					
c	0.500	0.3620	0.6942	(0.0600; 2.2370)	β	0.600	0.5030	0.4603	(0.1610; 1.6020)
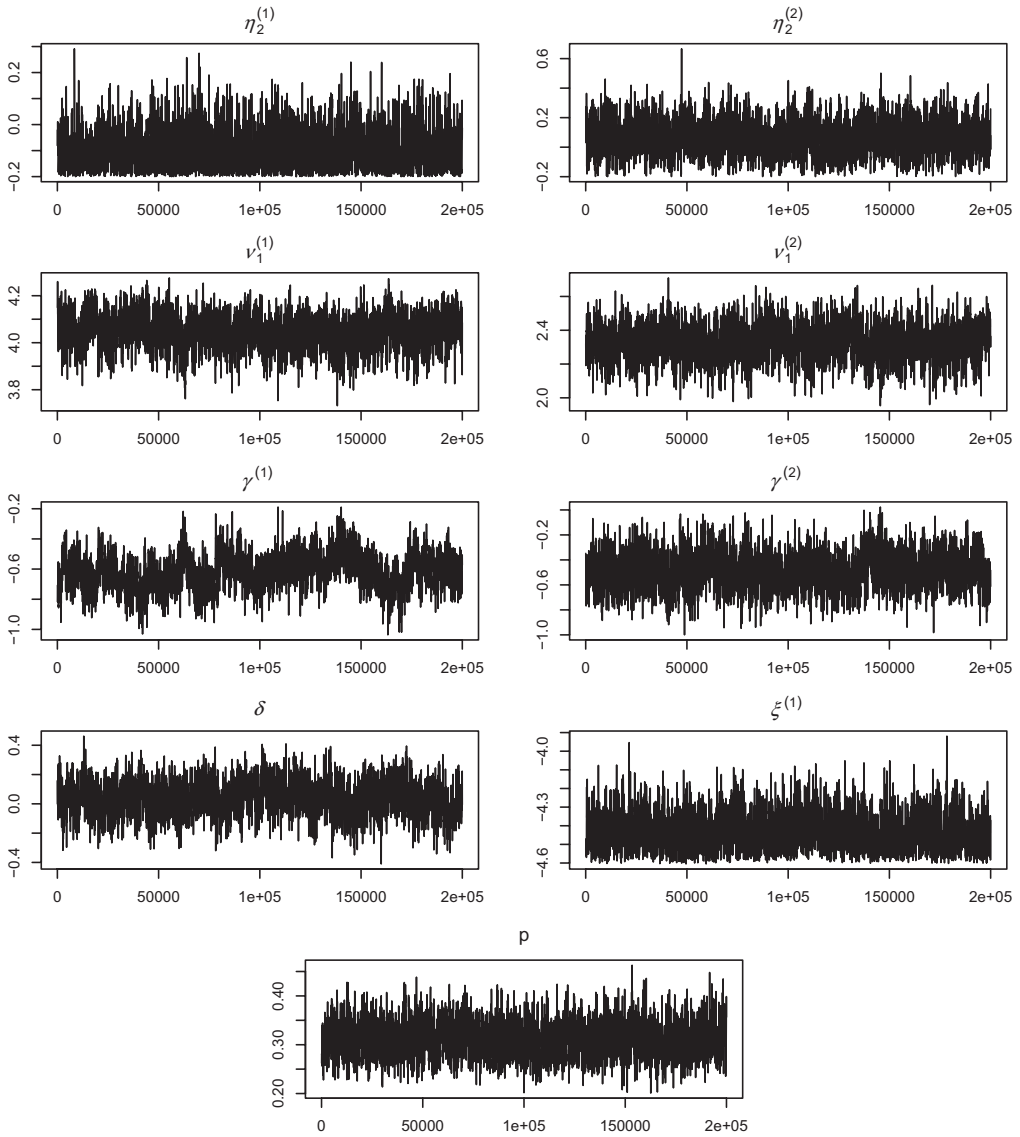
Figure A.1. MCMC sample paths of some parameters.

Appendix B. Simulated Data Set 2

Table B.1. Simulated triangle 1 (data set 2)

		1	2	3	4	5	6	7	8	9	10
						Development year					
Accident year	1	47.16	36.33	18.58	10.63	3.93	0.38	0.45	0.00	0.00	0.13
	2	102.30	49.12	18.47	17.05	3.30	1.70	1.77	1.04	0.00	
	3	101.87	56.91	14.75	24.29	1.46	1.15	0.83	0.17		
	4	97.09	35.96	27.80	10.86	3.93	3.71	0.43			
	5	107.07	34.34	20.49	19.73	6.65	1.70				
	6	107.10	66.55	27.03	17.09	2.38					
	7	123.60	37.41	32.77	15.53						
	8	107.03	50.50	18.30							
	9	105.93	42.02								
	10	109.09									

Table B.2. Simulated triangle 2 (data set 2)

		1	2	3	4	5	6	7	8	9	10
						Development year					
Accident year	1	19.61	45.29	44.23	28.82	24.48	3.15	3.23	3.30	1.94	0.73
	2	33.52	41.13	39.33	41.78	22.46	8.69	2.09	4.85	1.88	
	3	24.39	43.40	34.06	59.94	22.00	13.90	5.54	1.62		
	4	27.78	37.03	41.41	31.12	31.73	5.92	7.69			
	5	24.46	41.96	36.55	23.42	20.88	9.61				
	6	26.36	38.68	58.52	36.25	27.15					
	7	30.05	36.18	52.14	41.98						
	8	30.32	53.54	52.87							
	9	42.37	42.25								
	10	46.49									

Appendix C.   Real Data Set

This data set is drawn from Côté *et al.* (2016).

Table C.1.   Bodily Injury line (cumulative claims)

		Premium					Development year					
			1	2	3	4	5	6	7	8	9	10
Accident year	1	85,421	3,488	14,559	27,249	37,979	49,561	55,957	58,406	60,862	63,280	63,864
	2	98,579	1,169	12,781	20,550	31,547	42,808	47,385	50,251	50,978	51,272	
	3	103,062	1,478	10,788	25,499	34,279	43,057	49,360	52,329	52,544		
	4	108,412	1,186	11,852	22,913	32,537	41,824	48,005	52,542			
	5	111,176	1,737	13,881	25,521	38,037	43,684	47,755				
	6	112,050	1,571	12,153	27,329	41,832	51,779					
	7	112,577	1,199	17,077	29,876	44,149						
	8	113,707	1,263	16,073	28,249							
	9	126,442	986	10,003								
	10	130,484	683									

Table C.2.   Accident Benefits (cumulative claims)

		Premium					Development year					
			1	2	3	4	5	6	7	8	9	10
Accident year	1	116,491	13,714	24,996	31,253	38,352	44,185	46,258	47,019	47,894	48,334	48,902
	2	111,467	6883	16,525	24,796	29,263	32,619	33,383	34,815	35,569	35,612	
	3	107,241	7933	22,067	32,801	38,028	44,274	44,948	46,507	46,665		
	4	105,687	7052	18,166	25,589	31,976	36,092	38,720	39,914			
	5	105,923	10,463	23,982	31,621	36,039	38,070	41,260				
	6	111,487	9697	28,878	41,678	47,135	50,788					
	7	113,268	11,387	37,333	48,452	55,757						
	8	121,606	12,150	32,250	40,677							
	9	110,610	5348	14,357								
	10	104,304	4,612									

		Development year
		1	2	3	4	5	6	7	8	9	10
Accident year	1	1.000	0.988	0.938	1.028	0.987	0.963	0.746	1.654	1.504	1.000
	2	1.132	1.070	0.906	0.941	1.281	0.997	0.854	0.697	0.243	
	3	1.131	1.035	1.172	0.767	1.106	1.027	0.963	0.210		
	4	1.165	0.863	0.959	0.992	0.932	1.169	1.594			
	5	1.254	1.079	1.201	1.037	0.664	0.842				
	6	0.983	0.883	0.994	1.111	1.009					
	7	0.834	0.996	0.917	1.117						
	8	0.706	1.087	0.953							
	9	0.904	1.013								
	10	1.000									

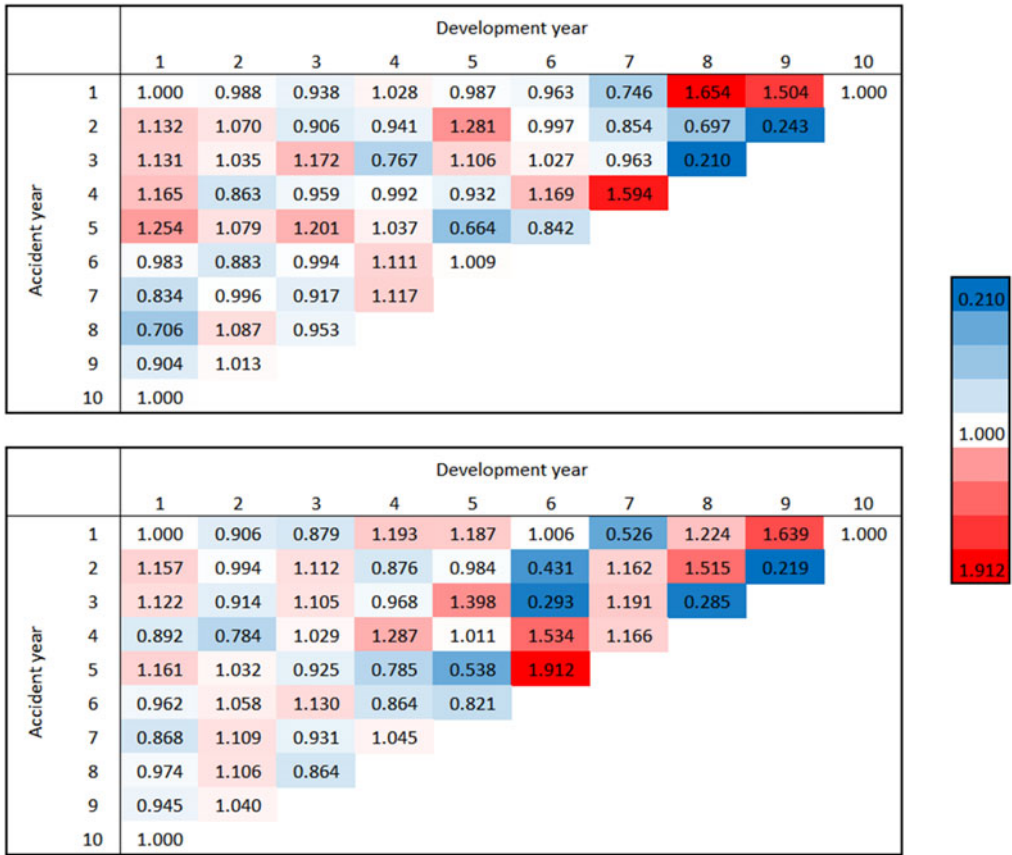| | | Development year |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Accident year | 1 | 1.000 | 0.906 | 0.879 | 1.193 | 1.187 | 1.006 | 0.526 | 1.224 | 1.639 | 1.000 |
| | 2 | 1.157 | 0.994 | 1.112 | 0.876 | 0.984 | 0.431 | 1.162 | 1.515 | 0.219 | |
| | 3 | 1.122 | 0.914 | 1.105 | 0.968 | 1.398 | 0.293 | 1.191 | 0.285 | | |
| | 4 | 0.892 | 0.784 | 1.029 | 1.287 | 1.011 | 1.534 | 1.166 | | | |
| | 5 | 1.161 | 1.032 | 0.925 | 0.785 | 0.538 | 1.912 | | | | |
| | 6 | 0.962 | 1.058 | 1.130 | 0.864 | 0.821 | | | | | |
| | 7 | 0.868 | 1.109 | 0.931 | 1.045 | | | | | | |
| | 8 | 0.974 | 1.106 | 0.864 | | | | | | | |
| | 9 | 0.945 | 1.040 | | | | | | | | |
| | 10 | 1.000 | | | | | | | | | |

Figure C.1. Heat maps of ratios of observed values to GLM fitted values (top: Bodily Injury line, bottom: Accident Benefits).

Table C.3. Posterior statistics of parameters from marginal estimation

	Median	SD	90% CI		Median	SD	90% CI
$\eta_2^{(1)}$	0.6390	0.0875	(0.5120; 0.7970)	$\eta_2^{(2)}$	0.7770	0.1119	(0.6230; 0.9830)
$\eta_3^{(1)}$	0.9010	0.1026	(0.7510; 1.0860)	$\eta_3^{(2)}$	1.0050	0.1508	(0.7860; 1.2770)
$\eta_4^{(1)}$	0.7340	0.1207	(0.5710; 0.9610)	$\eta_4^{(2)}$	1.3690	0.1773	(1.1220; 1.6980)
$\eta_5^{(1)}$	0.9680	0.1461	(0.7650; 1.2390)	$\eta_5^{(2)}$	0.9980	0.1390	(0.8010;1.2550)
$\eta_6^{(1)}$	0.8260	0.1149	(0.6580; 1.0310)	$\eta_6^{(2)}$	1.4150	0.2656	(1.0520; 1.9120)
$\eta_7^{(1)}$	1.2260	0.1464	(1.0140; 1.4880)	$\eta_7^{(2)}$	1.5060	0.1727	(1.2560; 1.8240)
$\eta_8^{(1)}$	0.8510	0.1276	(0.6760; 1.0940)	$\eta_8^{(2)}$	1.2390	0.2478	(0.9200; 1.7170)
$\eta_9^{(1)}$	0.7230	0.0772	(0.6060; 0.8600)	$\eta_9^{(2)}$	1.2450	0.2061	(0.9440; 1.6230)
$\eta_{10}^{(1)}$	0.2200	0.0592	(0.1480; 0.3360)	$\eta_{10}^{(2)}$	1.7260	0.3244	(1.2560; 2.3350)
$\nu_1^{(1)}$	0.0160	0.0022	(0.0130; 0.0200)	$\nu_1^{(2)}$	0.0590	0.0080	(0.0470; 0.0740)
$\nu_2^{(1)}$	0.1430	0.0192	(0.1140; 0.1770)	$\nu_2^{(2)}$	0.1050	0.0139	(0.0840; 0.1300)
$\nu_3^{(1)}$	0.1270	0.0136	(0.1060; 0.1510)	$\nu_3^{(2)}$	0.0670	0.0095	(0.0530; 0.0840)
$\nu_4^{(1)}$	0.0930	0.0111	(0.0760; 0.1120)	$\nu_4^{(2)}$	0.0310	0.0040	(0.0250; 0.0390)
$\nu_5^{(1)}$	0.1190	0.0153	(0.0970; 0.1470)	$\nu_5^{(2)}$	0.0300	0.0032	(0.0250; 0.0350)
$\nu_6^{(1)}$	0.0510	0.0088	(0.0380; 0.0670)	$\nu_6^{(2)}$	0.0160	0.0019	(0.0130; 0.0190)
$\nu_7^{(1)}$	0.0400	0.0065	(0.0310; 0.0520)	$\nu_7^{(2)}$	0.0150	0.0018	(0.0120; 0.0180)
$\nu_8^{(1)}$	0.0100	0.0010	(0.0090; 0.0120)	$\nu_8^{(2)}$	0.0050	0.0008	(0.0040; 0.0070)
$\nu_9^{(1)}$	0.0200	0.0040	(0.0140; 0.0270)	$\nu_9^{(2)}$	0.0020	0.0003	(0.0020; 0.0030)
$\nu_{10}^{(1)}$	0.0050	0.0008	(0.0040; 0.0060)	$\nu_{10}^{(2)}$	0.0030	0.0004	(0.0020; 0.0030)
$\gamma^{(1)}$	0.1400	0.0372	(0.0900; 0.2120)	$\gamma^{(2)}$	0.1580	0.0431	(0.1030; 0.2430)
δ	0.3240	0.0732	(0.2220; 0.4610)	p	1.8290	0.0660	(1.7120; 1.9260)

Table C.4. Posterior statistics of parameters from multivariate estimation

	Median	SD	90% CI
c	1.0080	4.6868	(0.0570; 17.2280)
β	3.0910	0.9413	(1.8920; 5.0220)