

# COMMON SHOCK MODELS FOR CLAIM ARRAYS

BY

BENJAMIN AVANZI, GREG TAYLOR AND BERNARD WONG

## ABSTRACT

The paper is concerned with multiple claim arrays. In recognition of the extensive use by practitioners of large correlation matrices for the estimation of diversification benefits in capital modelling, we develop a methodology for the construction of such correlation structures (to any dimension). Indeed, the literature does not document any methodology by which practitioners, who often parameterise those correlations by means of informed guesswork, may do so in a disciplined and parsimonious manner.

We construct a broad and flexible family of models, where dependency is induced by common shock components. Models incorporate dependencies between observations both within arrays and between arrays. Arrays are of general shape (possibly with holes), but include the usual cases of claim triangles and trapezia that appear in the literature. General forms of dependency are considered with cell-, row-, column-, diagonal-wise, and other forms of dependency as special cases. Substantial effort is applied to practical interpretation of such matrices generated by the models constructed here.

Reasonably realistic examples are examined, in which an expression is obtained for the general entry in the correlation matrix in terms of a limited set of parameters, each of which has a straightforward intuitive meaning to the practitioner. This will maximise chance of obtaining a reliable matrix. This construction is illustrated by a numerical example.

## KEYWORDS

Claim array, claim dependency, common shock, correlation, loss reserving.  
JEL codes: G22, C52, C55

## 1. INTRODUCTION

### 1.1. Background and motivation

Calculation of a capital margin for an insurance operation is usually based on the statistical distributions of the operation's total assets and liabilities at defined

points of time. It is usual for the claim liabilities to be estimated business segment by business segment, where the segments may be lines of business, or subsets of these.

It is rare that all of the segments will be considered stochastically independent, and so their aggregation within a distribution of total liabilities requires some information on the dependencies between them. The segments may be numerous; it would not be unusual for those recognised by a large insurer to number 100 or so.

If each of these is represented by just a single claim triangle of even modest  $10 \times 10$  dimension (i.e., 55 observations), then one requires a dependency structure that extends over thousands of observations. To the extent that one might wish to represent dependencies by correlations (this is often the case; see, for instance, IAA, 2009), the number of independent entries in the correlation matrix would be of the order of  $10^6$  or more.

The selection of these individually is clearly impractical. Moreover, piecemeal selection may not produce a positive definite matrix. It follows that one must construct some underlying structure that will generate the required correlation matrix, or other dependency representation, with a manageable number of parameters.

## 1.2. Literature review

The literature considering dependencies between claim triangles is relatively recent and small, but is developing rapidly.

The earliest papers considered cell-wise dependence, wherein dependency exists between the  $(i, j)$  cells of different triangles, i.e., between cells with identical locations in different triangles, but not otherwise (Braun, 2004; Hess *et al.*, 2006; Taylor and McGuire, 2007). This structure was continued by a number of subsequent authors (Zhang, 2010; Shi and Frees, 2011; Zhang and Dukic, 2013; Avanzi *et al.*, 2016b).

One might reasonably ask why dependency would exist between cell  $(i, j)$  of triangle  $m$  and cell  $(i, j)$  of triangle  $n$ , but not between cell  $(i, j)$  of triangle  $m$  and not say cell  $(i, j + 1)$  of triangle  $n$ . For this reason, other dependency structures are examined in the literature.

Happ *et al.* (2014) introduced row-wise dependence (dependence between row  $i$  of triangle  $m$  and row  $i$  of triangle  $n$ ), and diagonal-wise dependence was examined by De Jong (2012) and Abdallah *et al.* (2015).

Some authors (Merz *et al.*, 2013; Shi, 2014; Wüthrich and Merz, 2015; Abdallah *et al.*, 2016) have considered more general structures, incorporating multiple dependencies, e.g., row, column and diagonal. Shi *et al.* (2012) introduced time series dependency across diagonals.

All of the above references deal with the dependency across multiple triangles, and some both within and across triangles. The literature contains a number of others concerned with only dependency within a single triangle.

In some cases, dependency structures, such as correlation matrices or copulas, are introduced with many free parameters, but no indication of the means of estimation. This does not invalidate the models, but leaves the user with a substantial estimation problem. The issue of parameter reduction is sometimes mentioned in the literature (e.g., Wüthrich and Merz, 2015, p. 38), but solutions are rarely provided. This problem is explicitly considered and addressed in the current paper.

It is worth noting that issues with excessive parameter numbers also occur in the area of financial portfolio construction (for instance, specification of a variance-covariance matrix in a mean-variance portfolio optimisation problem). In the finance literature, broadly speaking, the covariance matrix is often constructed via so called “factor models” (see, for example, Sharpe, 1963; Luenberger, 2014). These are similar to the “common shock models” in this paper. Naturally, the detailed dependence to be introduced via such constructions is specific to the problem at hand, and the main considerations will differ significantly between problems in asset returns modelling and claim loss triangles as discussed in this paper.

**Remark 1.1.** *In Avanzi et al. (2016b), we illustrated how careful modelling could explain most of the dependencies appearing in claims triangles, thus, wiping off most of the apparent correlations. This was also observed by others and with different data sets; see, e.g., Meyers (2016).*

*However, we also emphasized the fact that this did not relieve us of the need to specify dependencies for future development of claims, which, due to the effect of random processes (e.g., inflation), would still display some dependencies.*

### 1.3. Paper aims and structure

The objective of the current paper is to produce a flexible family of models that:

- incorporate dependency both within and across triangles (actually, more general arrays than triangles);
- include row, column and diagonal dependence, as well as more general forms of dependency;
- include dependencies of the time series type, in which the strength of dependency (e.g., diagonal) declines as the distance between diagonals increases;
- formulate these dependencies in such a manner as to restrict the number of parameters involved, and therefore requiring estimation.

A generalisation of common shock models (e.g., Meyers, 2007; Wüthrich and Merz, 2015; Avanzi et al., 2016a) is used to achieve these objectives. Our paper generalises Wüthrich and Merz (2015, e.g., first equation on p. 210), who use a Bayesian model with log normal likelihood, by being model-independent. Specifically, we move away from the Bayesian and distributional framework therein, and replace it with distribution-free common shocks. Emphasis is placed on the algebraic formulation of the models. Although an objective is to minimise the number of parameters, there is little emphasis on their statistical

estimation. However, we discuss how large correlation matrices can be otherwise specified in a practical setting.

Section 2 establishes the mathematical framework. Section 3 defines the family of models, and Section 4 discusses a number of matters relevant to them. Section 5 comments briefly on parameter estimation, with special emphasis on heuristic estimation, possibly guesswork, which is of especial interest to practitioners, and a numerical example is provided in Section 5.3. Concluding comments are found in Section 6.

## 2. FRAMEWORK AND NOTATION

The present paper will be concerned with insurance *claim arrays*, which are generalisations of *claim triangles*, the latter as in standard texts on loss reserving (Taylor, 2000; Wüthrich and Merz, 2008). The generalisations can be found, for example, in Kuang *et al.* (2008) (called “index sets” in their terminology) and Taylor (2017).

A claim array  $\mathcal{A}$  will be defined here as a 2-dimensional array of random variables  $X_{ij}$ , indexed by integers  $i, j$ , with  $1 \leq i \leq I, 1 \leq j \leq J$  for some fixed integers  $I, J$ . For any given pair  $i, j$ , the random variable  $X_{ij}$  may or may not be present.

The subscripts  $i, j$  typically index accident period (row) and development period (column), respectively, and the  $X_{ij}$  represent observations on claims, commonly claim counts or amounts. In the special case  $I = J$  and  $\mathcal{A} = \{X_{ij} : 1 \leq i \leq I, 1 \leq j \leq I - i + 1\}$ , the array reduces to the well-known claim triangle.

Define  $t = i + j - 1$ , so that  $t = 1, 2, \dots, I + J - 1$ . Observations with common  $t$  lie on the  $t$ th diagonal of  $\mathcal{A}$ .

Subsequent sections will often involve the simultaneous consideration of *multiple business segments* with one array for each segment. A segment could be a *line of business*.

It will be necessary in this case to consider a collection  $\mathbb{A} = \{\mathcal{A}^{(n)}, n = 1, 2, \dots, N\}$  of claim arrays, where  $\mathcal{A}^{(n)}$  denotes the array for segment  $n$ . It will be assumed that all  $\mathcal{A}^{(n)}$  are *congruent*, i.e., are of the same dimensions  $I, J$ , and that they have missing observations in the same  $i, j$  locations.

The  $i, j$  observation of  $\mathcal{A}^{(n)}$  will be denoted  $X_{ij}^{(n)}$ ; the entire  $i$ th row of  $\mathcal{A}^{(n)}$  denoted  $\mathcal{R}_i^{(n)}$ ; and the entire  $j$ th column  $\mathcal{C}_j^{(n)}$ .

It will also be useful to consider *diagonals* of  $\mathcal{A}^{(n)}$ , where the  $t$ th diagonal is defined as the subset  $\{X_{ij}^{(n)} \in \mathcal{A}^{(n)} : i + j - 1 = t\}$ , and represents claim observations from the  $t$ th calendar period,  $t = 1$  denoting the calendar period in which the first accident period falls. The entire  $t$ th diagonal of  $\mathcal{A}^{(n)}$  will be denoted  $\mathcal{D}_t^{(n)}$ .

When only a single array  $\mathcal{A}$  is under consideration, the suffix  $n$  will be omitted, yielding the notation  $\mathcal{R}_i, \mathcal{C}_j, \mathcal{D}_t$ .

An array, under this general definition, may refer to past or future observations. Specification of a correlation structure in relation to past observations may be required for the modelling of them; a correlation structure in relation to future observations may be required for the forecast of them.

It will be convenient to arrange the observations  $X_{ij}^{(n)}$  as a vector  $X^{(n)}$ . For the purpose of the following sections, it will be assumed that, for fixed  $n$ , the  $X_{ij}^{(n)}$  are arranged in dictionary order within  $X^{(n)}$ . However, the order is fundamentally irrelevant, and the following sections may be easily adapted to any order of the  $X_{ij}^{(n)}$  within  $X^{(n)}$ .

Let  $I^{(N)}$  denote the  $N \times N$  identity matrix, and let  $I^{(\bullet)}$  denote the identity matrix with dimension equal to the number of observations in each array  $\mathcal{A}^{(n)}$ .

Henceforth,  $X_p^{(n)}$ , with a single subscript, will be understood to denote the  $p$ th component of the vector  $X^{(n)}$ . It is, of course, identical to  $X_{ij}^{(n)}$  for some particular  $i, j$ . The array  $\mathcal{A}^{(n)}$  will then be said to have been represented in *array vector notation*.

The vectors  $X^{(n)}$  may now be stacked into a vector  $X$ :

$$X = \begin{bmatrix} X^{(1)} \\ \vdots \\ X^{(N)} \end{bmatrix}.$$

Similarly, any other array  $\mathcal{B}^{(n)} = \{Y_{ij}^{(n)}\}$  may be represented without comment in following sections in array vector notation  $Y^{(n)} = \{Y_p^{(n)}\}$ , and the  $Y^{(n)} = \{Y_p^{(n)}\}$  may be stacked into a vector  $Y$  without comment.

For a vector  $Y$ ,  $\mathbb{D}(Y^{(n)})$  will denote the diagonal matrix  $\text{diag}(Y_1^{(n)}, Y_2^{(n)}, \dots)$ , and  $\mathbb{D}(Y)$  will denote the diagonal matrix constructed as diagonal block matrix with blocks  $\mathbb{D}(Y^{(n)})$ .

Further, for a set of constants  $c^{(q)}$ ,  $q = 1, \dots, Q$ , let  $\mathbb{D}(c^{(\bullet)})$  denote the matrix  $\text{diag}(c^{(1)}, \dots, c^{(Q)})$ . The quantity  $Q$  will assume different meanings in different circumstances; for example, it might denote the number of observations in an array, the number of rows in an array, etc., depending on context.

For the same vector  $Y$ ,  $\mathbb{B}(Y)$  will denote the block diagonal matrix as

$$\mathbb{B}(Y) = \begin{bmatrix} Y^{(1)} & \dots & \dots & \dots \\ \vdots & Y^{(2)} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & Y^{(N)} \end{bmatrix}.$$

Thus, three representations of the vector  $Y$  have been defined, namely  $Y$  itself, the diagonal matrix  $\mathbb{D}(Y)$  and the block diagonal matrix  $\mathbb{B}(Y)$ .

For a set of matrices of constants  $M^{(n)}$ ,  $n = 1, \dots, N$ , let  $\mathbb{B}(M^{(\bullet)})$  denote the block diagonal matrix whose block  $(n, n)$  is matrix  $M^{(n)}$ .

Define  $\Delta_{\mathcal{R}_i}$  as the diagonal matrix whose  $(p, p)$  element is

$$(e_{\mathcal{R}_i})_{pp} = \begin{cases} 1 & \text{if } Y_p^{(n)} \in \mathcal{R}_i^{(n)} \text{ for some } n; \\ 0 & \text{otherwise;} \end{cases}$$

i.e., the components of the vector  $\Delta_{\mathcal{R}_i} Y^{(n)}$  are  $Y_p^{(n)}$  whenever  $Y_p^{(n)} \in \mathcal{R}_i^{(n)}$  for some  $n$ , and zero otherwise. Note that  $\Delta_{\mathcal{R}_i}$  does not depend on  $n$  because of the congruence of the  $\mathcal{A}^{(n)}$ .

Similarly,  $\Delta_{\mathcal{C}_j}$  and  $\Delta_{\mathcal{D}_t}$  denote matrices that select the  $j$ th column and the  $t$ th diagonal, respectively, of the arrays  $\mathcal{A}^{(n)}$ ,  $n = 1, \dots, N$ .

### 3. COMMON SHOCK MODELS

#### 3.1. Introduction

The common shock model was adapted to claim triangles in some papers such as De Jong (2006, 2012) or Avanzi *et al.* (2016a). In its most basic form, the model for two variates  $A, B$ , subject to common shock, is as follows:

$$A = \alpha_A W + \beta_A Z_A, \tag{3.1}$$

$$B = \alpha_B W + \beta_B Z_B, \tag{3.2}$$

where  $W$  is a non-degenerate random variable representing the common shock,  $Z_A, Z_B$  are random variables contributing to  $A, B$ , respectively, and  $\alpha_A, \beta_A, \alpha_B, \beta_B \geq 0$  are constants. The component random variables  $W, Z_A, Z_B$  are mutually independent.

It is evident that the common shock structure induces positive dependency between  $A$  and  $B$ . Indeed, it follows from (3.1) and (3.2) that

$$\text{Cov}[A, B] = \alpha_A \alpha_B \sigma_W^2 \geq 0, \text{ and hence} \tag{3.3}$$

$$\text{Corr}[A, B] = \left[ 1 + \frac{\beta_A^2 \sigma_A^2}{\alpha_A^2 \sigma_W^2} \right]^{-\frac{1}{2}} \left[ 1 + \frac{\beta_B^2 \sigma_B^2}{\alpha_B^2 \sigma_W^2} \right]^{-\frac{1}{2}} \geq 0, \tag{3.4}$$

where  $\sigma_W^2 = \text{Var}[W]$ ,  $\sigma_A^2 = \text{Var}[Z_A]$ , and where  $\sigma_B^2 = \text{Var}[Z_B]$ .

Models (3.1) and (3.2) will be referred to as a *linear common shock* model; strictly, *bivariate linear common shock* model, as there will be occasion to introduce a more general multivariate version later.

There is an alternative form of common shock which will prove useful in relation to claim arrays. This is the *log-linear common shock* model, defined as follows:

$$A = \exp(\alpha_A W + \beta_A Z_A), \tag{3.5}$$

$$B = \exp(\alpha_B W + \beta_B Z_B). \tag{3.6}$$

The expression of covariance of  $A$  and  $B$  in terms of the moments of  $W$ ,  $Z_A$ ,  $Z_B$  requires that these variates be equipped with distributions. Assume for the remainder of this subsection that all are normally distributed (such as in Wüthrich and Merz, 2015, for instance):

$$W \sim N(\mu_W, \sigma_W^2), \quad Z_A \sim N(\mu_A, \sigma_A^2), \quad Z_B \sim N(\mu_B, \sigma_B^2).$$

By elementary calculations,

$$\text{Var}[I] = \mathbb{E}^2[I] \{ \exp(\alpha_I^2 \sigma_W^2 + \beta_I^2 \sigma_I^2) - 1 \} \text{ for } I = A, B, \text{ and} \quad (3.7)$$

$$\text{Cov}[A, B] = \mathbb{E}[A]\mathbb{E}[B] \{ \exp(\alpha_A \alpha_B \sigma_W^2) - 1 \}, \quad (3.8)$$

where

$$\mathbb{E}[I] = \exp \left\{ (\alpha_I \mu_W + \beta_I \mu_I) + \frac{1}{2} (\alpha_I^2 \sigma_W^2 + \beta_I^2 \sigma_I^2) \right\}. \quad (3.9)$$

It then follows that

$$\text{Corr}[A, B] = \frac{\exp[(\alpha_A \sigma_W)(\alpha_B \sigma_W)] - 1}{[\exp(\alpha_A \sigma_W)^2 \exp(\beta_A \sigma_A)^2 - 1]^{\frac{1}{2}} [\exp(\alpha_B \sigma_W)^2 \exp(\beta_B \sigma_B)^2 - 1]^{\frac{1}{2}}}. \quad (3.10)$$

### 3.2. A general model of dependency between corresponding array locations

The present section will be concerned with dependencies between observations in different regions of a single array. At first, for completeness, a very general dependency structure will be defined, but the focus will turn very quickly to special cases of practical interest.

Let  $\mathcal{P}^{(n)}$  be a partition of  $\mathcal{A}^{(n)} \in \mathbb{A}$ , i.e.,  $\mathcal{P}^{(n)} = \{\mathcal{P}_1^{(n)}, \dots, \mathcal{P}_P^{(n)}\}$ , where the  $\mathcal{P}_p^{(n)}$  are subsets of  $\mathcal{A}^{(n)}$  with  $\mathcal{P}_p^{(n)} \cap \mathcal{P}_q^{(n)} = \emptyset$  for all  $p, q = 1, \dots, P$ ,  $p \neq q$  and  $\bigcup_{p=1}^P \mathcal{P}_p^{(n)} = \mathcal{A}^{(n)}$ . Suppose that all partitions are the same in the sense that, for each  $p$ , the  $(i, j)$  positions of the elements of  $\mathcal{A}^{(n)}$  included in  $\mathcal{P}_p^{(n)}$  are the same for different  $n$ .

Now consider the following dependency structure on the elements  $X_{ij}^{(n)}$ :

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} W_{\pi(i,j)} + \beta_{ij}^{(n)} W_{\pi(i,j)}^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \quad (3.11)$$

where  $\pi(i, j) = p$  such that  $X_{i,j}^{(n)} \in \mathcal{P}_p^{(n)}$ , a unique mapping;  $W_{\pi(i,j)}$ ,  $W_{\pi(i,j)}^{(n)}$ ,  $Z_{ij}^{(n)}$  are independent stochastic variates, and  $\alpha_{ij}^{(n)}$ ,  $\beta_{ij}^{(n)}$ ,  $\phi_{ij}^{(n)} \geq 0$  are fixed and known constants.

It is evident that  $W_p$  is a common shock across all  $n$ , but affecting only subsets  $\mathcal{P}_p^{(n)}$  for fixed  $p$ ;  $W_p^{(n)}$  is similarly a common shock across  $\mathcal{P}_p^{(n)}$ , but now for fixed  $n$  and  $p$ ; and  $Z_{ij}^{(n)}$  is an idiosyncratic component of  $X_{ij}^{(n)}$ , specific to  $i, j$ .

The common shock  $W_p^{(n)}$  creates dependency between observations within the subset  $\mathcal{P}_p^{(n)}$  of array  $\mathcal{A}^{(n)}$ . Since the partitions  $\mathcal{P}^{(n)}$  are the same across  $n$ , the

common shock  $W_p$  creates dependency between observations in the subsets  $\mathcal{P}_p^{(n)}$  of the same or different arrays.

Indeed, the results of Section 3.1 may be extended to the general case (3.11). Let  $\sigma_{W_p}^2 = \text{Var}[W_p]$ ,  $\sigma_{W_p^{(m)}}^2 = \text{Var}[W_p^{(m)}]$ ,  $\sigma_{Z_{ij}^{(n)}}^2 = \text{Var}[Z_{ij}^{(n)}]$ . Then, for the linear common shock model,

$$\begin{aligned} \text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \delta_{\pi(i,j)\pi(k,l)} \alpha_{ij}^{(m)} \alpha_{kl}^{(n)} \sigma_{W_{\pi(i,j)}}^2 \\ &+ \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \beta_{ij}^{(m)} \beta_{kl}^{(m)} \sigma_{W_{\pi(i,j)}^{(m)}}^2 + \delta_{ik} \delta_{jl} \delta_{mn} \left( \phi_{ij}^{(m)} \right)^2 \sigma_{Z_{ij}^{(m)}}^2, \end{aligned} \tag{3.12}$$

where  $\delta_{\#}$  is the Kronecker delta defined such that

$$\delta_{AB} = \begin{cases} 1 & \text{if } A = B; \\ 0 & \text{otherwise.} \end{cases}$$

It will be assumed henceforth that  $\sigma_{W_p}^2, \sigma_{W_p^{(m)}}^2, \sigma_{Z_{ij}^{(n)}}^2 > 0$ . The case in which any one of them is zero can be achieved by setting the relevant  $\alpha, \beta$  or  $\phi$  coefficient to zero in (3.12). Under this assumption,

$$\begin{aligned} \text{Corr} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \frac{\delta_{\pi(i,j)\pi(k,l)} + \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \left( \frac{\beta_{ij}^{(m)}}{\alpha_{ij}^{(m)}} \right)^2 \frac{\sigma_{W_{\pi(i,j)}^{(m)}}^2}{\sigma_{W_{\pi(i,j)}}^2} + \delta_{ik} \delta_{jl} \delta_{mn} \left( \frac{\phi_{ij}^{(m)}}{\alpha_{ij}^{(m)}} \right)^2 \frac{\sigma_{Z_{ij}^{(m)}}^2}{\sigma_{W_{\pi(i,j)}}^2}}{\left[ 1 + \left( \frac{\beta_{ij}^{(m)}}{\alpha_{ij}^{(m)}} \right)^2 \frac{\sigma_{W_{\pi(i,j)}^{(m)}}^2}{\sigma_{W_{\pi(i,j)}}^2} + \left( \frac{\phi_{ij}^{(m)}}{\alpha_{ij}^{(m)}} \right)^2 \frac{\sigma_{Z_{ij}^{(m)}}^2}{\sigma_{W_{\pi(i,j)}}^2} \right]^{\frac{1}{2}} \left[ 1 + \left( \frac{\beta_{kl}^{(n)}}{\alpha_{kl}^{(n)}} \right)^2 \frac{\sigma_{W_{\pi(k,l)}^{(n)}}^2}{\sigma_{W_{\pi(k,l)}}^2} + \left( \frac{\phi_{kl}^{(n)}}{\alpha_{kl}^{(n)}} \right)^2 \frac{\sigma_{Z_{kl}^{(n)}}^2}{\sigma_{W_{\pi(k,l)}}^2} \right]^{\frac{1}{2}}}. \end{aligned} \tag{3.13}$$

In the log-linear common shock model corresponding to (3.8),

$$\begin{aligned} \text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= E \left[ X_{ij}^{(m)} \right] E \left[ X_{kl}^{(n)} \right] \\ &\times \left\{ \exp \left[ \delta_{\pi(i,j)\pi(k,l)} \alpha_{ij}^{(m)} \alpha_{kl}^{(n)} \sigma_{W_{\pi(i,j)}}^2 + \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \left( \beta_{ij}^{(m)} \right)^2 \sigma_{W_{\pi(i,j)}^{(m)}}^2 \right. \right. \\ &\left. \left. + \delta_{ik} \delta_{jl} \delta_{mn} \left( \phi_{ij}^{(m)} \right)^2 \sigma_{Z_{ij}^{(m)}}^2 \right] - 1 \right\}. \end{aligned} \tag{3.14}$$

**Remark 3.1.** *It is evident from (3.11) and the independence of the random components that if  $\alpha_{ij}^{(n)} = 0$ , for all  $i, j, n$ , the model reduces to within-array dependence. In this case,  $\text{Cov}[X_{ij}^{(m)}, X_{kl}^{(n)}] = 0$  for  $m \neq n$ .*



**Remark 3.2.** *It is noteworthy that the parameter  $\sigma_{W_{\pi(i,j)}}^2$  in (3.12) may be absorbed into the associated coefficient by the simple substitution  $\alpha_{ij}^{(m)} \leftarrow \alpha_{ij}^{(m)} \sigma_{W_{\pi(i,j)}}$ ,  $\sigma_{W_{\pi(i,j)}} \leftarrow 1$ . The other  $\sigma$  terms in (3.12) may be treated similarly so that, without loss of generality, (3.12) may be represented in the alternative form*

$$\begin{aligned} \text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \delta_{\pi(i,j)\pi(k,l)} \alpha_{ij}^{(m)} \alpha_{kl}^{(n)} \\ &\quad + \delta_{mn} \delta_{\pi(i,j),\pi(k,l)} \beta_{ij}^{(m)} \beta_{kl}^{(m)} + \delta_{ik} \delta_{jl} \delta_{mn} \left( \phi_{ij}^{(m)} \right)^2. \end{aligned} \tag{3.15}$$

The  $\sigma$  terms may be similarly eliminated from all covariances set out in the remainder of this section.

### 3.3. Special cases of dependency between corresponding array locations

A number of special cases derive directly from (3.11). These are discussed in this subsection.

3.3.1. *Array-wide dependence.* Define  $\mathcal{P}^{(n)}$  to be the minimal partition of  $\mathcal{A}^{(n)}$  in which  $P = 1$ ,  $\mathcal{P}_1^{(n)} = \mathcal{A}^{(n)}$ . Then,  $\pi(i, j) = 1$  for all  $i, j$  in which case the subscript  $\pi(i, j)$  may be omitted from (3.11) and (3.12). Moreover,  $\delta_{\pi(i,j)\pi(k,l)} = 1$ , and the two relations in question reduce to the following:

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} W + \beta_{ij}^{(n)} W^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \tag{3.16}$$

and

$$\text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] = \alpha_{ij}^{(m)} \alpha_{kl}^{(n)} \sigma_W^2 + \delta_{mn} \left[ \beta_{ij}^{(m)} \beta_{kl}^{(n)} \sigma_{W^{(m)}}^2 + \delta_{ik} \delta_{jl} \left( \phi_{ij}^{(m)} \right)^2 \sigma_{Z_{ij}^{(m)}}^2 \right]. \tag{3.17}$$

Let  $\text{Var}[X]$  denote the covariance matrix comprising the quantities on the left side of (3.17), with observations  $X_{ij}^{(n)}$  arranged as a vector  $X$ , as described in Section 2.

In addition, let  $\alpha$ ,  $\beta$  and  $\phi$  denote the vectors with entries  $\alpha_{ij}^{(n)}$ ,  $\beta_{ij}^{(n)}$  and  $\phi_{ij}^{(n)}$ , respectively, arranged in the same order as the  $X_{ij}^{(n)}$  within  $X$ . Now write  $X_{ij}^{(m)}$ ,  $X_{kl}^{(n)}$ ,  $\alpha_{ij}^{(m)}$ ,  $\alpha_{kl}^{(n)}$ ,  $\beta_{ij}^{(m)}$ ,  $\beta_{ij}^{(n)}$ ,  $\phi_{ij}^{(m)}$ ,  $\phi_{ij}^{(n)}$  as  $X_p^{(m)}$ ,  $X_q^{(n)}$ ,  $\alpha_p^{(m)}$ ,  $\alpha_q^{(n)}$ ,  $\beta_p^{(m)}$ ,  $\beta_q^{(n)}$ ,  $\phi_p^{(m)}$ ,  $\phi_q^{(n)}$ , respectively, using the array vector notation established in Section 2.

Then, (3.17) becomes

$$\text{Cov} \left[ X_p^{(m)}, X_q^{(n)} \right] = \alpha_p^{(m)} \alpha_q^{(n)} \sigma_W^2 + \delta_{mn} \left[ \beta_p^{(m)} \beta_q^{(n)} \sigma_{W^{(m)}}^2 + \delta_{pq} \left( \phi_p^{(m)} \right)^2 \sigma_{Z_p^{(m)}}^2 \right]. \tag{3.18}$$

The covariance matrix can then be expressed as

$$\text{Var}[X] = \sigma_W^2 (\alpha \alpha^T) + \mathbb{B} (\sigma_{W^{(\cdot)}} \beta^{(\cdot)}) \mathbb{B}^T (\sigma_{W^{(\cdot)}} \beta^{(\cdot)}) + \mathbb{D}^2(\phi) \mathbb{D} (\sigma_Z^2). \tag{3.19}$$

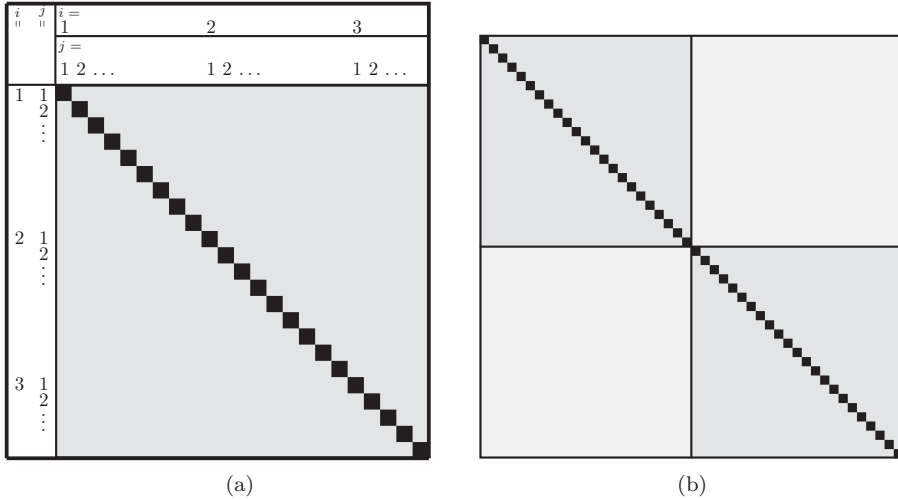


FIGURE 1: Schematic covariance matrix for array-wide dependence ( $AY = i, DY = j$ ). (a)  $N = 1$ . (b)  $N = 2$ .

It is helpful to illustrate covariance (3.18) diagrammatically as in Figure 1. The left half of the diagram illustrates the case  $N = 1$ , where the rows and columns of the covariance matrix are indexed by accident year (AY) and development year (DY). The half right of the diagram illustrates the case  $N = 2$ . For fixed  $n$ , rows and columns appear in the same order as in the case  $N = 1$  but, in view of the small scale, have not been labelled. The figure is schematic only, insensitive to small differences in covariances, but broadly heavier shading (darker grey or black) indicates larger values, with white indicating zero covariance.

3.3.2. *Cell-wise dependence.* Define each of the  $\mathcal{P}_p^{(n)}$  to consist of a single observation, e.g.,  $\mathcal{P}_1^{(n)} = \{X_{11}^{(n)}\}$ ,  $\mathcal{P}_2^{(n)} = \{X_{12}^{(n)}\}$ , etc. Then, for given  $i, j$ ,  $W_{\pi(i,j)}$  appears in (3.11) only for observations  $X_{ij}^{(n)}$ ,  $n = 1, \dots, N$ . Then,  $W_{\pi(i,j)}$  may be denoted  $V_{ij}$ , indicating the dependence on  $i, j$  but not  $n$ . Further,  $W_{\pi(i,j)}^{(n)}$  relates only to the single observation  $X_{ij}^{(n)}$ , and so functions in the same manner as  $Z_{ij}^{(n)}$ . It may therefore be omitted, whereupon (3.11) and (3.12) become

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} V_{ij} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \tag{3.20}$$

and

$$\text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] = \delta_{ik} \delta_{jl} \left[ \alpha_{ij}^{(m)} \alpha_{ij}^{(n)} \sigma_{V_{ij}}^2 + \delta_{mn} \left( \phi_{ij}^{(m)} \right)^2 \sigma_{Z_{ij}^{(m)}}^2 \right], \tag{3.21}$$

where  $\sigma_{V_{ij}}^2 = \text{Var}[V_{ij}]$ , and it is recognised that  $\delta_{\pi(i,j)\pi(k,l)} = \delta_{ik} \delta_{jl}$ . The member of (3.20) involving the  $\alpha$  coefficient creates dependency between individual cells

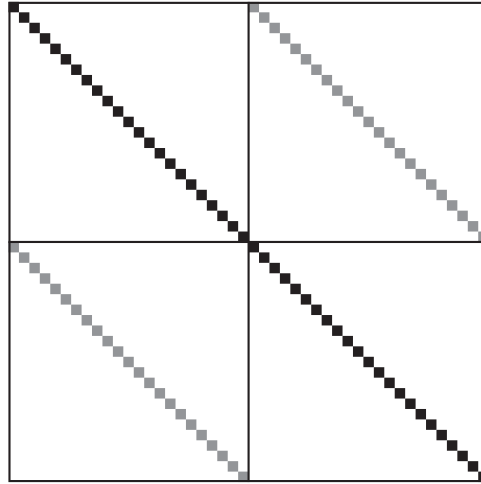


FIGURE 2: Schematic covariance matrix for cell-wise dependence.

in the same location of different arrays. This case can be found in Braun (2004) and Avanzi *et al.* (2016a).

The matrix representation of (3.21) can be given by

$$\text{Var}[X] = C(\alpha)\mathbb{D}(\sigma_{\tilde{r}}^2)C(\alpha)^T + \mathbb{D}^2(\phi)\mathbb{D}(\sigma_Z^2), \quad (3.22)$$

where

$$C(\alpha) = \begin{bmatrix} \mathbb{D}(\alpha^{(1)}) \\ \vdots \\ \mathbb{D}(\alpha^{(N)}) \end{bmatrix}.$$

Figure 2 is the counterpart of the case  $N = 2$  in Figure 1, but this time for covariances in the presence of cell-wise dependence.

**3.3.3. Row-wise dependence.** Define each of the  $\mathcal{P}_p^{(n)}$  to consist of the observations of a single row of  $\mathcal{A}^{(n)}$ , i.e.,  $\mathcal{P}_p^{(n)} = \mathcal{R}_i^{(n)}$ ,  $p = 1, \dots, I$ . Then, for a given  $i$ ,  $W_{\pi(i,j)}$  appears in (3.11) for all observations  $X_{ij}^{(n)}$ ,  $n = 1, \dots, N$ . Then,  $W_{\pi(i,j)}$  may be denoted  $R_i$ , indicating the dependence on  $i$ , but not  $j, n$ . Similarly,  $W_{\pi(i,j)}^{(n)}$  may be denoted  $R_i^{(n)}$ , and (3.11) and (3.12) become

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} R_i + \beta_{ij}^{(n)} R_i^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)} \quad (3.23)$$

and

$$\text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] = \delta_{ik} \left\{ \alpha_{ij}^{(m)} \alpha_{il}^{(n)} \sigma_{R_i}^2 + \delta_{mn} \left[ \beta_{ij}^{(m)} \beta_{il}^{(n)} \sigma_{R_i^{(m)}}^2 + \delta_{jl} \left( \phi_{ij}^{(m)} \right)^2 \sigma_{Z_{ij}^{(m)}}^2 \right] \right\}, \tag{3.24}$$

where  $R_i$  is a stochastic variate, common to all arrays but specific to the  $i$ th row,  $R_i^{(n)}$  is a similar stochastic variate common to observations on the  $i$ th row but specific to array  $\mathcal{A}^{(n)}$ ,  $\sigma_{R_i}^2 = \text{Var}[R_i]$  and  $\sigma_{R_i^{(n)}}^2 = \text{Var}[R_i^{(n)}]$ .

Consider the term  $\delta_{ik} \alpha_{ij}^{(m)} \alpha_{il}^{(n)}$  for fixed  $i$ , i.e.,  $\alpha_{ij}^{(m)} \alpha_{kl}^{(n)}$  whenever  $i = k$ . These, and only these, quantities will appear in the matrix  $\Delta_{\mathcal{R}_i} (\alpha^{(m)} (\alpha^{(n)})^T) \Delta_{\mathcal{R}_i}$ , and so the contribution of the first member of (3.24) to the  $(m, n)$  block of the matrix  $\text{Var}[X]$  is  $\sigma_{R_i}^2 \Delta_{\mathcal{R}_i} (\alpha^{(m)} (\alpha^{(n)})^T) \Delta_{\mathcal{R}_i}$ . Then, the contribution of this first member to the entire matrix  $\text{Var}[X]$  (still for fixed  $i$ ) is  $\sigma_{R_i}^2 (\Delta_{\mathcal{R}_i} \otimes I^{(N)}) (\alpha \alpha^T) (\Delta_{\mathcal{R}_i} \otimes I^{(N)})$ , where

$$\Delta_{\mathcal{R}_i} \otimes I^{(N)} = \begin{bmatrix} \Delta_{\mathcal{R}_i} & \cdots & \cdots & \cdots \\ \vdots & \Delta_{\mathcal{R}_i} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \Delta_{\mathcal{R}_i} \end{bmatrix}. \tag{3.25}$$

The second member of (3.24) (involving  $\beta$ 's) is similar except that it contributes only diagonal blocks to  $\text{Var}[X]$ , and the multipliers  $\sigma_{R_i^{(m)}}^2$  vary over those blocks. It may be noted that  $\mathbb{B}(\sigma_{R_i^{(n)}} \beta)$  is a block diagonal matrix whose block  $(n, n)$  is  $\sigma_{R_i^{(n)}} \beta^{(n)}$ . Then, the contribution to  $\text{Var}[X]$  of the  $\beta$  member of (3.24) is  $\Delta_{\mathcal{R}_i} \mathbb{B}(\sigma_{R_i^{(n)}} \beta^{(n)}) \mathbb{B}^T(\sigma_{R_i^{(n)}} \beta^{(n)}) \Delta_{\mathcal{R}_i}$ .

Collection of these contributions to  $\text{Var}[X]$ , with allowance for variation of  $i$ , converts (3.24) to the form

$$\begin{aligned} \text{Var}[X] &= \sum_{i=1}^I (\Delta_{\mathcal{R}_i} \otimes I^{(N)}) \left[ \sigma_{R_i}^2 (\alpha \alpha^T) + \mathbb{B}(\sigma_{R_i^{(n)}} \beta^{(n)}) \mathbb{B}^T(\sigma_{R_i^{(n)}} \beta^{(n)}) \right] \\ &\quad \times (\Delta_{\mathcal{R}_i} \otimes I^{(N)}) + \mathbb{D}^2(\phi) \mathbb{D}(\sigma_Z^2), \end{aligned} \tag{3.26}$$

where the final member is the same as in (3.22), and where  $\Delta_{\mathcal{R}_i} \otimes I^{(N)}$  is defined in (3.25).

Figure 3 is the counterpart of Figure 1, but this time for covariances in the presence of row-wise dependence (in claim triangles). As in the earlier figure, the left half of the diagram illustrates the case  $N = 1$ , with the same ordering of rows and columns, and the right half illustrates the case  $N = 2$ .

**3.3.4. Column-wise dependence.** Column-wise dependence is constructed in a manner similar to row-wise dependence. In this case  $\mathcal{P}_1^{(n)} = \{X_{11}^{(n)}, X_{21}^{(n)}, \dots\}$ ,

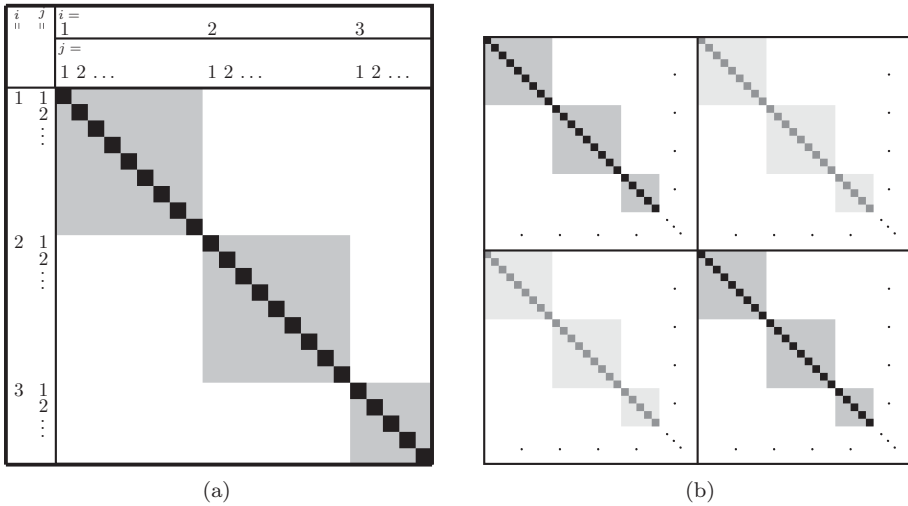


FIGURE 3: Schematic covariance matrix for row-wise dependence ( $AY = i, DY = j$ ). (a)  $N = 1$ . (b)  $N = 2$ .

$\mathcal{P}_2^{(n)} = \{X_{12}^{(n)}, X_{22}^{(n)}, \dots\}$  etc., and (3.11) and (3.12) become

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} C_j + \beta_{ij}^{(n)} C_j^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \tag{3.27}$$

where  $C_j$  is a stochastic variate, common to all arrays but specific to the  $j$ th column,  $C_j^{(n)}$  is a similar stochastic variate, common to observations on the  $j$ th column but specific to array  $\mathcal{A}^{(n)}$ ,  $\sigma_{C_j}^2 = \text{Var}[C_j]$  and  $\sigma_{C_j^{(n)}}^2 = \text{Var}[C_j^{(n)}]$ .

The expression for the covariance  $\text{Cov}[X_{ij}^{(m)}, X_{kl}^{(n)}]$  is similar to that of row-wise dependence in (3.24), except that the indicator function is now used for the selection of cells in a common column  $j = l$ , and common shocks  $C_j$  and  $C_j^{(n)}$  to column  $j$  replace common row shocks  $R_i$  and  $R_i^{(n)}$ . Similarly, the covariance matrix is the same as (3.26) except that rows are replaced by columns. Figure 3, with rows and columns interchanged, illustrates the covariance matrix in this context.

**3.3.5. Diagonal-wise dependence.** Diagonal-wise dependence is constructed in a manner similar to row-wise and column-wise. In this case,  $\mathcal{P}_1^{(n)} = \{X_{11}^{(n)}\}$ ,  $\mathcal{P}_2^{(n)} = \{X_{12}^{(n)}, X_{21}^{(n)}, \dots\}$  etc., with  $\mathcal{P}_t^{(n)}$  consisting of the  $t$ th diagonal, and (3.11) and (3.12) become

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} D_t + \beta_{ij}^{(n)} D_t^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \tag{3.28}$$

where  $D_t$  is a stochastic variate, common to all arrays but specific to the  $t$ th diagonal,  $D_t^{(n)}$  is a similar stochastic variate, common to observations on the

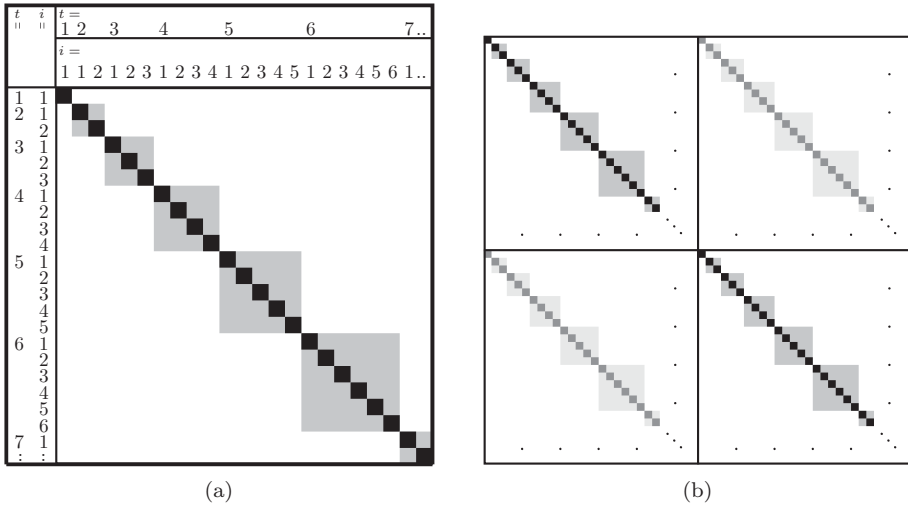


FIGURE 4: Schematic covariance matrix for diagonal-wise dependence ( $AY = i, CY = t$ ). (a)  $N = 1$ . (b)  $N = 2$ .

$t$ th diagonal but specific to array  $\mathcal{A}^{(n)}$ , all variates  $D_t$  and  $D_t^{(n)}$  are independent for all  $t$  and  $n$ ,  $\sigma_{D_t}^2 = \text{Var}[D_t]$  and  $\sigma_{D_t^{(n)}}^2 = \text{Var}[D_t^{(n)}]$ .

The expression for the covariance  $\text{Cov}[X_{ij}^{(n)}, X_{kl}^{(n)}]$  is similar to that of row-wise dependence in (3.24), except that the indicator function is now used for the selection of cells in a common diagonal  $t = i + j - 1 = k + l - 1$ , and common shocks  $D_t$  and  $D_t^{(\cdot)}$  to diagonal  $t$  replace common row shocks  $R_i$  and  $R_i^{(\cdot)}$ . Similarly, the covariance matrix is the same as (3.26) except that rows are replaced by diagonals.

Figure 4 is the counterpart of Figure 1, but now for covariances in the presence of diagonal-wise dependence. As in the earlier figure, the left half of the diagram illustrates the case  $N = 1$ , and the right half  $N = 2$ . Cells have been ordered according to  $AY$  within calendar year ( $CY$ ).

### 3.4. Adding structure with multiple shocks

3.4.1. *Simultaneous array, cell, row, column and diagonal shock.* It is evident that array, cell, row, column and diagonal shocks can be all present simultaneously. One simply adds the cell, row, columns and diagonal effects in (3.16), (3.20), (3.23), (3.27) and (3.28) into a single representation of  $X_{ij}^{(n)}$ , thus,

$$\begin{aligned}
 X_{ij}^{(n)} = & \alpha_{(w)ij}^{(n)} W + \beta_{(w)ij}^{(n)} W^{(n)} + \alpha_{(v)ij}^{(n)} V_{ij} + \alpha_{(r)ij}^{(n)} R_i + \beta_{(r)ij}^{(n)} R_i^{(n)} + \alpha_{(c)ij}^{(n)} C_j \\
 & + \beta_{(c)ij}^{(n)} C_j^{(n)} + \alpha_{(d)ij}^{(n)} D_t + \beta_{(d)ij}^{(n)} D_t^{(n)} + \phi_{ij}^{(n)} Z_{ij}^{(n)}, \quad (3.29)
 \end{aligned}$$

where it is now assumed that all the  $W, W^{(n)}, V_{ij}, R_i, R_i^{(n)}, C_j, C_j^{(n)}, D_t, D_t^{(n)}, Z_{ij}^{(n)}$  are stochastically independent, and the  $\alpha$  and  $\beta$  coefficients are now specific to array, cell, row, column and diagonal shocks.

Because of the independence of the variates appearing on the right side of (3.29), the covariance structure follows easily from variance expressions of array-wide, cell-, row-, column-, diagonal-wise dependence and the result is obtained simply by addition of summands from these relations. Although obvious, it is tedious to write out, and is not given here.

3.4.2. *Shocks creating time series effects.* One may require that the row, column or diagonal effects illustrated in Section 3.3 exhibit time series dependency. For example, the diagonal effect might exhibit AR(1) dependency (see also Wüthrich and Merz, 2015, p. 154). This could be achieved by defining  $D_t$  and  $D_t^{(n)}$  from (3.28) as follows:

$$D_t = \theta D_{t-1} + \varepsilon_t, \quad E[\varepsilon_t] = 0, \quad \text{Var}[\varepsilon_t] = \sigma_{\varepsilon_t}^2, \tag{3.30}$$

$$D_t^{(n)} = \theta^{(n)} D_{t-1}^{(n)} + \varepsilon_t^{(n)}, \quad E[\varepsilon_t^{(n)}] = 0, \quad \text{Var}[\varepsilon_t^{(n)}] = \sigma_{\varepsilon_t^{(n)}}^2, \tag{3.31}$$

for  $t = -v, -v + 1, \dots, 0, 1, 2, \dots$ , where  $v$  is an arbitrary natural number,  $0 < \theta^\# < 1, \sigma_{\varepsilon_t^\#}^2$  constant with  $t$ , all  $\varepsilon_t$  and  $\varepsilon_t^{(n)}$  stochastically independent, and independent of all variates on the right side of (3.29).

For brevity henceforth, let  $D_t^\#$  denote either  $D_t$  or  $D_t^{(n)}$  according to whether  $\#$  takes the value  $(n)$  or is absent. Introduce similar notation for other quantities appearing in (3.30) and (3.31). In this case,

$$D_t^\# = \sum_{u=s+1}^t (\theta^\#)^{t-u} \varepsilon_u + (\theta^\#)^{t-s} D_s^\#, \quad t > s, \tag{3.32}$$

whence

$$\text{Cov} [D_s^\#, D_t^\#] = (\theta^\#)^{t-s} \text{Var} [D_s^\#] \rightarrow (\theta^\#)^{t-s} \frac{\sigma_{\varepsilon_t^\#}^2}{1 - (\theta^\#)^2}, \quad t > s, \quad v \rightarrow \infty. \tag{3.33}$$

If (3.32) is substituted into (3.28), then the covariance is given by the following approximation:

$$\begin{aligned} \text{Cov} [X_{ij}^{(m)}, X_{kl}^{(n)}] &= \alpha_{ij}^{(m)} \alpha_{kl}^{(n)} \theta^{|(i+j)-(k+l)|} \frac{\sigma_\varepsilon^2}{1 - \theta^2} \\ &+ \delta_{mn} \left[ \beta_{ij}^{(m)} \beta_{kl}^{(n)} (\theta^{(m)})^{|(i+j)-(k+l)|} \frac{\sigma_{\varepsilon^{(m)}}^2}{1 - (\theta^{(m)})^2} + \delta_{ik} \delta_{jl} (\phi_{ij}^{(m)})^2 \sigma_{Z_{ij}^{(m)}}^2 \right]. \end{aligned} \tag{3.34}$$

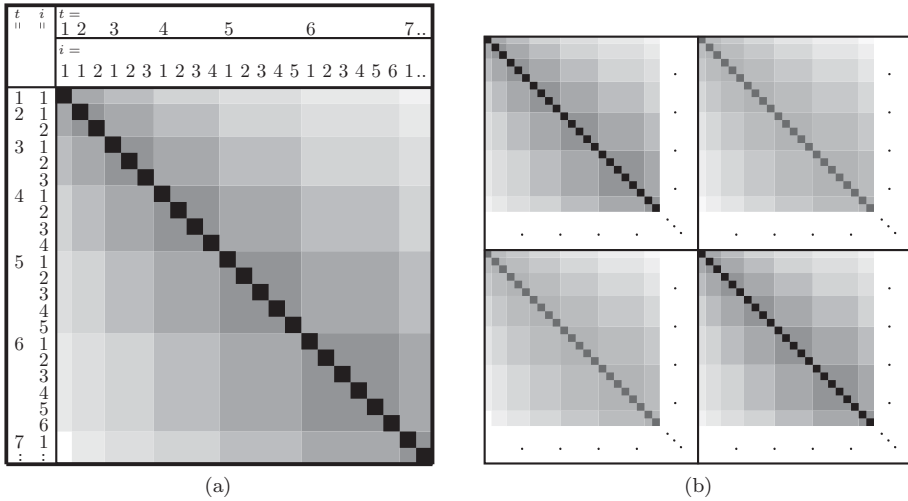


FIGURE 5: Schematic covariance matrix for diagonal-wise AR(1) dependence ( $AY = i, CY = t$ ). (a)  $N = 1$ . (b)  $N = 2$ .

Note that this result is not obtained by direct substitution for  $\sigma_{D_{t+j-1}}^2$  in the covariance expression for diagonal-wise dependence, since  $D_{i+j-1}, D_{k+l-1}$  are no longer independent as in that case.

Figure 5 is the counterpart of Figure 4 but now for the AR(1) case of diagonal-wise dependence. As in the earlier figure, the left half of the diagram illustrates the case  $N = 1$ , and the right half  $N = 2$ . Cells have been ordered according to  $AY$  within  $CY$ .

#### 4. DISCUSSION

##### 4.1. Positive definiteness

The covariance matrices associated with claim arrays are typically large. For an isolated  $n \times n$  triangle, the number of observations is  $n(n+1)/2$ , in which case the number of independent entries in the covariance matrix is  $n(n+1)(n^2+n+2)/8$ . Even in the fairly basic case  $n = 10$ , this amounts to 1,540 independent entries. It is evident that this count increases rapidly as additional triangles, with covariances between them, are added. Indeed, the issue is even larger than this when the claim arrays are modelled for the purpose of loss reserving. The number of covariances per triangle is then  $n^2$ , rather than  $n(n+1)/2$ , with corresponding increasing in dimension of the large covariance matrix.

If covariance matrices of this type are to be constructed for analysis of multiple business segments, it is necessary to ensure that they are at least non-negative definite, and, for all practical purposes, positive definite. Undisciplined, or piecemeal, construction may not achieve this.



It is useful, therefore, to note that the common shock models are assured of producing positive definite covariance structures. This occurs simply because the models formulate all observations as functions of other latent random variables, in which case any vector of observations must have a genuine (non-negative definite) covariance structure. In the absence of zero-variance latent variables, the covariance matrix of the vector of observations must be positive definite.

## 4.2. Model coefficient

Now consider a simple model such as (3.16). No word has been said about the nature of the coefficients  $\alpha_{ij}^{(n)}, \beta_{ij}^{(n)}, \phi_{ij}^{(n)} \geq 0$  other than to note that they are fixed and known constants. Their purpose will be discussed further in the next couple of sub-sections. The discussion will extrapolate readily to other models in Section 3.

4.2.1. *Compatibility of common shock and idiosyncratic components in linear common shock models.* One possible purpose of the coefficients is the adjustment of members of (3.16) to produce a tractable distribution of  $X_{ij}^{(n)}$ . An example of this occurs in Avanzi *et al.* (2016a), in which the components of (3.20) are Tweedie distributed. They define a special case of the model of cell-wise dependence in (3.16) (see their equations (2.11) and (2.19)), in which  $\beta_{ij}^{(n)} = 0, \phi_{ij}^{(n)} = 1$ :

$$X_{ij}^{(n)} = \alpha_{ij}^{(n)} W + Z_{ij}^{(n)}. \quad (4.1)$$

In that case, it was assumed that

$$W \sim \text{Tweedie}_p(\xi, \eta), \quad (4.2)$$

$$Z_{ij}^{(n)} \sim \text{Tweedie}_p(\kappa_i^{(n)} v_j^{(n)}, \phi^{(n)}), \quad (4.3)$$

where  $\text{Tweedie}_p(\mu, \phi)$  denotes the member of the Tweedie family with mean  $\mu$  and variance  $\phi\mu^p$  (Tweedie, 1984).

It is shown in Jørgensen (1997, ch. 3 and 4) that

$$\alpha_{ij}^{(n)} W \sim \text{Tweedie}_p\left(\alpha_{ij}^{(n)} \xi, \left[\alpha_{ij}^{(n)}\right]^{2-p} \eta\right).$$

It then follows also from Jørgensen (1997, ch. 3 and 4) that, for  $X_{ij}^{(n)}$  defined by (4.1) and (4.3),

$$X_{ij}^{(n)} \sim \text{Tweedie}_p\left(\alpha_{ij}^{(n)} \xi + \kappa_i^{(n)} v_j^{(n)}, \phi^*\right),$$

$$\text{with } \phi^* = \left(\alpha_{ij}^{(n)} \eta / \xi^{1-p}\right) \left(\alpha_{ij}^{(n)} \xi + \kappa_i^{(n)} v_j^{(n)}\right)^{1-p},$$

provided that  $\alpha_{ij}^{(n)} \eta / \xi^{1-p} = \phi^{(n)} / \left( \kappa_i^{(n)} \nu_j^{(n)} \right)^{1-p}$ , i.e.,

$$\alpha_{ij}^{(n)} = \left( \frac{\xi}{\kappa_i^{(n)} \nu_j^{(n)}} \right)^{1-p} \frac{\phi^{(n)}}{\eta}. \tag{4.4}$$

Even though other choices may be allowed for some special cases of the Tweedie family, this choice of  $\alpha_{ij}^{(n)}$  will always ensure that a neat Tweedie distribution is obtained for  $X_{ij}^{(n)}$ . Thus, the coefficients  $\alpha_{ij}^{(n)}$  may be used to scale  $W$  relative to  $Z_{ij}^{(n)}$  in order to achieve this distributional result.

4.2.2. *Variation of common shock magnitude over columns in linear common shock models.* In a typical array, the *column factors*  $\nu_j^{(n)}$  peak at some low value of  $j$ , and decline over higher values. These factors at high  $j$  may become small relative to the peak value and other values at low  $j$ .

In this situation, it may be desirable for  $\alpha_{ij}^{(n)}$  to vary over  $i, j$  in a manner similar to  $\kappa_i^{(n)} \nu_j^{(n)}$ . Consider the alternative, e.g.,  $\alpha_{ij}^{(n)} = 1$  so that (4.1) yields

$$E \left[ X_{ij}^{(n)} \right] = E[W] + E[Z_{ij}^{(n)}] = E[W] + \kappa_i^{(n)} \nu_j^{(n)}.$$

In case of an  $E[W]$  comparable with the larger values of  $\kappa_i^{(n)} \nu_j^{(n)}$ ,  $W$  will swamp the cells with smaller values of  $\kappa_i^{(n)} \nu_j^{(n)}$ . If this is corrected by setting  $E[W]$  comparable with the smaller values of  $\kappa_i^{(n)} \nu_j^{(n)}$ , then  $W$  will contribute almost immaterially to the cells with larger values of  $\kappa_i^{(n)} \nu_j^{(n)}$ , creating low dependency, and the purpose of the common shock for generating dependency will be lost.

One possible course of action to rectify this situation is to form rough estimates  $\tilde{\kappa}_i^{(n)}, \tilde{\nu}_j^{(n)}$  of the  $\kappa_i^{(n)}, \nu_j^{(n)}$ . A simple alternative would be to adopt those estimates  $\tilde{\kappa}_i^{(n)}, \tilde{\nu}_j^{(n)}$  obtained in the case  $\alpha_{ij}^{(n)} = 0$ . This is the case of no dependency, and so the estimates are obtained by application of the chain ladder separately to each claim array  $1, \dots, N$ .

Once the  $\tilde{\kappa}_i^{(n)}, \tilde{\nu}_j^{(n)}$  have been found, it is possible to set, for example,

$$\alpha_{ij}^{(n)} = \tilde{\kappa}_i^{(n)} \tilde{\nu}_j^{(n)}, \tag{4.5}$$

in which case

$$\begin{aligned} E[X_{ij}^{(n)}] &= \alpha_{ij}^{(n)} E[W] + \kappa_i^{(n)} \nu_j^{(n)} = \xi \tilde{\kappa}_i^{(n)} \tilde{\nu}_j^{(n)} + \kappa_i^{(n)} \nu_j^{(n)} \\ &= \kappa_i^{(n)} \nu_j^{(n)} \left[ 1 + \xi \frac{\tilde{\kappa}_i^{(n)} \tilde{\nu}_j^{(n)}}{\kappa_i^{(n)} \nu_j^{(n)}} \right], \end{aligned} \tag{4.6}$$

and the common shock term (the second member inside the square bracket) is seen to vary in rough proportion with the idiosyncratic term over the entire array.

4.2.3. *Log-linear common shock models.* Return to the log-linear common shock model, whose prototype appears as (3.5) and (3.6), and take the log-linear parallel of (3.11):

$$\ln X_{ij}^{(n)} = \alpha_{ij}^{(n)} \ln W_{\pi(i,j)} + \beta_{ij}^{(n)} \ln W_{\pi(i,j)}^{(n)} + \phi_{ij}^{(n)} \ln Z_{ij}^{(n)}, \quad (4.7)$$

where

$$\ln W_{\pi(i,j)} \sim N\left(\mu_{W_{\pi(i,j)}}, \sigma_{W_{\pi(i,j)}}^2\right), \quad (4.8)$$

$$\ln W_{\pi(i,j)}^{(n)} \sim N\left(\mu_{W_{\pi(i,j)}^{(n)}}, \sigma_{W_{\pi(i,j)}^{(n)}}^2\right), \quad (4.9)$$

$$\ln Z_{ij}^{(n)} \sim N\left(\mu_{Z_{ij}^{(n)}}, \sigma_{Z_{ij}^{(n)}}^2\right). \quad (4.10)$$

The covariance structure is given by (3.14). One may also note the cell mean

$$\begin{aligned} E[X_{ij}^{(n)}] &= \exp\left\{\alpha_{ij}^{(n)} \mu_{W_{\pi(i,j)}} + \frac{1}{2} \left(\alpha_{ij}^{(n)}\right)^2 \sigma_{W_{\pi(i,j)}}^2\right\} \\ &\times \exp\left\{\beta_{ij}^{(n)} \mu_{W_{\pi(i,j)}^{(n)}} + \frac{1}{2} \left(\beta_{ij}^{(n)}\right)^2 \sigma_{W_{\pi(i,j)}^{(n)}}^2\right\} \times \exp\left\{\phi_{ij}^{(n)} \mu_{Z_{ij}^{(n)}} + \frac{1}{2} \left(\phi_{ij}^{(n)}\right)^2 \sigma_{Z_{ij}^{(n)}}^2\right\}. \end{aligned} \quad (4.11)$$

A noteworthy special case is that discussed in Section 3.3.2 in which the  $\mathcal{P}_p^{(n)}$  consist of a single observation, e.g.,  $\mathcal{P}_1^{(n)} = \{X_{11}^{(n)}\}$ ,  $\mathcal{P}_2^{(n)} = \{X_{12}^{(n)}\}$ , etc., and the following additional properties hold:

$$\beta_{ij}^{(n)} = 0 \text{ for all } i, j, n, \quad (4.12)$$

$$\phi_{ij}^{(n)} = 1 \text{ for all } i, j, n, \quad (4.13)$$

$$E[Z_{ij}^{(n)}] = \kappa_i^{(n)} \nu_j^{(n)}. \quad (4.14)$$

One property of this model is that a cell mean (4.11) takes the multiplicative form

$$E[X_{ij}^{(n)}] = \gamma_{ij}^{(n)} \kappa_i^{(n)} \nu_j^{(n)}, \quad (4.15)$$

where  $\gamma_{ij}^{(n)}$  denotes the first of the three factors on the right side of (4.11) and, as noted in Section 3.3.2, is dependent on  $i, j, n$ .

If the  $\gamma_{ij}^{(n)}$  term is removed (i.e.,  $\alpha_{ij}^{(n)} = 0$ ), then (4.15) reduces to cross-classified form of the chain ladder, as in Renshaw (1989) and many subsequent

papers, with independence between all arrays. The  $\gamma_{ij}^{(n)}$  factor is induced by  $W_{\pi(i,j)}$  — the common shock that creates the dependency between the arrays since  $W_{\pi(i,j)}$  is the same for all  $n$ . This produces a multiplicative version of the additive cell-wise dependency model of Section 3.3.2.

Similar multiplicative versions of the other models of Section 3.3 can be developed. As a single example, set  $\pi(i, j)$  to be the  $(i + j - 1)$ th diagonal in (4.7), as the diagonal-wise dependence examples of Section 3.3.5 and Section 3.4.2.

Assume, further that (4.12)–(4.14) continue to hold and that  $\alpha_{ij}^{(n)}$  depends only on  $i + j$ . In this case, (4.15) is replaced by the following

$$E[X_{ij}^{(n)}] = \gamma_i^{(n)} \kappa_i^{(n)} \nu_j^{(n)}, \quad (4.16)$$

which is a generalisation of the chain ladder model that takes account of diagonal effects as well as row and column, and with diagonal-wise dependency between arrays.

## 5. ESTIMATION OF DEPENDENCE STRUCTURES IN A PRACTICAL SETTING

While parameter estimation should be based on data wherever possible, there may be some forecast models for which data-based estimation is difficult. This is particularly true of dependency models, which may involve numerous parameters. Consider *formal estimation*, the rigorous statistical process by which parameters estimates are derived from a data set. As explained in Section 1, such estimation, for the models formulated in Section 3, is not the primary purpose of the present paper.

Instead, one of the prime objectives of this paper is to address the practical issue of specifying large correlation matrices for the modelling needs of practice, in a way that is as disciplined and rigorous as possible. In this section, we explain how the models developed in the previous section can be used to achieve that aim.

Notwithstanding the abovementioned goal, we start with a brief discussion of formal estimation dealt with in the literature in Section 5.1. Section 5.2 provides a practical and parsimonious heuristic estimation procedure, which is illustrated in Section 5.3.

### 5.1. Formal estimation

First, let us consider linear common shock models. Section 4.2.1 mentions Avanzi *et al.* (2016a) in connection with the model (4.1). Section 4 of the same paper discussed parameter estimation by means of Markov Chain Monte Carlo (MCMC) when the model parameters are random subject to prior distributions. Sometimes, these estimates can be made to approach non-Bayesian (fixed effects) estimates by allowing the priors to approach to uninformative forms, although this strategy may be defeated by difficulties with convergence. Other

linear common shock models, such as those considered in Section 3, can be dealt with similarly, subject again to timely convergence of the MCMC algorithm.

The case of log-linear common shock models is considered in detail in Wüthrich and Merz (2015, ch. 5). In particular, their Corollary 5.3 deals with parameter estimation.

## 5.2. Heuristic estimation

Irrespective of whether (formal) estimation difficulties arise, many practitioners base their dependency models on educated guesswork. The present section is intended to provide comment that will assist this heuristic estimation of parameters in the models of Section 3.

Parameter estimation will vary according to the form of model under consideration, and it is difficult to formulate a fully general protocol. However, reasonable generality may be obtained by considering first the general model of Section 3.2, and then grafting on time series dependencies of the sort introduced in Section 3.4.2.

It will be assumed in the following that each array has already been thoroughly modelled in isolation and that, as a consequence estimates of  $\text{Var}[X_{ij}^{(n)}]$  are available for all  $i, j, n$ . The estimation of only covariances now requires consideration.

Consider the model (3.11), with covariance structure (3.12). It is timely to issue a reminder here that (3.12) decomposes the covariance into three components (in order):

- within-subset covariance (subsets defined in Section 3.2 by the function  $\pi(., .)$ );
- within-subset within-segment covariance;
- variance of individual observations.

For formal expression of this, recall covariance (3.15), a special case of which is

$$\text{Var}[X_{ij}^{(m)}] = V_{ij(\text{sub})}^{(m)} + V_{ij(\text{seg})}^{(m)} + V_{ij(\text{cell})}^{(m)}, \quad (5.1)$$

where

$$V_{ij(\text{sub})}^{(m)} = \left(\alpha_{ij}^{(m)}\right)^2, \text{ relating to the subset that corresponds to } \pi(i, j),$$

$$V_{ij(\text{seg})}^{(m)} = \left(\beta_{ij}^{(m)}\right)^2, \text{ relating to the subset that corresponds to } \pi(i, j),$$

within segment  $m$ ,

$$V_{ij(\text{cell})}^{(m)} = \left(\phi_{ij}^{(m)}\right)^2, \text{ relating to cell } (i, j) \text{ within segment } m.$$

Then, (3.15) may be re-expressed in the form

$$\begin{aligned} \text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \delta_{\pi(i,j)\pi(k,l)} \left( V_{ij(\text{sub})}^{(m)} V_{kl(\text{sub})}^{(n)} \right)^{\frac{1}{2}} \\ &\quad + \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \left( V_{ij(\text{seg})}^{(m)} V_{kl(\text{seg})}^{(n)} \right)^{\frac{1}{2}} + \delta_{ik} \delta_{jl} \delta_{mn} V_{ij(\text{cell})}^{(m)}, \end{aligned} \tag{5.2}$$

or, equivalently, with due allowance for the Kronecker deltas,

$$\begin{aligned} \text{Cov} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \left\{ \text{Var}[X_{ij}^{(m)}] \text{Var}[X_{kl}^{(n)}] \right\}^{\frac{1}{2}} \left\{ \delta_{\pi(i,j)\pi(k,l)} \left( \gamma_{ij(\text{sub})}^{(m)} \gamma_{kl(\text{sub})}^{(n)} \right)^{\frac{1}{2}} \right. \\ &\quad \left. + \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \left( \gamma_{ij(\text{seg})}^{(m)} \gamma_{kl(\text{seg})}^{(n)} \right)^{\frac{1}{2}} + \delta_{ik} \delta_{jl} \delta_{mn} \gamma_{ij(\text{cell})}^{(m)} \right\}, \end{aligned} \tag{5.3}$$

where

$$\gamma_{ij(S)}^{(m)} = V_{ij(S)}^{(m)} / \text{Var}[X_{ij}^{(m)}], \quad S = \text{sub, seg, or cell.} \tag{5.4}$$

Hence,

$$\begin{aligned} \text{Corr} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \delta_{\pi(i,j)\pi(k,l)} \left( \gamma_{ij(\text{sub})}^{(m)} \gamma_{kl(\text{sub})}^{(n)} \right)^{\frac{1}{2}} \\ &\quad + \delta_{mn} \delta_{\pi(i,j)\pi(k,l)} \left( \gamma_{ij(\text{seg})}^{(m)} \gamma_{kl(\text{seg})}^{(n)} \right)^{\frac{1}{2}} + \delta_{ik} \delta_{jl} \delta_{mn} \left( 1 - \gamma_{ij(\text{sub})}^{(m)} - \gamma_{ij(\text{seg})}^{(m)} \right), \end{aligned} \tag{5.5}$$

where it has been recognised that  $\gamma_{ij(\text{sub})}^{(m)} + \gamma_{ij(\text{seg})}^{(m)} + \gamma_{ij(\text{cell})}^{(m)} = 1$ .

The inclusion of AR(1) dependencies between subsets leads to a simple modification of (5.5). For example, it was seen in Section 3.4.2, specifically (3.34), that an AR(1) dependency between diagonals across all segments led to the appearance of the factor  $\theta^{|(i+j)-(k+l)|}$ ; a similar dependency between diagonals in just segment  $m$  led to the appearance of the factor  $(\theta^{(m)})^{|(i+j)-(k+l)|}$ .

If these dependencies are included in the present case, then (5.5) is replaced by the following:

$$\begin{aligned} \text{Corr} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \left( \gamma_{ij(\text{sub})}^{(m)} \gamma_{kl(\text{sub})}^{(n)} \right)^{\frac{1}{2}} \theta^{|(i+j)-(k+l)|} \\ &\quad + \delta_{mn} \left( \gamma_{ij(\text{seg})}^{(m)} \gamma_{kl(\text{seg})}^{(n)} \right)^{\frac{1}{2}} (\theta^{(m)})^{|(i+j)-(k+l)|} \\ &\quad + \delta_{ik} \delta_{jl} \delta_{mn} \left( 1 - \gamma_{ij(\text{sub})}^{(m)} - \gamma_{ij(\text{seg})}^{(m)} \right). \end{aligned} \tag{5.6}$$

This result, though relatively simple, provides a rich correlation structure. It depends on only four families of parameters, whose interpretations are summarised in Table 1.

Other variations of (5.6), e.g., AR(1) structures on rows of the arrays are easily constructed.

TABLE 1  
INTERPRETATION OF QUANTITIES REQUIRED FOR CORRELATION CONSTRUCTION.

Quantity	Applicable To	Interpretation
$\gamma_{ij(\text{sub})}^{(m)}$	Contribution to variance of cell $(i, j)$ of business segment $m$ specific to the <i>subset</i> to which $(i, j)$ belongs	Proportion of total cell variance
$\gamma_{ij(\text{seg})}^{(m)}$	Contribution to variance of cell $(i, j)$ of business segment $m$ specific to that <i>segment</i>	Proportion of total cell variance
$\theta$	AR(1) structure on subsets in relation to the component of dependency that runs across all segments	Autoregression parameter
$\theta^{(m)}$	AR(1) structure on subsets in relation to the component of dependency specific to segment $m$	Autoregression parameter

### 5.3. Numerical illustration

The present section will provide a simple numerical example of formula (5.6) at work. Simplicity is necessary if display of the correlation matrix is to be feasible, since even the modest case  $I = J = 4$ ,  $N = 2$  (two segments, each represented by  $4 \times 4$  claim triangles) generates a  $20 \times 20$  correlation matrix; see also Table 3 below. Note that only upper triangles are dealt with in this illustration. In practical cases, unlimited by display on a page, much larger correlation matrices could, and should, be produced.

As noted in Section 5.2, the parameters requiring estimation are  $\theta; \theta^{(m)}, m = 1, \dots, N, \gamma_{ij(S)}^{(m)}, S = \text{sub, seg}; i = 1, \dots, I, j = 1, \dots, J, m = 1, \dots, N$ . In some cases, it may be considered reasonable to assume that  $\gamma_{ij(S)}^{(m)} = \gamma_{(S)}^{(m)}$ , independent of  $i, j$ . Then, (5.6) simplifies slightly:

$$\begin{aligned} \text{Corr} \left[ X_{ij}^{(m)}, X_{kl}^{(n)} \right] &= \left( \gamma_{(\text{sub})}^{(m)} \gamma_{(\text{sub})}^{(n)} \right)^{\frac{1}{2}} \theta^{|(i+j)-(k+l)|} \\ &+ \delta_{mn} \gamma_{(\text{seg})}^{(m)} \left( \theta^{(m)} \right)^{|(i+j)-(k+l)|} + \delta_{ik} \delta_{jl} \delta_{mn} \left( 1 - \gamma_{(\text{sub})}^{(m)} - \gamma_{(\text{seg})}^{(m)} \right). \end{aligned} \quad (5.7)$$

The number of parameters to be estimated is now  $3N + 1$ .

The example below will involve two  $4 \times 4$  claim triangles ( $I = J = 4, N = 2$ ), and so total of 7 parameters, all of which have a highly intuitive interpretation. The triangles will relate to Property and Liability classes, denoted  $m = 1, 2$ , respectively. It will be assumed that:

TABLE 2  
ASSUMED PARAMETER VALUES.

Parameter	Assumed Value
$\gamma_{(\text{sub})}^{(1)}$	0.1
$\gamma_{(\text{sub})}^{(2)}$	0.1
$\gamma_{(\text{seg})}^{(1)}$	0.3
$\gamma_{(\text{seg})}^{(2)}$	0.1
$\theta$	0.2
$\theta^{(1)}$	0.3
$\theta^{(2)}$	0.6

- simple process error (i.e., unrelated to dependencies) accounts for a lesser proportion of cell variance in the Property class than in the Liability (that is,  $1 - \gamma_{(\text{sub})}^{(1)} - \gamma_{(\text{seg})}^{(1)} < 1 - \gamma_{(\text{sub})}^{(2)} - \gamma_{(\text{seg})}^{(2)}$ );
- diagonals are subject to positive AR(1) dependency both within and between classes (that is,  $\theta^{(m)} > 0$  and  $\theta > 0$ );
- the within-Property dependency between diagonals is stronger than the within-Liability dependency (that is,  $\gamma_{(\text{seg})}^{(1)} > \gamma_{(\text{seg})}^{(2)}$ );
- the between-class dependency with respect to diagonals is weak (that is,  $\gamma_{(\text{sub})}^{(m)}$  are close to 0);
- the within-class AR(1) dependency decays (within increasing distance between diagonals) more rapidly in the Property class than the Liability (that is,  $\theta^{(1)} < \theta^{(2)}$ ).

The specific parameter values assumed for the example are set out in Table 2, and reflect the above properties.

Table 3 displays the correlation matrix obtained on application of (5.7) with parameter values as assumed in Table 2. For maximum clarity of the trends in the matrix, entries are ordered by AY within CY. Matrix entries that correlate observations from the same diagonal (CY), or diagonals that are identically located in the two triangles, are indicated by shading.

## 6. CONCLUSION

This paper is concerned with multiple claim arrays, with observations indexed by array number  $n$ , accident period  $i$  and development period  $j$ . It constructs a number of models that incorporate dependencies between observations both within arrays and between arrays. Arrays are of general shape (possibly with holes), but include the usual cases of claim triangles and trapezia that appear in the literature.



TABLE 3  
CORRELATION MATRIX ( $AY = i, CY = t$ ).

		Class # 1												2									
		$t = 1$	2	2	3	3	3	4	4	4	4	1	2	2	3	3	3	4	4	4	4		
Class #	$t =$	$i =$	$i = 1$	1	2	1	2	3	1	2	3	4	1	1	2	3	3	3	4	4	4	4	
1	1	1	1.00	0.11	0.11	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.10	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2	1	0.11	1.00	0.40	0.11	0.11	0.11	0.03	0.03	0.03	0.03	0.02	0.10	0.10	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.00
	2	2	0.11	0.40	1.00	0.11	0.11	0.11	0.03	0.03	0.03	0.03	0.02	0.10	0.10	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.00
	3	1	0.03	0.11	0.11	1.00	0.40	0.40	0.11	0.11	0.11	0.11	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.02
	3	2	0.03	0.11	0.11	0.40	1.00	0.40	0.11	0.11	0.11	0.11	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.02
	3	3	0.03	0.11	0.11	0.40	0.40	1.00	0.11	0.11	0.11	0.11	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.02
	4	1	0.01	0.03	0.03	0.11	0.11	0.11	1.00	0.40	0.40	0.40	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.10
	4	2	0.01	0.03	0.03	0.11	0.11	0.11	0.40	1.00	0.40	0.40	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.10
	4	3	0.01	0.03	0.03	0.11	0.11	0.11	0.40	0.40	1.00	0.40	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.10
	4	4	0.01	0.03	0.03	0.11	0.11	0.11	0.40	0.40	0.40	1.00	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.10
2	1	1	0.10	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.08	0.08	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.02
	2	1	0.02	0.10	0.10	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.08	1.00	0.20	0.08	0.08	0.08	0.04	0.04	0.04	0.04	0.04
	2	2	0.02	0.10	0.10	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.08	0.20	1.00	0.08	0.08	0.08	0.04	0.04	0.04	0.04	0.04
	3	1	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.04	0.08	0.08	1.00	0.20	0.20	0.08	0.08	0.08	0.08	0.08
	3	2	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.04	0.08	0.08	0.20	1.00	0.20	0.08	0.08	0.08	0.08	0.08
	3	3	0.00	0.02	0.02	0.10	0.10	0.10	0.02	0.02	0.02	0.02	0.04	0.08	0.08	0.20	0.20	1.00	0.08	0.08	0.08	0.08	0.08
	4	1	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.02	0.04	0.04	0.08	0.08	0.08	1.00	0.20	0.20	0.20	0.20
	4	2	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.02	0.04	0.04	0.08	0.08	0.08	0.20	1.00	0.20	0.20	0.20
	4	3	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.02	0.04	0.04	0.08	0.08	0.08	0.20	0.20	1.00	0.20	0.20
	4	4	0.00	0.00	0.00	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.02	0.04	0.04	0.08	0.08	0.08	0.20	0.20	0.20	1.00	0.20

The dependencies are defined in reasonable generality. At the most general level, each claim array is partitioned into a collection of subsets, and dependencies between observations are defined in terms of the subsets to which the observations belong.

The subsets are quite general, and may be chosen arbitrarily. Specific examples are cells, rows, columns and diagonals. These choices yield cell-wise, row-wise, column-wise and diagonal-wise forms of dependency. Combinations of these forms of dependency are also permitted.

At the most basic level of (say) diagonal-wise dependency, observations are dependent if they belong to the same diagonal within an array, or the same diagonals in different arrays, but are independent otherwise. Further, structure is introduced to incorporate dependency between different diagonals. This would usually be structured so that dependency diminishes as the distance between the diagonals increases.

The dependency structures are constructed, for the most part, by means of common shocks. The explicit forms of covariance are given in the case of each model.

As remarked in Section 1, a model of dependency between such a large number of observations is likely to involve a large number of parameters, estimates of all of which will be required if the model is to be operationalised. Section 5.1 makes a very brief discussion of parameter estimation in general but, as also pointed out in Section 1, this is not the major emphasis of the paper.

Instead, some attention is paid to “heuristic estimation” in Section 5.2. This is the kind of guesswork by practitioners that one encounters, usually involving a large correlation matrix with respect to all observations. These matrices are often of dubious provenance, with construction lacking rigour, and potentially even lacking positive definiteness. This is not to express disapproval of the practitioner’s approach, but rather to seek more organised correlation matrix construction, while accepting that parameter values will be obtained heuristically. Section 5.2 considers a reasonably generic model, and obtains an expression for the general entry in the correlation matrix in terms of a limited set of parameters, each of which will have a straightforward intuitive meaning to the practitioner. This will maximise the likelihood of a reliable matrix.

This approach to estimation is illustrated by a numerical example in Section 5.3. In this example, it is shown that the correlation matrix over  $N$  claims triangles, each of dimension  $J$ , can be constructed on the basis of  $3N + 1$  parameters even though the matrix contains  $NJ(J + 1)[NJ(J + 1) - 2]/8$  free entries. In the specific numerical example,  $N = 2$ ,  $J = 4$  so that 190 free correlations are specified by 7 parameters, and all 7 with strong intuitive meaning.

Finally, the correlation matrix is combined with estimates of tail dependency to arrive at a multivariate  $t$  model for the construction of capital margins dealing with the extreme right tail of liabilities.

## ACKNOWLEDGEMENTS

This paper was presented at an Australian Actuaries Institute *Insights event* in October 2016 in Sydney, Australia. The authors are grateful for constructive comments received from colleagues who attended this event, as well as for remarks from three reviewers, which all led to substantial improvements to the paper. Finally, the authors are also grateful to Gwendal Combot, Phuong Anh Vu and Alan Xian for their research assistance.

This research was supported by a grant from the Australian Actuaries Institute, as well as under Australian Research Council's Linkage Projects funding scheme (Project no. LP130100723, with funding partners Allianz Australia Insurance Ltd., Insurance Australia Group Ltd. and Suncorp Metway Ltd.). The views expressed herein are those of the authors and are not necessarily those of the supporting organisations.

## REFERENCES

- ABDALLAH, A., BOUCHER, J.-P. and COSSETTE, H. (2015) Modeling dependence between loss triangles with hierarchical Archimedean copulas. *ASTIN Bulletin*, **45**(3), 577–599.
- ABDALLAH, A., BOUCHER, J.-P., COSSETTE, H. and TRUFIN, J. (2016) Sarmanov family of bivariate distributions for multivariate loss reserving analysis. *North American Actuarial Journal*, **20**(2), 184–200.
- AVANZI, B., TAYLOR, G.C., VU, P.A. and WONG, B. (2016a) Stochastic loss reserving with dependence: A flexible multivariate Tweedie approach. *Insurance: Mathematics and Economics*, **71**, 63–78.
- AVANZI, B., TAYLOR, G.C. and WONG, B. (2016b) Correlations between insurance lines of business: An illusion or a real phenomenon? Some methodological considerations. *ASTIN Bulletin*, **46**(2), 225–263.
- BRAUN, C. (2004) The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin*, **34**, 399–423.
- 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.
- HAPP, S., MAIER, R. and MERZ, M. (2014) Multivariate Bühlmann–Straub credibility model applied to claims reserving for correlated run-off triangles. *Variance*, **8**, 23–42.
- HESS, K.T., SCHMIDT, K.D. and ZOCHER, M. (2006) Multivariate loss prediction in the multivariate additive model. *Insurance: Mathematics and Economics*, **39**(2), 185–191.
- IAA (2009) *Note on Enterprise Risk Management for Capital and Solvency Purposes in the Insurance Industry*. Sydney, Australia: International Actuarial Association.
- JØRGENSEN, B. (1997) The theory of dispersion models. In *Monographs on Statistics and Applied Probability*, vol. 76. London: Chapman & Hall.
- KUANG, D., NIELSEN, B. and NIELSEN, J. (2008) Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 987–991.
- LUENBERGER, D. (2014) *Investment Science*. Oxford, UK: Oxford University Press.
- MERZ, M., WÜTHRICH, M.V. and HASHORVA, E. (2013) Dependence modelling in multivariate claims run-off triangles. *Annals of Actuarial Science*, **7**, 3–25.
- MEYERS, G.G. (2007) The common shock model for correlated insurance losses. *Variance*, **1**(1), 40–52.
- MEYERS, G.G. (2016, Winter) Dependencies in stochastic loss reserve models. Casualty Actuarial Society Forum.

- RENSHAW, A. (1989) Chain ladder and interactive modelling (claims reserving and GLIM). *Journal of the Institute of Actuaries*, **116**, 559–587.
- SHARPE, W.F. (1963) A simplified model for portfolio analysis. *Management Science*, **9**(2), 277–293.
- 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. and MEYERS, G.G. (2012) A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal*, **16**(1), 29–51.
- SHI, P. and FREES, E.W. (2011) Dependent loss reserving using copulas. *ASTIN Bulletin*, **41**(2), 449–486.
- TAYLOR, G. (2000) Loss reserving: An actuarial perspective. In *Huebner International Series on Risk, Insurance and Economic Security*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- TAYLOR, G. (2017) Existence and uniqueness of chain ladder solutions. *ASTIN Bulletin*, **47**(1), 1–41.
- TAYLOR, G. and MCGUIRE, G. (2007) A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. *North American Actuarial Journal*, **11**(3), 70–88.
- TWEEDIE, M. (1984) An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions*, pp. 579–604.
- WÜTHRICH, M. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Chichester, UK: John Wiley & Sons.
- WÜTHRICH, M.V. and MERZ, M. (2015) *Stochastic claims reserving manual: Advances in dynamic modeling*. SSRN Manuscript no. 2649057.
- ZHANG, Y. (2010) A general multivariate chain ladder model. *Insurance: Mathematics and Economics*, **46**(3), 588–599.
- ZHANG, Y. and DUKIC, V. (2013) Predicting multivariate insurance loss payments under the Bayesian copula framework. *Journal of Risk and Insurance*, **80**(4), 891–919.

BENJAMIN AVANZI

*School of Risk and Actuarial Studies*

*UNSW Sydney Business School*

*UNSW Sydney, NSW 2052, Australia*

*Département de Mathématiques et de Statistique*

*Université de Montréal*

*Montréal, Quebec H3T 1J4, Canada*

*E-Mail: b.avanzi@unsw.edu.au*

GREG TAYLOR (Corresponding author)

*School of Risk and Actuarial Studies*

*UNSW Sydney Business School*

*UNSW Sydney, NSW 2052, Australia*

*E-Mail: greg.taylor@unsw.edu.au*

BERNARD WONG

*School of Risk and Actuarial Studies*

*UNSW Sydney Business School*

*UNSW Sydney, NSW 2052, Australia*

*E-Mail: bernard.wong@unsw.edu.au*