CHAIN LADDER CORRELATIONS

Greg Taylor

Taylor Fry Consulting Actuaries
Level 11, 55 Clarence Street
Sydney NSW 2000
Australia

Professorial Associate
Centre for Actuarial Studies
Faculty of Economics and Commerce
University of Melbourne
Parkville VIC 3052
Australia

Phone: 61 2 9249 2901
Fax: 61 2 9249 2999
greg.taylor@taylorfry.com.au

August 2011

Abstract
Correlations of future observations are investigated within the recursive and non-recursive chain ladder models. The recursive models considered are the Mack and ODP Mack models; the non-recursive models are the ODP cross-classified models. Distinct similarities are found between the correlations within the recursive and non-recursive models, but distinct differences also emerge. The ordering of corresponding correlations within the recursive and non-recursive models is also investigated.

Keywords: chain ladder, correlation, Mack model, non-recursive model, ODP cross-classified model, ODP Mack model, recursive model,

1. Introduction

The actuarial literature identifies two families of chain ladder models categorised by Verrall (2000) as **recursive** and **non-recursive** models respectively. Although the model formulations are fundamentally different, both are found to yield the same maximum likelihood estimators of age-to-age factors and the same forecasts of loss reserve. The properties of these models are studied by Taylor (2011).

Despite the identical forecasts of the different models, their different formulations are liable to lead to different correlation structures. This means that the correlations can be regarded as providing one means of differentiating between recursive and non-recursive models. The purpose of the present paper is the investigation of these correlation structures.

There is independence between rows in all the models considered, so the correlations of greatest interest are those between future observations conditional on information up to a defined point of time, specifically $Corr[X_{k,j+m}, X_{k,j+m+n} \mid X_{kj}]$ where X_{kj} denotes cumulative claims experience (notifications, payments, etc) up to an including development period j in respect of accident period k.

2. Framework and notation

2.1 Claims data

Consider a K x J rectangle of claims observations Y_{kj} with:

- accident periods represented by rows and labelled k = 1, 2, .., K;
- development periods represented by columns and labelled by $j = 1, 2, \ldots, J \le K$.

Within the rectangle identify a **development trapezoid** of **past** observations

$$\mathcal{D}_K = \left\{ Y_{kj} : 1 \le k \le K \text{ and } 1 \le j \le \min\left(J, K-k+1\right) \right\}$$

The complement of this subset, representing **future** observations is

$$\mathcal{D}_K^c = \left\{ Y_{kj} : 1 \le k \le K \text{ and } \min\left(J, K-k+1\right) < j \le J \right\}$$
$$= \left\{ Y_{kj} : K-J+1 < k \le K \text{ and } K-k+1 < j \le J \right\}$$

Also let

$$\mathcal{D}_K^+ = \mathcal{D}_K \cup \mathcal{D}_K^c$$

In general, the problem is to predict $\mathcal{D}_K^c$ on the basis of observed $\mathcal{D}_K$.

The usual case in the literature (though often not in practice) is that in which $J = K$, so that the trapezoid becomes a triangle. The more general trapezoid will be retained throughout the present paper.

Define the **cumulative row sums**

$$X_{kj} = \sum_{i=1}^{j} Y_{ki} \tag{2.1}$$

and the full **row and column sums** (or horizontal and vertical sums)

$$H_k = \sum_{j=1}^{\min(J,K-k+1)} Y_{kj}$$

$$V_j = \sum_{k=1}^{K-j+1} Y_{kj} \tag{2.2}$$

Also define, for $k = K - J + 2, \ldots, K$,

$$R_k = \sum_{j=K-k+2}^{J} Y_{kj} = X_{kJ} - X_{k,K-k+1} \tag{2.3}$$

$$R = \sum_{k=K-J+2}^{K} R_k \tag{2.4}$$

Note that R is the sum of the (future) observations in $\mathcal{D}_K^c$. It will be referred to as the total amount of **outstanding losses**. Likewise, R_k denotes the amount of outstanding losses in respect of accident period k. The objective stated earlier is to forecast the R_k and R.

Let $\sum^{\mathcal{R}(k)}$ denote summation over the entire row k of $\mathcal{D}_K$, i.e. $\sum_{j=1}^{\min(J,K-k+1)}$ for fixed k.

Similarly, let $\sum^{C(j)}$ denote summation over the entire column of $\mathcal{D}_K$, i.e. $\sum_{k=1}^{K-j+1}$ for fixed j. For example, (2.2) may be expressed as

$$V_j = \sum^{C(j)} Y_{kj}$$

Finally, let $\sum^{\mathcal{T}}$ denote summation over the entire trapezoid of (*k*,*j*) cells, i.e.

$$\sum^{\mathcal{T}} = \sum_{k=1}^{K} \sum_{j=1}^{\min(J,K-k+1)} = \sum_{k=1}^{K} \sum^{\mathcal{R}(k)}$$
$$= \sum_{j=1}^{J} \sum_{k=1}^{K-j+1} = \sum_{j=1}^{J} \sum^{C(j)}$$

For a random variable A_{kj} with $(k,j) \in \mathcal{D}_K, A_K$ will denote the entire array $\{A_{kj} : (k,j) \in \mathcal{D}_K\}$. The first column $(j=1)$ of A_K will be denoted by A_K^1.

2.2 Families of distributions

2.2.1 Exponential dispersion family

The **exponential dispersion family** (EDF) (Nelder & Wedderburn, 1972) consists of those variables *Y* with log-likelihoods of the form

$$\ell(y,\theta,\phi) = [y\theta - b(\theta)] / a(\phi) + c(y,\phi) \tag{2.5}$$

for parameters θ (canonical parameter) and ϕ (scale parameter) and suitable functions *a*, *b* and *c*, with *a* continuous, *b* differentiable and one-one, and *c* such as to produce a total probability mass of unity.

For *Y* so distributed,

$$E[Y] = b'(\theta) \tag{2.6}$$

$$Var[Y] = a(\phi) b''(\theta) \tag{2.7}$$

If μ denotes *E*[*Y*], then (2.6) establishes a relation between μ and θ, and so (2.7) may be expressed in the form

$$Var[Y] = a(\phi) V(\mu) \tag{2.8}$$

for some function *V*, referred to as the **variance function**.

The notation $Y \sim EDF(\theta,\phi;a,b,c)$ will be used to mean that a random variable *Y* is subject to the EDF likelihood (2.5).

2.2.2 Tweedie family

The **Tweedie family** (Tweedie, 1984) is the sub-family of the EDF for which

$$a(\phi) = \phi \tag{2.9}$$

$$V(\mu) = \mu^p, \; p \le 0 \text{ or } p \ge 1 \tag{2.10}$$

For this family,

$$b(\theta) = (2-p)^{-1}\left[(1-p)\theta\right]^{(2-p)/(1-p)} \quad (2.11)$$

$$\mu = [(1 - p)\theta]^{1/(1 - p)} \quad (2.12)$$

$$\ell(y;\mu,\phi) = \left[y\mu^{1-p}/(1-p) - \mu^{2-p}/(2-p)\right]/\phi + c(y,\phi) \quad (2.13)$$

$$\partial\ell/\partial\mu = \left(y\mu^{-p} - \mu^{1-p}\right)/\phi \quad (2.14)$$

The notation $Y \sim Tw(\mu,\phi,p)$ will be used to mean that a random variable Y is subject to the Tweedie likelihood with parameters μ,ϕ,p. The abbreviated form $Y \sim Tw(p)$ will mean that Y is a member of the sub-family with specific parameter p.

2.2.3 Over-dispersed Poisson family

The **over-dispersed Poisson** (ODP) family is the Tweedie sub-family with p = 1. The limit of (2.12) as $p \to 1$ gives

$$E[Y] = \mu = \exp\theta \quad (2.15)$$

By (2.8) – (2.10),

$$Var[Y] = \phi\mu \quad (2.16)$$

By (2.14),

$$\partial\ell/\partial\mu = (y-\mu)/\phi\mu \quad (2.17)$$

The notation $Y \sim ODP(\mu,\phi)$ means $Y \sim Tw(\mu,\phi,1)$.

3. Chain ladder models

3.1 Heuristic chain ladder

The chain ladder was originally (pre-1975) devised as a heuristic algorithm for forecasting outstanding losses. It had no statistical foundation. The algorithm is as follows.

Define the following factors:

$$\hat{f}_j = \sum_{k=1}^{K-j} X_{k,j+1} / \sum_{k=1}^{K-j} X_{kj},\ j = 1,2,...,J-1 \quad (3.1)$$

Note that $\hat{f}_j$ can be expressed in the form

$$\hat{f}_j = \sum_{k=1}^{K-j} w_{kj} \left(X_{k,j+1} / X_{kj} \right) \tag{3.2}$$

with

$$w_{kj} = X_{kj} / \sum_{k=1}^{K-j} X_{kj} \tag{3.3}$$

i.e. as a weighted average of factors $X_{k,j+1} / X_{kj}$ for fixed j.

Then define the following forecasts of $Y_{kj} \in \mathcal{D}_K^c$:

$$\hat{Y}_{kj} = X_{k,K-k+1} \; \hat{f}_{K-k+1} \; \hat{f}_{K-k+2} \ldots \hat{f}_{j-2} \left(\hat{f}_{j-1} - 1 \right) \tag{3.4}$$

Call these **chain ladder forecasts**. They yield the additional chain ladder forecasts:

$$\hat{X}_{kj} = X_{k,K-k+1} \; \hat{f}_{K-k+1} \ldots \hat{f}_{j-1} \tag{3.5}$$

$$\hat{R}_k = \hat{X}_{kJ} - \hat{X}_{k,K-k+1} \tag{3.6}$$

$$\hat{R} = \sum_{k=K-J+2}^{K} \hat{R}_k \tag{3.7}$$

3.2 Recursive models

A recursive model takes the general form

$$E\left[X_{k,j+1} \mid X_{kj} \right] = \text{function of } \mathcal{D}_{k+j-1} \text{ and some parameters} \tag{3.8}$$

where $\mathcal{D}_{k+j-1}$ is the data sub-array of $\mathcal{D}_K$ obtained by deleting diagonals on the right side of $\mathcal{D}_K$ until X_{kj} is contained in its right-most diagonal.

3.2.1 Mack model

The Mack model (Mack, 1993) is defined by the following assumptions.

(M1) Accident periods are stochastically independent, i.e. $Y_{k_1 j_1}, Y_{k_2 j_2}$ are stochastically independent if $k_1 \neq k_2$.

(M2) For each $k = 1, 2, \ldots, K$, the X_{kj} (j varying) form a Markov chain.

(M3) For each $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, J - 1$,
(a) $E\left[X_{k,j+1} \mid X_{kj} \right] = f_j X_{kj}$ for some parameters $f_j > 0$; and

(b) $Var\left[X_{k,j+1}\middle|X_{kj}\right]=\sigma_j^2 X_{kj}$ for some parameters $\sigma_j^2>0$.

3.2.2 ODP Mack model

Taylor (2011) defined the over–dispersed Poisson (ODP) Mack model as that satisfying assumptions (M1), (M2) and

$(ODPM3)$ For each k=1,2,…,K and j=1,2,…,J-1,

$$Y_{k,j+1} \mid X_{kj} \sim ODP\left(\left(f_j-1\right)X_{kj},\phi_{k,j+1}\right)$$

where now $f_j \geq 1$.

Assumption (ODPM3) implies (M3a). Moreover, in the special case $\phi_{k,j+1}=\phi_{j+1}$ independent of k, (ODPM3) also implies (M3b) with $\sigma_j^2=\phi_{j+1}\left(f_j-1\right)$.

It is evident that, for this model to be valid, it is necessary that all $Y_{k,j} \geq 0$. Note also that, under $(ODPM3)$, $X_{kj}=0$ implies that $X_{k,j+m}=0$ for all $m>0$. This means that, for each k, either $Y_{k1}>0$ or or $X_{kj}=0$ for all j.

A summary of these requirements in terms of the data array $\mathcal{D}_K$ is as follows.

(R1) $Y_{kj} \geq 0$ for all $Y_{kj} \in \mathcal{D}_K$

(R2) For each k = 1, 2, …, K, either:

(a) $Y_{k1}>0$; or

(b) $Y_{kj}=0$ for all $1 \leq j \leq \min\left(J,K-k+1\right)$

A data array satisfying these requirements will be called **ODPM-regular**.

Assumption (ODPM3) may be expressed in the following form, suitable for GLM implementation of the OPD Mack model:

$$Y_{k,j+1} \mid X_{kj} \sim ODP\left(\exp\left[\ln X_{kj}+\ln\left(f_j-1\right)\right],\phi/w_{k,j+1}\right) \tag{3.9}$$

where

$$w_{k,j+1}=\phi/\phi_{k,j+1} \tag{3.10}$$

In this form, the GLM of the $Y_{k,j+1}$ has log link, offsets $\ln X_{kj}$, parameters $\ln\left(f_j-1\right)$, and weights $w_{k,j+1}$.

It is shown by Taylor (2011) that the chain ladder estimates of age-to-age factors (3.1) are maximum likelihood for this model.

3.3 Non-recursive models

Taylor (2011) also defined the **ODP cross-classified model** as that satisfying the following assumptions:

(ODPCC1) The random variables $Y_{kj} \in \mathcal{D}_K^+$ are stochastically independent.

(ODPCC2) For each $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, J$,

(a) $Y_{kj} \sim ODP\left(\mu_{kj}, \phi_{kj}\right)$;

(b) $\mu_{kj} = \alpha_k \beta_j$ for some parameters $a_k, \beta_j > 0$; and

(c) $\sum_{j=1}^{J} \beta_j = 1$

Assumption (ODPCC2b) may be expressed in the following form, suitable for GLM implementation of the ODP cross-classified model:

$$Y_{kj} \sim ODP\left(\exp\left(\ln \alpha_k + \ln \beta_j\right), \phi / w_{kj}\right) \tag{3.11}$$

In this form, the GLM of the Y_{kj} has log link, parameters $\ln \alpha_k$ and $\ln \beta_j$, and weights w_{kj} satisfying

$$w_{kj} = \phi / \phi_{kj} \tag{3.12}$$

Assumption (ODPCC2b) removes one degree of redundancy from the parameter set, and would be reflected by the aliasing of one parameter in the GLM.

It has long been known for the case $\phi / w_{kj} = 1$ that the maximum likelihood forecasts of future Y_{kj} in this model are the same as the chain ladder forecasts (3.5)-(3.7) (see e.g. Hachemeister & Stanard, 1975; Renshaw & Verrall, 1998; Taylor, 2000). It is shown by England & Verrall (2002) that this result continues to hold in the more general case $\phi / w_{kj} = \phi \neq 1$.

Thus the ODP Mack and ODP cross-classified models produce the same maximum likelihood forecasts of loss reserves despite their fundamentally different formulations. This means that their respective correlation structures can be viewed as a means of differentiating between them.

4. Correlation between observations

4.1 Background common to recursive and non-recursive models

Consider the models defined in Sections 3.2 and 3.3, and specifically the conditional covariance $Cov\left[X_{k_1,j_1+m}, X_{k_2,j_2+m+n} \middle| X_{k_1,j_1}, X_{k_2 j_2}\right]$ with $m>0, n\geq 0$. The following lemma is immediate from assumption (M1) or (ODPCC1).

Lemma 4.1. The following is true for each of the Mack, ODP Mack and ODP cross-classified models:

$Cov\left[X_{k_1,j_1+m}, X_{k_2,j_2+m+n} \middle| X_{k_1,j_1}, X_{k_2 j_2}\right] = 0$ for $k_1 \neq k_2$ □

In view of this result, attention will be focused on **within-row covariances** $Cov\left[X_{k,j+m}, X_{k,j+m+n} \middle| X_{kj}\right]$**.** This quantity will be denoted $c_{k,j+m,j+m+n|j}$. It is evaluated as follows:

$$c_{k,j+m,j+m+n|j} = E\left[\left\{X_{k,j+m}, X_{k,j+m+n} \middle| X_{kj}\right]\right\} \times \left\{X_{k,j+m+n} - \mathrm{E}\left[X_{k,j+m+n} \middle| X_{kj}\right]\right\} \middle| X_{kj}\right]$$
$$= E\left[\left\{X_{k,j+m} - E\left[X_{k,j+m} \middle| X_{kj}\right]\right\} \times E\left[X_{k,j+m+n} - E\left[X_{k,j+m+n} \middle| X_{kj}\right] \middle| X_{k,j+m}\right] \middle| X_{kj}\right]$$
$$= E\left[\left\{X_{k,j+m} - E\left[X_{k,j+m} \middle| X_{kj}\right]\right\} \times \left\{E\left[X_{k,j+m+n} \mid X_{k,j+m}\right] - E\left[X_{k,j+m+n} \middle| X_{kj}\right]\right\} \middle| X_{kj}\right] \quad (4.1)$$

4.2 Recursive models

4.2.1 Mack model

By recursive application of (M3a),
$E\left[X_{k,j+m+n} \middle| X_{k,j+m}\right] = f_{j+m+n-1} f_{j+m+n-2} \cdots f_{j+m} X_{j+m}$
and so

$$E\left[X_{k,j+m+n} \middle| X_{k,j+m}\right] - E\left[X_{k,j+m+n} \middle| X_{kj}\right]$$
$$= f_{j+m+n-1} \cdots f_{j+m} \left\{X_{k,j+m} - E\left[X_{k,j+m} \middle| X_{kj}\right]\right\} \quad (4.2)$$

Substitution of (4.2) into (4.1) yields

$$c_{k,j+m,j+m+n|j} = f_{j+m+n-1} \cdots f_{j+m} Var\left[X_{k,j+m} \middle| X_{kj}\right] \quad (4.3)$$

The variance term here is evaluated by Mack (1993, p.218) as

$$Var\left[X_{k,j+m} \middle| X_{kj}\right] = X_{kj} \sum_{i=j}^{j+m-1} f_{j+m-1}^2 \cdots f_{i+1}^2 \sigma_i^2 f_{i-1} \cdots f_j \quad (4.4)$$

Substitution of (4.4) into (4.3) yields

$$c_{k,j+m,j+m+n|j} = f_{j+m+n-1 \ldots} f_{j+m} X_{kj} \sum_{i=j}^{j+m-1} f_{j+m-1}^2 \cdots f_{i+1}^2 \sigma_i^2 f_{i-1} \cdots f_j \quad (4.5)$$

It then follows that

$$Corr\left[X_{k,j+m,}\, X_{k,j+m+n}\middle|X_{kj}\right] = \frac{c_{k,j+m,j+m+n|j}}{\left[c_{k,j+m+n,j+m+n|j}c_{k,j+m,j+m|j}\right]^{\frac{1}{2}}}$$

$$= \left[1 + B_{j+m,j+m+n|j}\right]^{-\frac{1}{2}} \tag{4.6}$$

where

$$B_{j+m,j+m+n|j} = \frac{\sum_{i=j+m}^{j+m+n-1} f_{j+m+n-1}^2 \cdots f_{i+1}^2 \sigma_i^2 f_{i-1} \cdots f_j}{\sum_{i=j}^{j+m-1} f_{j+m+n-1}^2 \cdots f_{i+1}^2 \sigma_i^2 \; f_{i-1} \cdots f_j} \tag{4.7}$$

An equivalent form is

$$B_{j+m,j+m+n|j} = \left(f_{j+m+n-1}^2 \cdots f_{j+m}^2\right)^{-1} \frac{\sum_{i=j+m}^{j+m+n-1} f_{j+m+n-1}^2 \cdots f_{i+1}^2 \sigma_i^2 f_{i-1} \cdots f_j}{\sum_{i=j}^{j+m-1} f_{j+m-1}^2 \cdots f_{i+1}^2 \sigma_i^2 \; f_{i-1} \cdots f_j} \tag{4.8}$$

Theorem 4.2. Consider an ODPM-regular data array subject to a Mack model, and consider a row k that is not identically zero. Let *j,m,n* be strictly positive integers and let $\rho_{k,j+m,j+m+n|j}$ denote $Corr\left[X_{k,j+m,}\, X_{k,j+m+n}\middle|X_{kj}\right]$**.** For a given schedule of values $\left\{f_i, \sigma_i^2\right\}$ each of the following propositions holds:

(a) $0 < \rho_{k,j+m,j+m+n|j} < 1$

(b) $\rho_{k,j+m,j+m+n+1|j} < \rho_{k,j+m,j+m+n|j}$

(c) $\rho_{k,j+m,j+m+n|j}$ increases as any $\sigma_i^2,\, j \le i \le j+m-1$ increases, or any $\sigma_i^2,\, j+m \le i \le j+m+n-1$ decreases.

(d) $\rho_{k,j+m,j+m+n|j}$ increases as any f_i, $j+1 \le i \le j+m+n-1$ increases and σ_i^2 changes such that:
σ_i^2 / f_i increases if $j \le i \le j+m-1$; or
σ_i^2 / f_i decreases if $j+m \le i \le j+m+n-1$.

Proof. (a) Follows from (4.6) and the fact that $B_{j+m,j+m+n|j} > 0$.

(b) By (4.7), write

$$B_{j+m,j+m+n+1|j} = \frac{\sigma_{j+m+n}^2 f_{j+m+n-1} \cdots f_j}{\sum_{i=j}^{j+m-1} f_{j+m+n}^2 \cdots f_{i+1}^2 \sigma_i^2 \; f_{i-1} \cdots f_j} + B_{j+m,j+m+n|j}$$

$$> B_{j+m,j+m+n|j}$$

The result then follows from (4.6).

(c) Obvious from (4.8).

(d) Divide numerator and denominator of (4.7) by $f_{j+m+n-1}^2 \cdots f_{j+m}^2 f_{j+m-1} \ldots f_j$ to obtain

$$B_{j+m,j+m+n|j} = \frac{\sum_{i=j+m}^{j+m+n-1} (\sigma_i^2 / f_i) f_i^{-1} \cdots f_{j+m}^{-1}}{\sum_{i=j}^{j+m-1} f_{j+m-1} \cdots f_{i+1} (\sigma_i^2 / f_i)}$$

and the result then follows from (4.6). □

4.2.2 OPD Mack model

Expression (4.7) may be adapted to the case of the ODP Mack model with column dependent scale parameter $\phi_{kj} = \phi_j$. Section 3.2.2 notes that, in this case,

$$\sigma_j^2 = \phi_{j+1}\left(f_j - 1\right) \tag{4.9}$$

and substitution of this result in (4.7) yields

$$B_{j+m,j+m+n|j} = \frac{\sum_{i=j+m}^{j+m+n-1} \phi_{i+1} f_{j+m+n-1}^2 \cdots f_{i+1}^2 \left(f_i - 1\right) f_{i-1} \cdots f_j}{\sum_{i=j}^{j+m-1} \phi_{i+1} f_{j+m+n-1}^2 \cdots f_{i+1}^2 \left(f_i - 1\right) f_{i-1} \cdots f_j} \tag{4.10}$$

Special case. An interesting case arises when $f_i = f$, $\phi_{i+1} = \phi$, $i = j, j+1, \ldots, j+m+n-1$. Then (4.10) becomes

$$B_{j+m,j+m+n|j} = f^{-n} \left(f^n - 1\right) \big/ \left(f^m - 1\right) \tag{4.11}$$

4.3 Non-recursive models

Once again consider $\rho_{k,j+m,j+m+n|j}$. Note that

$$X_{k,j+m+n} = X_{k,j+m} + \sum_{i=j+m+1}^{j+m+n} Y_{ki}$$

where all terms on the right side are mutually stochastically independent.

Therefore

$$c_{k,j+m,j+m+n|j} = Var\left[X_{k,j+m} \middle| X_{kj}\right] \tag{4.12}$$

$$= Var\left[X_{kj} + \sum_{i=j+1}^{j+m} Y_{ki} \middle| X_{kj}\right]$$

$$= \sum_{i=j+1}^{j+m} Var\left[Y_{ki}\right] \tag{4.13}$$

by (ODPCC1).

By (4.12),

$$\rho^2_{j+m,j+m+n|j} = Var\left[X_{k,j+m} \middle| X_{kj}\right] \Big/ Var\left[X_{k,j+m+n} \middle| X_{kj}\right]$$
$$= \sum_{i=j}^{j+m-1} \phi_{i+1}\beta_{i+1} \Big/ \sum_{i=j}^{j+m+n-1} \phi_{i+1}\beta_{i+1} \qquad (4.14)$$

by (4.13) and (ODPCC2a-b).

Thus

$$\rho_{j+m,j+m+n|j} = (1 + D_{j+m,j+m+n|j})^{-\frac{1}{2}} \qquad (4.15)$$

with

$$D_{j+m,j+m+n|j} = \sum_{i=j+m}^{j+m+n-1} \phi_{i+1}\beta_{i+1} \Big/ \sum_{i=j}^{j+m-1} \phi_{i+1}\beta_{i+1} \qquad (4.16)$$

Now recall (3.18) and (3.19) and note that they yield

$$\beta_{i+1} = \frac{f_1 \cdots f_{i-1}(f_i - 1)}{\sum_{r=1}^{J-1} f_1 \cdots f_{r-1}(f_r - 1)} \qquad (4.17)$$

and this, combined with (4.16), gives

$$D_{j+m,j+m+n|j} = \frac{\sum_{i=j+m}^{j+m+n-1} \phi_{i+1}\, f_1 \cdots f_{i-1}(f_i - 1)}{\sum_{i=j}^{j+m-1} \phi_{i+1} f_1 \cdots f_{i-1}(f_i - 1)} \qquad (4.18)$$

$$= \frac{\sum_{i=j+m}^{j+m+n-1} f_{j+m} \cdots f_{i-1}(f_i - 1)\phi_{i+1}}{\sum_{i=j}^{j+m-1} [(1 - f_i^{-1})\phi_{i+1}]\, f_{i+1}^{-1} \cdots f_{j+m-1}^{-1}} \qquad (4.19)$$

Theorem 4.3. Consider an ODPM-regular data array subject to an OPD cross-classified model, and consider a row k that is not identically zero. Let j,m,n be strictly positive integers and let $\rho_{k,j+m,j+m+n|j}$ denote $Corr\left[X_{k,j+m}, X_{k,j+m+n} \middle| X_{kj}\right]$**.** For a given schedule of values $\{\beta_i, \phi_i\}$ each of the following propositions holds:

(a) $0 < \rho_{k,j+m,j+m+n|j} < 1$

(b) $\rho_{k,j+m,j+m+n+1|j} < \rho_{k,j+m,j+m+n|j}$

(c) $\rho_{k,j+m,j+m+n|j}$ increases as any ϕ_i or β_i, $j+1 \le i \le j+m$ increases, or any ϕ_i or β_i, $j+m+1 \le i \le j+m+n$ decreases.

(d) $\rho_{k,j+m,j+m+n|j}$ increases as any f_i, $j+1 \le i \le j+m+n-1$ decreases and ϕ_{i+1} changes such that $(1 - f_i^{-1})\phi_{i+1}$ increases.

Proof. (a) Follows directly from (4.14).

(b)-(c) Follow directly from (4.15) and (4.16).

(d) Follows directly from (4.15) and (4.19). □

It is interesting to compare the results of Theorems 4.2(d) and 4.3(d). The former shows that, subject to the condition on the dispersion parameter, an increase in an f_i causes $\rho_{k,j+m,j+m+n|j}$ to increase in the Mack model, whereas the latter yields the opposite result in the ODP cross-classified model.

Special case. An interesting special case arises when $\phi_i = \phi$, independent of i. Then (4.14) reduces

$$\rho^2_{k,j+m,j+m+n|j} = \sum_{i=j}^{j+m-1} \beta_{i+1} \Big/ \sum_{i=j}^{j+m+n-1} \beta_{i+1} \tag{4.20}$$

Special case. As in Section 4.2.2, the case $f_i = f$, $\phi_{i+1} = \phi$, $i = j, j+1, ..., j+m+n-1$ is interesting. Here, (4.18) yields

$$D_{j+m,j+m+n|j} = f^m \left(f^n - 1\right) \Big/ \left(f^m - 1\right) \tag{4.21}$$

Adopt the notation

$$\mu_{k,j+1,j+m} = E\left[\sum_{i=j+1}^{j+m} Y_{ki} \right] = \alpha_k \sum_{i=j}^{j+m-1} \beta_{i+1} \tag{4.22}$$

by (ODPCC2b).

4.4 Comparison between recursive and non-recursive models

Let $\rho^R_{k,j+m,j+m+n|j}$ denote $\rho_{k,j+m,j+m+n|j}$ in the special case of the (recursive) ODP Mack model. Likewise, let $\rho^{NR}_{k,j+m,j+m+n|j}$ apply to the (non-recursive) ODP cross-classified model.

Further, let $\pi_{j+m,j+m+n|j}$ denote the ratio $D_{j+m,j+m+n|j} / B_{j+m,j+m+n|j}$.

With subscripts suppressed, ρ^R and ρ^{NR} are related through π as follows. By (4.6),

$$B = 1 \Big/ \left(\rho^R\right)^2 - 1$$

Then, by (4.15),

$$\left(\rho^{NR}\right)^2 = 1 \Big/ \left\{ 1 + \pi \left[1 \Big/ \left(\rho^R\right)^2 - 1 \right] \right\}$$

and hence

$$\rho^{NR} = \pi^{-\frac{1}{2}} \rho^R \Big/ \left[1 + \frac{1-\pi}{\pi} \left(\rho^R\right)^2 \right]^{\frac{1}{2}} \tag{4.23}$$

For comparative purposes, it is useful to convert (4.6) and (4.10) for the ODP Mack model into a form involving $\beta's$ as in (4.14).

Note that (4.10) may be may be expressed in the alternative form

$$B_{j+m,j+m+n|j} = \frac{\sum_{i=j+m}^{j+m+n-1} \phi_{i+1} \; f_{j+m+n-1}^2 \cdots f_{i+1}^2 \left(f_i - 1 \right) f_{i-1} \cdots f_1}{\sum_{i=j}^{j+m-1} \phi_{i+1} \; f_{j+m+n-1}^2 \cdots f_{i+1}^2 \left(f_i - 1 \right) f_{i-1} \cdots f_1}$$

$$= \frac{\sum_{i=j+m}^{j+m+n-1} \phi_{i+1} \; \beta_{i+1} \left(f_{j+m+n-1}^2 \cdots f_{i+1}^2 \right)}{\sum_{i=j}^{j+m-1} \phi_{i+1} \; \beta_{i+1} \left(f_{j+m+n-1}^2 \cdots f_{i+1}^2 \right)} \tag{4.24}$$

by (4.17).

Theorem 4.4. Consider an ODPM-regular data array $\mathcal{D}_k^+$, and a row k within it that is not identically zero. Then

(a) $f_{j+m}^2 \le D_{j+m,j+m+n|j} / B_{j+m,j+m+n|j} \le f_{j+m+n-1}^2 \cdots f_{j+1}^2$

(b) $\pi_{k,j+m,j+m+n|j} \ge 1$. Hence $\rho_{k,j+m,j+m+n|j}^R \ge \rho_{k,j+m,j+m+n|j}^{NR}$.

(c) $\pi_{k,j+m,j+m+n|j} \to 1$ as $j \to \infty$. Hence $\rho_{k,j+m,j+m+n|j}^R / \rho_{k,j+m,j+m+n|j}^{NR} \to 1$ as $j \to \infty$.

Proof. (a) The largest multiplier of $\phi_{i+1} \beta_{i+1}$ in the numerator of (4.24) is $f_{j+m+n-1}^2 \cdots f_{j+m+1}^2$ (for $i = j+m$) while the smallest multiplier in the denominator is $f_{j+m+n-1}^2 \cdots f_{j+m}^2$ $\left(i = j+m-1\right)$. By (4.16), this proves that

$$B_{j+m,j+m+n|j} / D_{j+m,j+m+n|j} \le \left(f_{j+m}^2 \right)^{-1}$$

and hence the left inequality of (a).

The right inequality is similarly proved by considering the case $i = j+m+n-1$ in the numerator of (4.24) and $i = j$ in the denominator.

(b) Since all f factors are not less than unity, it follows from (a) that

$$B_{j+m,j+m+n|j} \le D_{j+m,j+m+n|j}$$

This, combined with (4.6) and (4.15), yields

$$\rho_{k,j+m,j+m+n|j}^R \ge \rho_{k,j+m,j+m+n|j}^{NR}$$

(c) As $j \to \infty$, $f_i \to 1$ for all $i \geq j$ in order that $\mathrm{E}\left[X_{kj}\right] = X_{k,K-k+1} f_{K-k+1} f_{K-k+2} \cdots f_{j-1}$ should converge as $j \to \infty$. It then follows from (a) that

$$D_{j+m,j+m+n|j} / B_{j+m,j+m+n|j} \to 1 \text{ as } j \to \infty$$

This, combined with (4.6) and (4.15), yields the stated result. □

5. Conclusion

The ODP Mack model is a special case of the Mack model and there is a simple translation between their correlation structures (Section 3.2.2).

The respective correlation structures associated with the recursive and non-recursive models considered here show a number of similarities but also distinct dissimilarities.

Theorems 4.2 and 4.3 show that, in both cases, correlation decreases with increasing time separation of future observations. The same theorems show that, in both cases, correlations $\rho_{k,j+m,j+m+n|j}$ generally increase as the dispersion coefficients of observations (σ_i^2 for the Mack model, and ϕ_i for the ODP Mack or ODP cross-classified model) up to time $j+m$ increase and as the dispersion of observations beyond this decreases.

However, the dependency of correlations on the mean development factors f_i differs as between the recursive and non-recursive models. For full details, see Theorems 4.2(d) and 4.3(d). In broad terms, increasing age-to-age factors cause correlations within the recursive models to increase and within the non-recursive models to decrease, though these results are subject to side-conditions that involve interaction between the age-to-age factors and dispersion coefficients.

If comparison is made between corresponding correlations in recursive and non-recursive models that are subject to consistent parameters, it is found that the recursive correlation is always the larger. However, as the development period on which the correlation between future observations is conditioned moves further into the development tail, the recursive and non-recursive correlations converge. Full details appear in Theorem 4.4.

References

England P D & Verrall R J (2002). Stochastic claims reserving in general insurance. **British Actuarial Journal**, 8(iii), 443-518.

Hachemeister C A & Stanard J N (1975). IBNR claims count estimation with static lag functions. **Spring meeting of the Casualty Actuarial Society.**

Mack T (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. **Astin Bulletin**, 23(2), 213-225.

Renshaw AE & Verrall RJ (1998). A stochastic model underlying the chain-ladder technique. **British Actuarial Journal**, 4(iv), 903-923.

Taylor G (2000). **Loss reserving: an actuarial perspective**. Kluwer Academic Publishers, Boston.

Taylor G (2011). Maximum likelihood and estimation efficiency of the chain ladder. **Astin Bulletin**, 41(1), 131-155.

Verrall RJ (2000). An investigation into stochastic claims reserving models and the chain-ladder technique. **Insurance: mathematics and economics**, 26(1), 91-99.