

# Multivariate Chain–Ladder

By Carsten Pröhl and Klaus D. Schmidt

Lehrstuhl für Versicherungsmathematik  
Technische Universität Dresden

## Abstract

In the present paper we propose a multivariate version of the chain–ladder method. The multivariate chain–ladder method is based on a stochastic model which is a multivariate version of the model of Schnaus. It is suitable for portfolios consisting of subportfolios with a certain dependence structure and resolves in a certain sense the problem of non–additivity of the univariate chain–ladder method.

## 1 Introduction

The chain–ladder method is a widely used method of loss reserving which can be modified and generalized in various ways; see Radtke and Schmidt [2004] for a survey of various methods of loss reserving and underlying stochastic models.

The chain–ladder method applies to a single run–off triangle and it is well–known that the chain–ladder predictors of the non–observable (future) aggregate claims of a portfolio consisting of several sub–portfolios differ, except for unrealistic special cases, from the respective sums of the chain–ladder predictors of the non–observable aggregate claims of the sub–portfolios; see Ajne [1994].

The reason for the non–coincidence of the chain–ladder predictors of sums and the sums of chain–ladder predictors has its origin in the univariate character of the chain–ladder method which neglects the dependence structure existing between the sub–portfolios of a portfolio.

In a recent paper published in this journal, Braun [2004] has proposed a bivariate stochastic model for loss reserving which extends the univariate model of Mack [1993] and takes into account correlation between two subportfolios of a portfolio. In the spirit of Mack [1993], Braun [2004] used his model as a foundation for the construction of estimators of the prediction errors of the univariate chain–ladder predictors, but he did not use his model to replace the univariate chain–ladder predictors by bivariate ones reflecting the correlation structure.

In the present paper, we propose a multivariate model which extends the univariate model of Schnaus, presented in Schmidt and Schnaus [1996], and the bivariate model of Braun [2004]. Under the assumptions of this model, we deduce multivariate chain–ladder predictors which are suitable for portfolios consisting of subportfolios with a certain dependence structure, which are optimal with respect to an essentially classical optimality criterion, and which resolve in a certain sense the problem of non–additivity of the univariate chain–ladder method.

Throughout this paper, let  $(\Omega, \mathcal{F}, P)$  be a probability space on which all random variables are defined. We assume that all random variables are square integrable and that all random vectors and random matrices have square integrable coordinates. Moreover, all equalities and inequalities involving random variables are understood to hold almost surely with respect to the probability measure  $P$ .

## 2 Univariate Chain–Ladder Prediction

In the present section we consider a single portfolio which is described by a family  $\{S_{i,k}\}_{i,k \in \{0,1,\dots,n\}}$  of random variables. We interpret  $S_{i,k}$  as the aggregate claim size of *accident year*  $i$  and *development year*  $k$  and we assume that  $S_{i,k}$  is *observable* for *calendar year*  $i + k \leq n$  and *non–observable* for *calendar year*  $i + k > n$ . Then the observable claim sizes can be represented by the following *run–off triangle*:

Accident Year	Development Year								
	0	1	...	$k$	...	$n-i$	...	$n-1$	$n$
0	$S_{0,0}$	$S_{0,1}$	...	$S_{0,k}$	...	$S_{0,n-i}$	...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$	...	$S_{1,k}$	...	$S_{1,n-i}$	...	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
$i$	$S_{i,0}$	$S_{i,1}$	...	$S_{i,k}$	...	$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$	...	$S_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
$n$	$S_{n,0}$								

We assume henceforth that each of the random variables  $S_{i,k}$  is strictly positive.

For  $i \in \{0, 1, \dots, n\}$  and  $k \in \{1, \dots, n\}$ , the random variable

$$F_{i,k} := \frac{S_{i,k}}{S_{i,k-1}}$$

is said to be the *individual development factor* of accident year  $i$  and development year  $k$ .

We can now define the *univariate chain–ladder method*:

**Univariate Chain–Ladder Method:** *The univariate chain–ladder method is defined by letting*

$$S_{i,n-i}^{\text{CL}} := S_{i,n-i}$$

as well as

$$F_k^{\text{CL}} := \sum_{j=0}^{n-k} \frac{S_{j,k-1}}{\sum_{h=0}^{n-k} S_{h,k-1}} F_{j,k}$$

and

$$S_{i,k}^{\text{CL}} := S_{i,k-1}^{\text{CL}} F_k^{\text{CL}}$$

for all  $i, k \in \{1, \dots, n\}$  such that  $i+k > n$ .

The random variables  $F_k^{\text{CL}}$  and  $S_{i,k}^{\text{CL}}$  are said to be the *univariate chain–ladder factor* of development year  $k$  and the *univariate chain–ladder predictor* of the aggregate claim size  $S_{i,k}$  of accident year  $i$  and development year  $k$ , respectively.

The univariate chain–ladder method is an algorithm which can be justified by various stochastic models and suitable optimality criteria. In particular, since each of the chain–ladder factors is a weighted mean of the observable individual development factors of the same development year, one may consider the question whether the weights used in the definition of the chain–ladder factors are optimal in some sense.

One of the models justifying the chain–ladder method is the model of Schnaus which is presented in Schmidt and Schnaus [1996] and extends the model of Mack [1993]. In Schmidt and Schnaus [1996] it is shown that the chain–ladder predictors for the aggregate claim sizes of the first non–observable calendar year  $n+1$  are indeed optimal under the assumptions of the model of Schnaus and with respect to a quite natural optimality criterion. Optimality of the chain–ladder predictors for the aggregate claim sizes of subsequent non–observable calendar years is studied in Schmidt [1997, 1999].

### 3 Multivariate Chain–Ladder Prediction

In the present section we consider  $m \in \mathbf{N}$  portfolios all having the same number of development years. The  $m$  portfolios may be interpreted as subportfolios of an aggregate portfolio.

For portfolio  $p \in \{1, \dots, m\}$ , we denote by

$$S_{i,k}^{(p)}$$

the aggregate claim size of accident year  $i \in \{0, 1, \dots, n\}$  and development year  $k \in \{0, 1, \dots, n\}$  and we denote by

$$F_{i,k}^{(p)}$$

the individual development factor of accident year  $i \in \{0, 1, \dots, n\}$  and development year  $k \in \{1, \dots, n\}$ .

For  $i, k \in \{0, 1, \dots, n\}$  we thus obtain the  $m$ -dimensional random vector of aggregate claims

$$\mathbf{S}_{i,k}$$

and for  $i \in \{0, 1, \dots, n\}$  and  $k \in \{1, \dots, n\}$  we obtain the  $m$ -dimensional random vector of individual development factors

$$\mathbf{F}_{i,k}$$

The observable aggregate claim size vectors are represented by the following run-off triangle:

Accident Year	Development Year								
	0	1	...	$k$	...	$n-i$	...	$n-1$	$n$
0	$\mathbf{S}_{0,0}$	$\mathbf{S}_{0,1}$	...	$\mathbf{S}_{0,k}$	...	$\mathbf{S}_{0,n-i}$	...	$\mathbf{S}_{0,n-1}$	$\mathbf{S}_{0,n}$
1	$\mathbf{S}_{1,0}$	$\mathbf{S}_{1,1}$	...	$\mathbf{S}_{1,k}$	...	$\mathbf{S}_{1,n-i}$	...	$\mathbf{S}_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
$i$	$\mathbf{S}_{i,0}$	$\mathbf{S}_{i,1}$	...	$\mathbf{S}_{i,k}$	...	$\mathbf{S}_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$\mathbf{S}_{n-k,0}$	$\mathbf{S}_{n-k,1}$	...	$\mathbf{S}_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$\mathbf{S}_{n-1,0}$	$\mathbf{S}_{n-1,1}$							
$n$	$\mathbf{S}_{n,0}$								

We assume henceforth that all coordinates of each of the vectors  $\mathbf{S}_{i,k}$  are strictly positive.

For a smooth definition of the multivariate chain-ladder method, it is convenient to represent the random vectors  $\mathbf{S}_{i,k}$  by the diagonal matrices

$$\Delta_{i,k} := \text{diag}(\mathbf{S}_{i,k})$$

Then we have  $\mathbf{S}_{i,k} = \Delta_{i,k} \mathbf{1}$  for all  $i, k \in \{0, 1, \dots, n\}$ , where  $\mathbf{1}$  is the vector in  $\mathbf{R}^m$  with all coordinates being equal to 1, and we also have

$$\mathbf{S}_{i,k} := \Delta_{i,k-1} \mathbf{F}_{i,k}$$

for all  $i \in \{0, 1, \dots, n\}$  and  $k \in \{1, \dots, n\}$ .

For  $k \in \{0, 1, \dots, n-1\}$  let  $\mathcal{G}_k$  denote the  $\sigma$ -algebra generated by the family  $\{\mathbf{S}_{i,l}\}_{i \in \{0,1,\dots,n\}, l \in \{0,1,\dots,k-1\}}$ .

While the univariate chain–ladder method can be defined as an algorithm without an underlying stochastic model, the multivariate chain–ladder method presented below is based on a stochastic model which

- in the case  $m = 1$  is the model of Schnaus presented in Schmidt and Schnaus [1996] and hence extends the model of Mack [1993] and
- in the case  $m = 2$  extends the model of Braun [2004]; see Section 6 for further details.

For  $k \in \{0, 1, \dots, n-1\}$ , let

$$\mathcal{G}_k$$

denote the  $\sigma$ –algebra generated by the family  $\{\mathbf{S}_{i,l}\}_{i \in \{0,1,\dots,n\}, l \in \{0,1,\dots,k-1\}}$ .

We can now introduce the *multivariate model of Schnaus*:

**The Multivariate Model of Schnaus:**

- (i) For each  $k \in \{1, \dots, n\}$ , there exists a  $\mathcal{G}_{k-1}$ –measurable random vector  $\Phi_k$  such that

$$E^{\mathcal{G}_{k-1}}(\mathbf{S}_{i,k}) = \Delta_{i,k-1} \Phi_k$$

holds for all  $i \in \{0, 1, \dots, n\}$ .

- (ii) For each  $k \in \{1, \dots, n\}$ , there exists a  $\mathcal{G}_{k-1}$ –measurable random matrix  $\Sigma_k$  which is symmetric and positive definite such that

$$\text{Cov}^{\mathcal{G}_{k-1}}(\mathbf{S}_{i,k}, \mathbf{S}_{j,k}) = \begin{cases} \Delta_{i,k-1}^{1/2} \Sigma_k \Delta_{i,k-1}^{1/2} & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j \end{cases}$$

holds for all  $i, j \in \{0, 1, \dots, n\}$ .

Under the assumptions of the multivariate model of Schnaus, we have

$$E^{\mathcal{G}_{k-1}}(\mathbf{F}_{i,k}) = \Phi_k$$

and

$$\text{Cov}^{\mathcal{G}_{k-1}}(\mathbf{F}_{i,k}, \mathbf{F}_{j,k}) = \begin{cases} \Delta_{i,k-1}^{-1/2} \Sigma_k \Delta_{i,k-1}^{-1/2} & \text{if } i = j \\ \mathbf{O} & \text{if } i \neq j \end{cases}$$

The following result is evident from the first part of Theorem 7.2 below:

**3.1 Theorem.** *Assume that the assumptions of the multivariate model of Schnaus are fulfilled. Then, for each  $k \in \{1, \dots, n\}$ , the random vector*

$$\mathbf{F}_k^{\text{CL}} := \left( \sum_{j=0}^{n-k} \Delta_{j,k-1}^{1/2} \Sigma_k^{-1} \Delta_{j,k-1}^{1/2} \right)^{-1} \sum_{j=0}^{n-k} \Delta_{j,k-1}^{1/2} \Sigma_k^{-1} \Delta_{j,k-1}^{1/2} \mathbf{F}_{j,k}$$

satisfies  $E^{\mathcal{G}_{k-1}}(\mathbf{F}_k^{\text{CL}}) = E^{\mathcal{G}_{k-1}}(\mathbf{F}_{i,k})$  for all  $i \in \{0, 1, \dots, n\}$  and is the unique random vector which, for each  $i \in \{0, \dots, n\}$  such that  $i+k > n$ , minimizes the conditional expected squared prediction error

$$E^{\mathcal{G}_{k-1}}\left(\left(\mathbf{F}_{i,k} - \mathbf{Y}\right)' \left(\mathbf{F}_{i,k} - \mathbf{Y}\right)\right)$$

over all  $m$ -dimensional random vectors  $\mathbf{Y}$  satisfying

$$\mathbf{Y} = \sum_{j=0}^{n-k} \mathbf{W}_{j,k} \mathbf{F}_{j,k}$$

with  $\mathcal{G}_{k-1}$ -measurable random matrices  $\mathbf{W}_{0,k}, \dots, \mathbf{W}_{n-k,k}$  such that  $\sum_{j=0}^{n-k} \mathbf{W}_{j,k} = \mathbf{I}$ .

The random vector  $\mathbf{F}_k^{\text{CL}}$  given by Theorem 3.1 is said to be the *multivariate chain-ladder factor* for development year  $k \in \{1, \dots, n\}$ . In the case  $m = 1$  it coincides with the usual (univariate) chain-ladder factor for development year  $k$ .

In the following result we change the point of view by looking at accident year  $i \in \{1, \dots, n\}$  instead of development year  $k \in \{1, \dots, n\}$ . In accident year  $i$ , the last observable vector of aggregate claims is  $\mathbf{S}_{i,n-i}$  and the first problem is to predict the first non-observable vector of aggregate claims  $\mathbf{S}_{i,n-i+1}$ . These vectors are related by the identity

$$\mathbf{S}_{i,n-i+1} = \Delta_{i,n-i} \mathbf{F}_{i,n-i+1}$$

It is therefore not too surprising that the best predictor of  $\mathbf{S}_{i,n-i+1}$  is obtained by replacing the vector of individual chain ladder factors  $\mathbf{F}_{i,n-i+1}$  in the previous identity by its best predictor given by Theorem 3.1:

**3.2 Corollary.** *Assume that the assumptions of the multivariate model of Schnaus are fulfilled. Then, for each  $i \in \{1, \dots, n\}$ , the random vector*

$$\mathbf{S}_{i,n-i+1}^{\text{CL}} := \Delta_{i,n-i} \left( \sum_{j=0}^{i-1} \Delta_{j,n-i}^{1/2} \Sigma_{n-i+1}^{-1} \Delta_{j,n-i}^{1/2} \right)^{-1} \sum_{j=0}^{i-1} \Delta_{j,n-i}^{1/2} \Sigma_{n-i+1}^{-1} \Delta_{j,n-i}^{1/2} \mathbf{F}_{j,n-i+1}$$

satisfies  $E^{\mathcal{G}_{n-i}}(\mathbf{S}_{i,n-i+1}^{\text{CL}}) = E^{\mathcal{G}_{n-i}}(\mathbf{S}_{i,n-i+1})$  and is the unique random vector which minimizes the conditional expected squared prediction error

$$E^{\mathcal{G}_{n-i}}\left(\left(\mathbf{S}_{i,n-i+1} - \mathbf{Y}\right)' \left(\mathbf{S}_{i,n-i+1} - \mathbf{Y}\right)\right)$$

over all  $m$ -dimensional random vectors  $\mathbf{Y}$  satisfying

$$\mathbf{Y} = \Delta_{i,n-i} \sum_{j=0}^{i-1} \mathbf{W}_{j,n-i+1} \mathbf{F}_{j,n-i+1}$$

with  $\mathcal{G}_{n-i}$ -measurable random matrices  $\mathbf{W}_{0,n-i+1}, \dots, \mathbf{W}_{i-1,n-i+1}$  such that

$$\sum_{j=0}^{i-1} \mathbf{W}_{j,n-i+1} = \mathbf{I}$$

Corollary 3.2 follows from the second part of Theorem 7.2 below.

We can now define the *multivariate chain–ladder method*:

**The Multivariate Chain–Ladder Method:** *Under the assumptions of the multivariate model of Schnaus, the multivariate chain–ladder method is defined by letting*

$$\mathbf{S}_{i,n-i}^{\text{CL}} := \mathbf{S}_{i,n-i}$$

for all  $i \in \{1, \dots, n\}$  and

$$\begin{aligned} \Delta_{i,k-1}^{\text{CL}} &:= \text{diag}(\mathbf{S}_{i,k-1}^{\text{CL}}) \\ \mathbf{S}_{i,k}^{\text{CL}} &:= \Delta_{i,k-1}^{\text{CL}} \mathbf{F}_k^{\text{CL}} \end{aligned}$$

for all  $i, k \in \{1, \dots, n\}$  such that  $i+k > n$ .

It can be shown as in the univariate case that, for each  $i \in \{1, \dots, n\}$ , the sequence  $\{\mathbf{S}_{i,k}^{\text{CL}}\}_{k \in \{n-i+1, \dots, n\}}$  satisfies the multivariate analogon of the sequential optimality criterion introduced by Schmidt [1999]:

**3.3 Corollary.** *Assume that the assumptions of the multivariate model of Schnaus are fulfilled. Then, for all  $i, k \in \{1, \dots, n\}$  such that  $i+k > n$ , the random vector*

$$\mathbf{S}_{i,k}^{\text{CL}}$$

*satisfies  $E^{\mathcal{G}_{k-1}}(\mathbf{S}_{i,k}^{\text{CL}}) = E^{\mathcal{G}_{k-1}}(\mathbf{S}_{i,k})$  and is the unique random vector which minimizes the conditional expected squared prediction error*

$$E^{\mathcal{G}_{k-1}}\left(\left(\mathbf{S}_{i,k} - \mathbf{Y}\right)' \left(\mathbf{S}_{i,k} - \mathbf{Y}\right)\right)$$

*over all  $m$ -dimensional random vectors  $\mathbf{Y}$  satisfying*

$$\mathbf{Y} = \Delta_{i,k-1}^{\text{CL}} \sum_{j=0}^{i-1} \mathbf{W}_{j,k} \mathbf{F}_{j,k}$$

*with  $\mathcal{G}_{k-1}$ -measurable random matrices  $\mathbf{W}_{0,k}, \dots, \mathbf{W}_{i-1,k}$  such that  $\sum_{j=0}^{i-1} \mathbf{W}_{j,k} = \mathbf{I}$ .*

The previous result can be applied to predict the sum of the coordinates of  $\mathbf{S}_{i,k}$ :

**3.4 Corollary.** *Assume that the assumptions of the multivariate model of Schnaus are fulfilled. Then, for all  $i, k \in \{1, \dots, n\}$  such that  $i+k > n$ , the random vector*

$$\mathbf{1}' \mathbf{S}_{i,k}^{\text{CL}}$$

*satisfies  $E^{\mathcal{G}_{k-1}}(\mathbf{1}' \mathbf{S}_{i,k}^{\text{CL}}) = E^{\mathcal{G}_{k-1}}(\mathbf{1}' \mathbf{S}_{i,k})$  and is the unique random vector which minimizes the conditional expected squared prediction error*

$$E^{\mathcal{G}_{k-1}}\left(\left(\mathbf{1}' \mathbf{S}_{i,k} - \mathbf{Y}\right)' \left(\mathbf{1}' \mathbf{S}_{i,k} - \mathbf{Y}\right)\right)$$

over all  $m$ -dimensional random vectors  $\mathbf{Y}$  satisfying

$$\mathbf{Y} = \mathbf{1}' \Delta_{i,k-1}^{\text{CL}} \sum_{j=0}^{i-1} \mathbf{W}_{j,k} \mathbf{F}_{j,k}$$

with  $\mathcal{G}_{k-1}$ -measurable random matrices  $\mathbf{W}_{0,k}, \dots, \mathbf{W}_{i-1,k}$  such that  $\sum_{j=0}^{i-1} \mathbf{W}_{j,k} = \mathbf{I}$ .

We remark that the problem of global optimality of the chain-ladder predictors has been solved in the negative in the case  $m = 1$ ; see Schmidt [1997].

## 4 Multivariate versus Univariate Chain-Ladder Prediction

In the case where  $k = n$  or  $\Sigma_k$  is diagonal, it is easily seen that the multivariate chain-ladder factor  $\mathbf{F}_k^{\text{CL}}$  does not depend on  $\Sigma_k$  and that, for each  $p \in \{1, \dots, m\}$ , the coordinates  $F_k^{\text{CL}(p)}$  of  $\mathbf{F}_k^{\text{CL}}$  satisfy

$$F_k^{\text{CL}(p)} = \sum_{j=0}^{n-k} \frac{S_{j,k-1}^{(p)}}{\sum_{h=0}^{n-k} S_{h,k-1}^{(p)}} F_{j,k}^{(p)}$$

and hence coincide with the chain-ladder factors for the subportfolio  $p$ .

In the case where each of the matrices  $\Sigma_1, \dots, \Sigma_{n-1}$  is diagonal, it follows that univariate chain-ladder prediction for every single subportfolio produces the same result as multivariate chain-ladder prediction for the aggregate portfolio.

In general, however, univariate chain-ladder prediction for every single subportfolio is not the same as multivariate chain-ladder prediction. The following partial result makes this statement more precise:

**4.1 Theorem.** *Assume that the assumptions of the model of Schnaus are fulfilled and assume that  $m = 2 \leq k < n$ . Then the following are equivalent:*

- (a) *Every coordinate of the chain-ladder predictor  $\mathbf{F}_k^{\text{CL}}$  depends only on the corresponding coordinates of  $\mathbf{F}_{0,k}, \dots, \mathbf{F}_{n-k,k}$ .*
- (b) *The matrix  $\Sigma_k$  is diagonal.*

**Proof.** Assume first that (a) holds. Then each of the matrices

$$\left( \sum_{h=0}^{n-k} \Delta_{h,k-1}^{1/2} \Sigma_k^{-1} \Delta_{h,k-1}^{1/2} \right)^{-1} \Delta_{j,k-1}^{1/2} \Sigma_k^{-1} \Delta_{j,k-1}^{1/2}$$

with  $j \in \{0, 1, \dots, n-k\}$  is diagonal and Theorem 7.3 implies that the matrices

$$\Delta_{j,k-1}^{1/2} \Sigma_k^{-1} \Delta_{j,k-1}^{1/2}$$

are either all diagonal or proportional to each other. In the first case, the matrix  $\Sigma_k$  is diagonal as well. In the second case, the matrices  $\Delta_{j,k-1}$  are proportional to each other such that there exist  $c_1, \dots, c_{n-k} \in (0, \infty)$  satisfying

$$\mathbf{S}_{j,k-1} = c_j \mathbf{S}_{0,k-1}$$

for all  $j \in \{1, \dots, n-k\}$  and hence

$$\text{Cov}^{\mathcal{G}^{k-2}}(\mathbf{S}_{j,k-1}, \mathbf{S}_{0,k-1}) = c_j \text{Var}^{\mathcal{G}^{k-2}}(\mathbf{S}_{0,k-1})$$

which is impossible since  $\text{Cov}^{\mathcal{G}^{k-2}}(\mathbf{S}_{j,k-1}, \mathbf{S}_{0,k-1}) = \mathbf{O} \neq \text{Var}^{\mathcal{G}^{k-2}}(\mathbf{S}_{0,k-1})$ . Therefore, (a) implies (b). The converse implication is obvious.  $\square$

## 5 The Problem of Non-Additivity of Chain-Ladder Predictors

It is well-known and not really surprising that, except for unrealistic cases, the chain-ladder predictors for the sum of run-off triangles differ from the sum of the chain-ladder predictors for the single run-off triangles; see Ajne [1994].

In some sense, the multivariate chain-ladder model resolves the problem of non-additivity of chain-ladder predictors. Since the multivariate chain-ladder model involves conditional correlation between the aggregate claims of a fixed development year, chain-ladder prediction for the sum of the aggregate claims over the subportfolios is meaningless since it ignores the correlation structure which relates the subportfolios to each other; this is also the case when the subportfolios are assumed to be uncorrelated. Therefore, the only reasonable predictor of  $\mathbf{1}'\mathbf{S}_{i,k}$  with  $i+k > n$  is the predictor  $\mathbf{1}'\mathbf{S}_{i,k}^{\text{CL}}$  which is the sum of the coordinates of the chain-ladder predictor of  $\mathbf{S}_{i,k}$ .

## 6 Related Models

The multivariate model of Schnaus presented in Section 3 is inspired by the bivariate model of Braun and its straightforward extension to the multivariate case which is defined as follows:

### The Multivariate Model of Braun:

- (i) *The accident years are independent in the sense that the family of  $\sigma$ -algebras  $\{\sigma(\{\mathbf{S}_{i,k}\}_{k \in \{0,1,\dots,n\}})\}_{i \in \{0,1,\dots,n\}}$  is independent.*
- (ii) *For each  $k \in \{1, \dots, n\}$ , there exists vector  $\varphi_k$  such that*

$$E^{\mathcal{G}^{k-1}}(\mathbf{S}_{i,k}) = \Delta_{i,k-1} \varphi_k$$

*holds for all  $i \in \{0, 1, \dots, n\}$ .*

- (iii) For each  $k \in \{1, \dots, n\}$ , there exists a matrix  $\boldsymbol{\sigma}_k$  which is symmetric and positive definite such that

$$\text{Var}^{\mathcal{G}_{k-1}}(\mathbf{S}_{i,k}) = \boldsymbol{\Delta}_{i,k-1}^{1/2} \boldsymbol{\sigma}_k \boldsymbol{\Delta}_{i,k-1}^{1/2}$$

holds for all  $i, j \in \{0, 1, \dots, n\}$ .

In the multivariate model of Brown, which

- in the case  $m = 1$  is due to Mack [1993] and
  - in the case  $m = 2$  is due to Braun [2004],
- the vector  $\boldsymbol{\varphi}_k$  and the matrix  $\boldsymbol{\sigma}_k$  are non-random.

It can be shown as in the univariate case that the multivariate model of Brown is a special case of the multivariate model of Schnaus; see Schmidt and Schnaus [1996]. Also, since the univariate model of Schnaus is strictly more general than the model of Mack, it is clear that the multivariate model of Schnaus is strictly more general than the multivariate model of Braun; see Hess and Schmidt [2002].

Braun [2004] used his bivariate model to construct estimators for the prediction errors of the univariate chain-ladder predictors of two correlated portfolios which take into account correlation between the portfolios and which are intended as to improve the estimators proposed by Mack [1993] neglecting correlation. As it is the case for the estimators proposed by Mack, the estimators proposed by Braun are constructed in a reasonable but heuristic way; in particular, in both cases it is not known whether these estimators have any particular statistical properties like, e. g., unbiasedness.

In the bivariate model of Braun, it follows from Theorem 4.1 that the univariate chain-ladder predictors are not optimal unless the matrices  $\boldsymbol{\sigma}_k$  are diagonal. This means that the estimators of the prediction errors proposed by Braun [2004] refer to non-optimal predictors of the non-observable aggregate claims. Thus, instead of improving the estimators of the prediction errors of non-optimal predictors, it would be more reasonable to start with the optimal predictors provided by the multivariate chain-ladder method under the assumptions of the multivariate model of Schnaus and then to try to develop reasonable estimators for the prediction errors of the multivariate chain-ladder predictors.

Another bivariate model of loss reserving, which is loosely related to the multivariate model of Schnaus, is the model of Quarg and Mack [2004]. Under the assumptions of their model, Quarg and Mack propose bivariate chain-ladder predictors for the paid and incurred aggregate claims of the same portfolio with the aim of reducing the gap between the univariate chain-ladder predictors for the paid and incurred aggregate claims of the same portfolio. The model of Quarg and Mack is not included in the multivariate model of Schnaus since it assumes a conditional correlation structure within the accident years instead of a completely specified conditional correlation structure between the paid and incurred aggregate claims.

## 7 Appendix

Throughout this appendix, let  $\mathcal{H}$  be a sub- $\sigma$ -algebra of  $\mathcal{F}$ .

We start with a general theorem:

**7.1 Theorem.** *Let  $\mathbf{X}$  and  $\mathbf{Z}$  be random vectors and assume that there exist an  $\mathcal{H}$ -measurable random vector  $\mathbf{H}$  and  $\mathcal{H}$ -measurable random matrices  $\mathbf{C}$  and  $\mathbf{E}$  such that  $E^{\mathcal{H}}(\mathbf{X}) = \mathbf{C}\mathbf{H}$  and  $E^{\mathcal{H}}(\mathbf{Z}) = \mathbf{E}\mathbf{H}$ . Let  $\mathbf{\Psi} := \text{cov}^{\mathcal{H}}(\mathbf{X}, \mathbf{Z})$  and assume that  $\mathbf{\Sigma} := \text{var}^{\mathcal{H}}(\mathbf{Z})$  is almost surely invertible.*

*Then the random vector*

$$\widehat{\mathbf{Y}} := \left( \mathbf{\Psi} + (\mathbf{C} - \mathbf{\Psi} \mathbf{\Sigma}^{-1} \mathbf{E})(\mathbf{E}' \mathbf{\Sigma}^{-1} \mathbf{E})^{-1} \mathbf{E}' \right) \mathbf{\Sigma}^{-1} \mathbf{Z}$$

*satisfies  $E^{\mathcal{H}}(\widehat{\mathbf{Y}}) = E^{\mathcal{H}}(\mathbf{Y})$  and is the unique random vector which minimizes the  $\mathcal{H}$ -conditional expected squared prediction error*

$$E^{\mathcal{H}}\left(\left(\mathbf{X} - \mathbf{Y}\right)' \left(\mathbf{X} - \mathbf{Y}\right)\right)$$

*over all random vectors  $\mathbf{Y}$  having the same dimension as  $\mathbf{X}$  and satisfying  $\mathbf{Y} = \mathbf{A}\mathbf{Z}$  with an  $\mathcal{H}$ -measurable random matrix  $\mathbf{A}$  such that  $\mathbf{A}\mathbf{E} = \mathbf{C}$ .*

*Moreover, if  $\mathbf{D}$  is an  $\mathcal{H}$ -measurable random matrix of suitable dimension, then the random vector  $\mathbf{D}\widehat{\mathbf{Y}}$  satisfies  $E^{\mathcal{H}}(\mathbf{D}\widehat{\mathbf{Y}}) = E^{\mathcal{H}}(\mathbf{D}\mathbf{Y})$  and is the unique random vector which minimizes the  $\mathcal{H}$ -conditional expected squared prediction error*

$$E^{\mathcal{H}}\left(\left(\mathbf{D}\mathbf{X} - \mathbf{Y}\right)' \left(\mathbf{D}\mathbf{X} - \mathbf{Y}\right)\right)$$

*over all random vectors  $\mathbf{Y}$  having the same dimension as  $\mathbf{X}$  and satisfying  $\mathbf{Y} = \mathbf{A}\mathbf{Z}$  with an  $\mathcal{H}$ -measurable random matrix  $\mathbf{A}$  such that  $\mathbf{A}\mathbf{E} = \mathbf{D}\mathbf{C}$ .*

**Proof.** The random matrix

$$\widehat{\mathbf{A}} := \left( \mathbf{\Psi} + (\mathbf{C} - \mathbf{\Psi} \mathbf{\Sigma}^{-1} \mathbf{E})(\mathbf{E}' \mathbf{\Sigma}^{-1} \mathbf{E})^{-1} \mathbf{E}' \right) \mathbf{\Sigma}^{-1}$$

is  $\mathcal{H}$ -measurable and satisfies  $\widehat{\mathbf{A}}\mathbf{E} = \mathbf{C}$ . For every  $\mathcal{H}$ -measurable random matrix  $\mathbf{A}$  satisfying  $\mathbf{A}\mathbf{E} = \mathbf{C}$  we have

$$\begin{aligned} E^{\mathcal{H}}(\mathbf{A}\mathbf{Z}) &= \mathbf{A} E^{\mathcal{H}}(\mathbf{Z}) \\ &= \mathbf{A}\mathbf{E}\mathbf{H} \\ &= \mathbf{C}\mathbf{H} \\ &= E^{\mathcal{H}}(\mathbf{X}) \end{aligned}$$

and

$$\text{Cov}^{\mathcal{H}}\left(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z}, \mathbf{A}\mathbf{Z} - \widehat{\mathbf{A}}\mathbf{Z}\right) = \text{Cov}^{\mathcal{H}}\left(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z}, (\mathbf{A} - \widehat{\mathbf{A}})\mathbf{Z}\right)$$

$$\begin{aligned}
&= \text{Cov}^{\mathcal{H}}(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z}, \mathbf{Z})(\mathbf{A} - \widehat{\mathbf{A}})' \\
&= (\boldsymbol{\Psi} - \widehat{\mathbf{A}}\boldsymbol{\Sigma})(\mathbf{A} - \widehat{\mathbf{A}})' \\
&= -(\mathbf{C} - \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\mathbf{E})(\mathbf{E}'\boldsymbol{\Sigma}^{-1}\mathbf{E})^{-1}\mathbf{E}'(\mathbf{A} - \widehat{\mathbf{A}})' \\
&= -(\mathbf{C} - \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\mathbf{E})(\mathbf{E}'\boldsymbol{\Sigma}^{-1}\mathbf{E})^{-1}(\mathbf{A}\mathbf{E} - \widehat{\mathbf{A}}\mathbf{E})' \\
&= \mathbf{O}
\end{aligned}$$

which yields

$$\begin{aligned}
\text{Var}^{\mathcal{H}}(\mathbf{X} - \mathbf{AZ}) &= \text{Var}^{\mathcal{H}}\left(\left(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z}\right) - \left(\mathbf{AZ} - \widehat{\mathbf{A}}\mathbf{Z}\right)\right) \\
&= \text{Var}^{\mathcal{H}}(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z}) + \text{Var}^{\mathcal{H}}(\mathbf{AZ} - \widehat{\mathbf{A}}\mathbf{Z})
\end{aligned}$$

and hence

$$\begin{aligned}
E^{\mathcal{H}}\left(\left(\mathbf{X} - \mathbf{AZ}\right)'(\mathbf{X} - \mathbf{AZ})\right) &= \text{trace}\left(E^{\mathcal{H}}\left(\left(\mathbf{X} - \mathbf{AZ}\right)(\mathbf{X} - \mathbf{AZ})'\right)\right) \\
&= \text{trace}\left(\text{Var}^{\mathcal{H}}(\mathbf{X} - \mathbf{AZ})\right) \\
&= \text{trace}\left(\text{Var}^{\mathcal{H}}(\mathbf{X} - \widehat{\mathbf{A}}\mathbf{Z})\right) \\
&\quad + \text{trace}\left(\text{Var}^{\mathcal{H}}(\mathbf{AZ} - \widehat{\mathbf{A}}\mathbf{Z})\right)
\end{aligned}$$

Since  $\text{trace}(\text{Var}^{\mathcal{H}}(\mathbf{AZ} - \widehat{\mathbf{A}}\mathbf{Z})) \geq 0$ , the first assertion follows.

The final assertion is then immediate from what has been shown before and the identities  $E^{\mathcal{H}}(\mathbf{DX}) = \mathbf{DCH}$  and  $\text{Cov}^{\mathcal{H}}(\mathbf{DX}, \mathbf{Z}) = \mathbf{D}\boldsymbol{\Psi}$ .  $\square$

The final part of the previous result applies, in particular, to the case where the matrix  $\mathbf{D}$  is the transpose of a unit vector. This means that the best predictor of a coordinate of  $\mathbf{X}$  is identical with the corresponding coordinate of the best predictor of  $\mathbf{X}$ .

The previous result can be specialized as follows:

**7.2 Theorem.** *Let  $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_r$  be random vectors and assume that exists an  $\mathcal{H}$ -measurable random vector satisfying  $E^{\mathcal{H}}(\mathbf{X}_i) = \mathbf{H}$  for all  $i \in \{0, 1, \dots, r\}$ . Assume further that  $\text{cov}^{\mathcal{H}}(\mathbf{X}_i, \mathbf{X}_j) = \mathbf{O}$  holds for all  $i, j \in \{0, 1, \dots, r\}$  such that  $i \neq j$  and that, for each  $i \in \{1, \dots, r\}$ ,  $\boldsymbol{\Sigma}_i := \text{var}^{\mathcal{H}}(\mathbf{X}_i)$  is almost surely invertible. Then the random vector*

$$\widehat{\mathbf{Y}} := \left( \sum_{i=1}^r \boldsymbol{\Sigma}_i^{-1} \right)^{-1} \sum_{i=1}^r \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i$$

*satisfies  $E^{\mathcal{H}}(\widehat{\mathbf{Y}}) = E^{\mathcal{H}}(\mathbf{X}_0)$  and is the unique random vector which minimizes the  $\mathcal{H}$ -conditional expected squared prediction error*

$$E^{\mathcal{H}}\left(\left(\mathbf{X}_0 - \mathbf{Y}\right)'(\mathbf{X}_0 - \mathbf{Y})\right)$$

over all random vectors  $\mathbf{Y}$  having the same dimension as  $\mathbf{X}_0$  and satisfying  $\mathbf{Y} = \sum_{i=1}^r \mathbf{A}_i \mathbf{X}_i$  with  $\mathcal{H}$ -measurable random matrices  $\mathbf{A}_1, \dots, \mathbf{A}_r$  such that  $\sum_{i=1}^r \mathbf{A}_i = \mathbf{I}$ . Moreover, if  $\mathbf{D}$  is an  $\mathcal{H}$ -measurable random matrix of suitable dimension, then the random vector  $\mathbf{D}\hat{\mathbf{Y}}$  satisfies  $E^{\mathcal{H}}(\mathbf{D}\hat{\mathbf{Y}}) = E^{\mathcal{H}}(\mathbf{D}\mathbf{X}_0)$  and is the unique random vector which minimizes the  $\mathcal{H}$ -conditional expected squared prediction error

$$E^{\mathcal{H}}\left(\left(\mathbf{D}\mathbf{X}_0 - \mathbf{Y}\right)' \left(\mathbf{D}\mathbf{X}_0 - \mathbf{Y}\right)\right)$$

over all random vectors  $\mathbf{Y}$  having the same dimension as  $\mathbf{X}_0$  and satisfying  $\mathbf{Y} = \sum_{i=1}^r \mathbf{A}_i \mathbf{X}_i$  with  $\mathcal{H}$ -measurable random matrices  $\mathbf{A}_1, \dots, \mathbf{A}_r$  such that  $\sum_{i=1}^r \mathbf{A}_i = \mathbf{D}$ .

Theorem 7.2 is a conditional version of a result to be found e. g. in Schmidt [2004]. Of course, the remark following Theorem 7.1 applies to Theorem 7.2 as well.

The following auxiliary result concerns diagonal matrices:

**7.3 Theorem.** *Assume that  $\mathbf{A}_1, \dots, \mathbf{A}_r$  are two-dimensional matrices which are symmetric and positive definite. Then the following are equivalent:*

(a) *For each  $i \in \{1, \dots, r\}$ , there exists a positive definite diagonal matrix  $\mathbf{D}_i$  such that*

$$\left(\sum_{j=1}^r \mathbf{A}_j\right)^{-1} \mathbf{A}_i = \mathbf{D}_i$$

(b) *The matrices  $\mathbf{A}_1, \dots, \mathbf{A}_r$  are diagonal or proportional to each other.*

**Proof.** Assume first that (a) holds. For all  $i \in \{1, \dots, r\}$ , we have

$$\mathbf{A}_i = \left(\sum_{j=1}^r \mathbf{A}_j\right) \mathbf{D}_i$$

The matrix  $\sum_{j=1}^r \mathbf{A}_j$  is either diagonal or not. If  $\sum_{j=1}^r \mathbf{A}_j$  is diagonal, then the previous identity implies that  $\mathbf{A}_1, \dots, \mathbf{A}_r$  are diagonal as well. If  $\sum_{j=1}^r \mathbf{A}_j$  is not diagonal, then the previous identity together with the symmetry of  $\mathbf{A}_1, \dots, \mathbf{A}_r$  yields the existence of  $d_1, \dots, d_r \in (0, \infty)$  such that  $\mathbf{D}_i = d_i \mathbf{I}$  and hence

$$d_i^{-1} \mathbf{A}_i = \sum_{j=1}^r \mathbf{A}_j$$

which implies that  $\mathbf{A}_1, \dots, \mathbf{A}_r$  are proportional to each other. Therefore, (a) implies (b). The converse implication is obvious.  $\square$

## Acknowledgement

This paper was written while the second author was at leave at ETH Zürich in spring 2005. He would like to thank Paul Embrechts for the invitation and his kind hospitality and Forschungsinstitut für Mathematik at ETH Zürich for financial support.

## References

- Ajne, B. [1994]:** *Additivity of chain-ladder projections.* ASTIN Bull. **24**, 313–318.
- Braun, C. [2004]:** *The prediction error of the chain-ladder method applied to correlated run-off triangles.* ASTIN Bull. **34**, 399–423.
- Hess, K. T., and Schmidt, K. D. [2002]:** *A comparison of models for the chain-ladder method.* Insurance, Mathematics and Economics **31**, 351–364.
- Mack, T. [1993]:** *Distribution-free calculation of the standard error of chain-ladder reserve estimates.* ASTIN Bull. **23**, 213–225.
- Quarg, G., and Mack, T. [2004]:** *Munich chain-ladder – A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses.* Blätter DGVM **xx**, 597–630.
- Radtko, M., and Schmidt, K. D. (eds.) [2004]:** *Handbuch zur Schadenreservierung.* Karlsruhe: Verlag Versicherungswirtschaft.
- Schmidt, K. D. [1997]:** *Non-optimal prediction by the chain-ladder method.* Insurance, Mathematics and Economics **21**, 17–24.
- Schmidt, K. D. [1999]:** *Chain ladder prediction and asset liability management.* Blätter DGVM **24**, 1–9.
- Schmidt, K. D. [2004]:** *Prediction.* In: *Encyclopedia of Actuarial Science*, vol. 3, pp. 1317–1321. Chichester: Wiley.
- Schmidt, K. D., and Schnaus, A. [1996]:** *An extension of Mack’s model for the chain-ladder method.* ASTIN Bull. **26**, 247–262.

Carsten Pröhl  
Klaus D. Schmidt  
Lehrstuhl für Versicherungsmathematik  
Technische Universität Dresden  
D-01062 Dresden

e-mail: [schmidt@math.tu-dresden.de](mailto:schmidt@math.tu-dresden.de)

10th June 2005