Prediction Error of the Multivariate Chain Ladder Reserving Method

Michael Merz* and Mario V. Wüthrich†

Abstract

In this paper we consider the claims reserving problem in a multivariate context: that is, we study the multivariate chain-ladder (CL) method for a portfolio of N correlated runoff triangles based on multivariate age-to-age factors. This method allows for a simultaneous study of individual runoff subportfolios and facilitates the derivation of an estimator for the mean square error of prediction (MSEP) for the CL predictor of the ultimate claim of the total portfolio. However, unlike the already existing approaches we replace the univariate CL predictors with multivariate ones. These multivariate CL predictors reflect the correlation structure between the subportfolios and are optimal in terms of a classical optimality criterion, which leads to an improvement of the estimator for the MSEP. Moreover, all formulas are easy to implement on a spreadsheet because they are in matrix notation. We illustrate the results by means of an example.

1. Motivation

1.1 Claims Reserving for Correlated Runoff Portfolios

Often claims reserves are the largest position on the liability side of a general insurance company's balance sheet. Therefore, given the available information about the past, the estimation of adequate claims reserves as well as the quantification of the uncertainties in these reserves is a major task in actuarial practice and science (see, e.g., Taylor 2000; Casualty Actuarial Society 1990; Teugels and Sundt 2004).

In this paper we consider the claims reserving problem in a multivariate context: that is, we consider a portfolio consisting of several correlated runoff subportfolios, and we use the chain-ladder (CL) method for estimating the claims reserves. In actuarial practice the mean square error of prediction (MSEP) is the most popular measure to quantify the uncertainties in claims reserves, and so we provide an MSEP estimator. Such studies of uncertainties for correlated runoff subportfolios are crucial in the development of new solvency guidelines for the quantification of risk profiles for different insurance companies. An alternative idea to the simultaneous study of several individual runoff subportfolios is to calculate the reserves and their uncertainties only for the total aggregated runoff portfolio. However, in most cases this is not a promising solution to the claims reserving problem on correlated runoff subportfolios because the total aggregated runoff portfolio usually does not satisfy the same homogeneity assumptions as the individual subportfolios do (e.g., the assumptions of the CL reserving method; see Anje 1994; Klemmt 2004).

* Michael Merz is Assistant Professor for Statistics, Risk and Insurance in the Faculty of Economics, University of Tübingen, Tübingen, D-72074 Germany, michael.merz@uni-tuebingen.de.

† Mario V. Wüthrich is Senior Researcher and Lecturer in the Department of Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland, wueth@math.ethz.ch.

1.2 Multivariate Claims Reserving Methods

Braun (2004), Pröhl and Schmidt (2005), Schmidt (2006), and Merz and Wüthrich (2008) have considered a multivariate version for the distribution-free CL model. Their work differs on the point of view of how the multidimensional CL parameters are estimated. Braun (2004) and Merz and Wüthrich (2008) use the classical (univariate) estimators, whereas Pröhl and Schmidt (2005) and Schmidt (2006) use multivariate estimators taking into account the dependence structure between the coordinates and that are optimal in terms of a classical optimality criterion. On the one hand Braun (2004) and Merz and Wüthrich (2008) provide an estimator of the MSEP for several correlated runoff portfolios. On the other hand, the studies of Pröhl and Schmidt (2005) and Schmidt (2006) for the multivariate estimators do not go beyond the study of first moments. In the present paper we close this gap and provide a MSEP estimator for the multivariate approach of Pröhl and Schmidt (2005) and Schmidt (2006).

The simultaneously study of several correlated runoff subportfolios is motivated by the following:

- In practice it is quite natural to subdivide a general insurance runoff portfolio into several correlated subportfolios, such that each subportfolio satisfies certain homogeneity properties (e.g., the CL assumptions).
- It addresses the problem of dependence between the runoff portfolios of different lines of business (e.g., bodily injury claims in auto liability and in general liability business).
- The multivariate approach has the advantage that by observing one runoff subportfolio we learn about the behavior of the other runoff subportfolios (e.g. subportfolios of small and large claims).
- It resolves the problem of additivity (i.e., the estimators of the ultimate claims for the whole runoff portfolio are obtained by summation over the estimators of the ultimate claims for the individual runoff subportfolios).

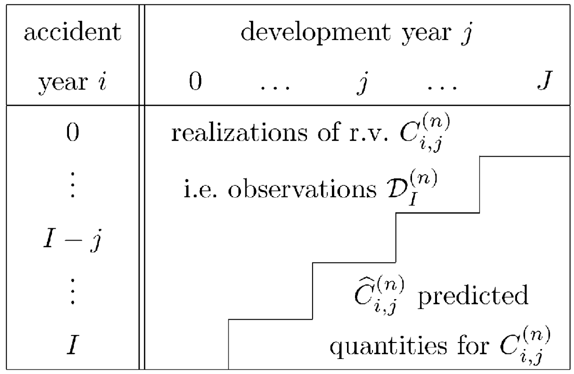
However, in the case of correlated subportfolios the analytical derivation of an estimate for the resulting conditional MSEP of the total runoff portfolio is rather sophisticated. In this respect the simulation-based approaches by Brehm (2002), Kirschner, Kerley, and Isaacs (2002), and Taylor and McGuire (2005, 2007), which extend the bootstrapping technique from a single runoff portfolio to several correlated runoff portfolios for estimating the uncertainties in the claims reserves, may seem more advantageous at times. However, in these approaches it is not always clear which statistical properties the produced samples actually have, and which correlations between the original runoff triangles are captured in the samples at all. Moreover, analytic solutions provide additional insight and intuition, and closed formulas are often easier to apply than purely numerical procedures. In addition, it is often difficult to interpret numerical results because parameter sensitivities are much harder to analyze. Therefore, apart from purely numerical procedures, we believe that analytical solutions provide valuable insights.

In Section 2 we provide the notation and data structure for our multivariate framework. In Section 3 we define the multivariate CL model and derive the properties of the multivariate CL estimators and predictors of the CL factors and ultimate claims. In Section 4 we give an estimation procedure for the conditional MSEP in the multivariate CL method. Our main results are stated in Results 4.8 and 4.10, which quantify the prediction errors in the multivariate CL method. Section 5 is dedicated to the estimation of the model parameters. Finally, in Section 6 we state an example and compare our results for the multivariate CL method to the results of Braun (2004) and Merz and Wüthrich (2008).

2. CLAIMS DEVELOPMENT TRIANGLE AND NOTATION

We assume that the subportfolios consist of $N \geq 1$ runoff triangles of observations of the same size. These runoff triangles (claims development data) have the structure shown in Figure 1. In these N runoff triangles the indices

Figure 1
Claims Development Triangle Number *n*



n, $1 \le n \le N$, refer to subportfolios (triangle)
i, $0 \le i \le I$, refer to accident years (rows)
j, $0 \le j \le J = I$, refer to development years (columns).

The cumulative claims (i.e., cumulative payments, claims incurred, or total number of reported claims) of runoff triangle n for accident year i and development year j are denoted by $C_{i,j}^{(n)}$.

Usually, at time I we have observations

$$\mathfrak{D}_I^{(n)} = \{C_{i,j}^{(n)};\ i + j \le I\} \tag{2.1}$$

for all runoff subportfolios $n \in \{1, \ldots, N\}$. This means that at time I (calendar year I) we have a total of observations over all subportfolios given by

$$\mathfrak{D}_I^N = \bigcup_{n=1}^{N} \mathfrak{D}_I^{(n)}, \tag{2.2}$$

and we need to predict the cumulative claims in its complement given by

$$\mathfrak{D}_I^{N,c} = \{C_{i,j}^{(n)};\ i + j > I,\ 1 \le n \le N\}. \tag{2.3}$$

For the derivation of an estimate of the MSEP it is convenient to write the data of the N runoff subportfolios in vector form. Thus, we define the N-dimensional random vector of cumulative claims by

$$\mathbf{C}_{i,j} = (C_{i,j}^{(1)}, \ldots, C_{i,j}^{(N)})' \tag{2.4}$$

for $i \in \{0, \ldots, I\}$ and $j \in \{0, \ldots, J\}$. Moreover, we define

$$\mathfrak{B}_k^N = \{C_{i,j}^{(n)} \in \mathfrak{D}_I^N;\ 0 \le j \le k\} \subseteq \mathfrak{D}_I^N \tag{2.5}$$

for $k \in \{0, \ldots, J\}$ and the N-dimensional column vector consisting of one's by $\mathbf{1} = (1, \ldots, 1)' \in \mathbb{R}^N$. The set $\mathfrak{B}_k^N$ consists of the cumulative claims in the N runoff triangles up to development year k, which are observed at time I. In particular, we have $\mathfrak{B}_J^N = \mathfrak{D}_I^N$.

3. MULTIVARIATE CHAIN-LADDER PREDICTION

The central objects of interest in the CL method are the development factors of the cumulative claims: we define for $n \in \{1, \ldots, N\}$, $i \in \{0, \ldots, I\}$, and $j \in \{1, \ldots, J\}$ the individual development factors for accident year i and development year j by

$$F_{i,j}^{(n)} = \frac{C_{i,j}^{(n)}}{C_{i,j-1}^{(n)}} \quad \text{and} \quad \mathbf{F}_{i,j} = (F_{i,j}^{(1)}, \ldots, F_{i,j}^{(N)})'. \tag{3.1}$$

In the following we denote by

$$\mathbf{D}(\mathbf{a}) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_N \end{pmatrix} \quad \text{and} \quad \mathbf{D}(\mathbf{a})^b = \begin{pmatrix} a_1^b & & 0 \\ & \ddots & \\ 0 & & a_N^b \end{pmatrix} \tag{3.2}$$

the $N \times N$ diagonal matrices of the N-dimensional vectors $\mathbf{a} = (a_1, \ldots, a_N)' \in \mathbb{R}^N$ and $(a_1^b, \ldots, a_N^b)' \in \mathbb{R}^N$ for an exponent $b \in \mathbb{R}$, respectively. Then we have

$$\mathbf{C}_{i,j} = \mathbf{D}(\mathbf{C}_{i,j-1}) \cdot \mathbf{F}_{i,j} = \mathbf{D}(\mathbf{F}_{i,j}) \cdot \mathbf{C}_{i,j-1} \tag{3.3}$$

for all $j = 1, \ldots, J$ and $i = 0, \ldots, I$.

The multivariate CL model is then given by the following definition.

Model Assumptions 3.1 (Multivariate CL Model)

- Cumulative claims $\mathbf{C}_{i,j}$ of different accident years i are independent.
- $(\mathbf{C}_{i,j})_{j\geq 0}$ form an N-dimensional Markov chain. There are N-dimensional deterministic vectors

$$\mathbf{f}_j = (f_j^{(1)}, \ldots, f_j^{(N)})' \tag{3.4}$$

with $f_j^{(n)} > 0$, $n = 1, \ldots, N$, and symmetric positive definite $N \times N$ matrices Σ_j for $j = 0, \ldots, J - 1$ such that for all $0 \leq i \leq I$ and $1 \leq j \leq J$ we have

$$\mathrm{E}[\mathbf{C}_{i,j}|\mathbf{C}_{i,j-1}] = \mathrm{D}(\mathbf{f}_{j-1}) \cdot \mathbf{C}_{i,j-1}, \tag{3.5}$$

$$\mathrm{Cov}(\mathbf{C}_{i,j}, \mathbf{C}_{i,j}|\mathbf{C}_{i,j-1}) = \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2} \cdot \Sigma_{j-1} \cdot \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2}. \tag{3.6}$$

□

REMARKS 3.2

- We call $(\mathbf{C}_{i,j})_{j\geq 0}$ a Markov chain to highlight that we consider a discrete-time model. In the literature the terminology ''Markov chain'' is sometimes used only for discrete-time countable state-space models (see, e.g., Rogers and Williams 1994, p. 228).
- Model 3.1 is a multivariate analogy of the univariate CL model of Mack (1993) and a special case of the multivariate CL model of Schnaus presented in Pröhl and Schmidt (2005) and Schmidt (2006) with independent accident years, deterministic CL factors $\mathbf{f}_j$, and deterministic $N \times N$ matrices Σ_j.
- The factors $\mathbf{f}_{j-1}$ are called N-dimensional CL factors, age-to-age factors, or link ratios. They indicate how $\mathbf{C}_{i,j}$ is linked to $\mathbf{C}_{i,j-1}$.
- Note that the dependence between different runoff subportfolios is defined through the correlation matrix Σ_{j-1}, which describes the conditional dependence between $C_{i,j}^{(n)}$ and $C_{i,j}^{(m)}$, given $\mathbf{C}_{i,j-1}$.

Lemma 3.3

Under Model Assumptions 3.1 we have for all $i \in \{1, \ldots, I\}$

$$\mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N] = \mathrm{E}[\mathbf{C}_{i,J}|\mathbf{C}_{i,I-i}] = \prod_{j=I-i}^{J-1} \mathrm{D}(\mathbf{f}_j) \cdot \mathbf{C}_{i,I-i}. \tag{3.7}$$

PROOF

The proof is similar to the proof in the univariate CL model (see Mack 1993). □

This result motivates a recursive algorithm for estimating the expected ultimate claim and predicting the ultimate claim, respectively, given the observation $\mathcal{D}_I^N$. If the N-dimensional CL factors $\mathbf{f}_j$ are known, the expected outstanding claims liabilities of accident year i for the N correlated runoff triangles based on $\mathcal{D}_I^N$ are estimated by

$$E[\mathbf{C}_{i,J}|\mathcal{D}_I^N] - \mathbf{C}_{i,I-i} = D(\mathbf{f}_{J-1}) \cdot \ldots \cdot D(\mathbf{f}_{I-i}) \cdot \mathbf{C}_{i,I-i} - \mathbf{C}_{i,I-i}. \tag{3.8}$$

However, in most practical applications the CL factors $\mathbf{f}_j$ are not known and need to be estimated. Pröhl and Schmidt (2005) and Schmidt (2006) propose the following multivariate age-to-age factor estimates for $\mathbf{f}_j$, $j = 0, \ldots, J-1$:

$$\begin{aligned}\widehat{\mathbf{f}}_j &= (\widehat{f}_j^{(1)}, \ldots, \widehat{f}_j^{(N)})' \\ &= \left(\sum_{i=0}^{I-j-1} D(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}D(\mathbf{C}_{i,j})^{1/2}\right)^{-1} \cdot \sum_{i=0}^{I-j-1} D(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}D(\mathbf{C}_{i,j})^{1/2} \cdot \mathbf{F}_{i,j+1}.\end{aligned} \tag{3.9}$$

Note that coordinate $\widehat{f}_j^{(n)}$ denotes the age-to-age factor estimate for development year j and runoff triangle $n \in \{1, \ldots, N\}$ based on the information $\mathcal{D}_I^N$.

Then define the following multivariate CL estimator for $E[\mathbf{C}_{i,j}|\mathcal{D}_I^N]$ by

$$\widehat{\mathbf{C}_{i,j}} = (\widehat{C_{i,j}^{(1)}}, \ldots, \widehat{C_{i,j}^{(N)}})' = \widehat{E}[\mathbf{C}_{i,j}|\mathcal{D}_I^N] = \prod_{l=I-i}^{j-1} D(\widehat{\mathbf{f}}_l) \cdot \mathbf{C}_{i,I-i} \tag{3.10}$$

for $i + j > I$.

REMARKS 3.4

- In the case $N = 1$ (i.e., only one runoff triangle) the age-to-age factor estimates (3.9) coincide with the classical age-to-age factor estimates

$$\widehat{f}_j = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \cdot F_{i,j+1}, \tag{3.11}$$

  which are used in the univariate CL method (see, e.g., Mack 1993).
- In the approaches of Braun (2004) and Merz and Wüthrich (2008) the predictors for the future cumulative claims of the N correlated runoff triangles are based on the univariate age-to-age factor estimates (3.11), which do not take into account the correlation structure between the different runoff triangles. In our matrix notation this means that the N-dimensional CL factors $\mathbf{f}_j$ are estimated by

$$\widehat{\mathbf{f}}_j^{(0)} = \left(\sum_{i=0}^{I-j-1} D(\mathbf{C}_{i,j})\right)^{-1} \cdot \sum_{i=0}^{I-j-1} D(\mathbf{C}_{i,j}) \cdot \mathbf{F}_{i,j+1}. \tag{3.12}$$

  Note that both $\widehat{\mathbf{f}}_j$ and $\widehat{\mathbf{f}}_j^{(0)}$ give conditionally unbiased estimators for the multivariate CL factors $\mathbf{f}_j$ (this is further described below). However, $\widehat{\mathbf{f}}_j$ is optimal in the sense that it has a smaller conditional expected squared loss to $\mathbf{f}_j$ (see Lemma 3.5 f below). If the $N \times N$ covariance matrix Σ_j is diagonal, then $\widehat{\mathbf{f}}_j$ and $\widehat{\mathbf{f}}_j^{(0)}$ coincide.
- If $\Sigma_0, \ldots, \Sigma_{J-2}$ are diagonal matrices, the following three predictions for the total ultimate claim coincide: (1) the sum of the univariate CL predictions on every individual runoff subportfolio, (2) the multivariate CL prediction proposed by Braun (2004) and Merz and Wüthrich (2008) based on the CL factor estimates (3.12), and (3) the multivariate CL prediction based on the multivariate CL factor estimates (3.9). However, in other cases it is more reasonable to use the multivariate CL factor estimates (3.9) (see Lemma 3.5 f below), whenever one has appropriate estimates for the covariance matrices Σ_j.
- Note that $\widehat{\mathbf{C}_{i,J}}$ is, on the one hand, used to estimate the conditional expectation $E[\mathbf{C}_{i,J}|\mathcal{D}_I^N]$ and, on the other hand, to predict the random variable $\mathbf{C}_{i,J}$.

Under Model Assumptions 3.1 we have the following properties. These are in analogy to the univariate case (see, e.g., Mack 1993; Taylor 2000) and justify the prediction of the ultimate claim $\mathbf{C}_{i,J}$ by $\widehat{\mathbf{C}_{i,J}}$:

LEMMA 3.5

Under Model Assumptions 3.1 we have

a. *Given* $\mathcal{B}_j^N$, $\widehat{\mathbf{f}}_j$ *is an unbiased estimator for* $\mathbf{f}_j$, *i.e.,* $\mathrm{E}[\widehat{\mathbf{f}}_j|\mathcal{B}_j^N] = \mathbf{f}_j$

b. $\widehat{\mathbf{f}}_j$ *is (unconditionally) unbiased for* $\mathbf{f}_j$, *i.e.,* $\mathrm{E}[\widehat{\mathbf{f}}_j] = \mathbf{f}_j$

c. $\widehat{\mathbf{f}}_j$ *and* $\widehat{\mathbf{f}}_k$ *are uncorrelated for* $j \neq k$, *i.e.,* $\mathrm{E}[\widehat{\mathbf{f}}_j \cdot \widehat{\mathbf{f}}_k'] = \mathbf{f}_j \cdot \mathbf{f}_k' = \mathrm{E}[\widehat{\mathbf{f}}_j] \cdot \mathrm{E}[\widehat{\mathbf{f}}_k]'$

d. *Given* $\mathbf{C}_{i,I-i}$, $\widehat{\mathbf{C}_{i,J}}$ *is an unbiased predictor for* $\mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N] = \mathrm{E}[\mathbf{C}_{i,J}|\mathbf{C}_{i,I-i}]$, *i.e.,* $\mathrm{E}[\widehat{\mathbf{C}_{i,J}}|\mathbf{C}_{i,I-i}] = \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]$

e. $\widehat{\mathbf{C}_{i,J}}$ *is an unbiased estimator for* $\mathrm{E}[\mathbf{C}_{i,J}]$, *i.e.,* $\mathrm{E}[\widehat{\mathbf{C}_{i,J}}] = \mathrm{E}[\mathbf{C}_{i,J}]$

f. $\widehat{\mathbf{f}}_j$ *minimizes (conditional on* $\mathcal{B}_j^N$*) the conditional expected squared loss to* $\mathbf{f}_j$ *among all unbiased linear combinations of the unbiased estimators* $(\mathbf{F}_{i,j+1})_{0\le i\le I-j-1}$ *for* $\mathbf{f}_j$, *i.e.,*

$$\mathrm{E}[(\mathbf{f}_j - \widehat{\mathbf{f}}_j)' \cdot (\mathbf{f}_j - \widehat{\mathbf{f}}_j)|\mathcal{B}_j^N] = \min_{\mathrm{W}_{i,j}\in\mathbb{R}^{N\times N}} \mathrm{E}\left[\left(\mathbf{f}_j - \sum_{i=0}^{I-j-1} \mathrm{W}_{i,j} \cdot \mathbf{F}_{i,j+1}\right)' \cdot \left(\mathbf{f}_j - \sum_{i=0}^{I-j-1} \mathrm{W}_{i,j} \cdot \mathbf{F}_{i,j+1}\right)\middle|\mathcal{B}_j^N\right]. \tag{3.13}$$

The proof of Lemma 3.5 is given in the Appendix. The lemma states that we have unbiased, optimal least squares estimators for the CL factors.

First conclusion: Observe that Lemma 3.5 d shows that we obtain an unbiased estimator $\widehat{\mathbf{C}_{i,J}}$ for the conditionally expected ultimate claim $\mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]$. This implies that the predictor for the aggregated ultimate claim of accident year $i \in \{1, \ldots, I\}$

$$\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}} = \mathbf{1}' \cdot \widehat{\mathbf{C}_{i,J}} \tag{3.14}$$

is, given $\mathbf{C}_{i,I-i}$, unbiased for $\Sigma_{n=1}^N \mathrm{E}[C_{i,J}^{(n)}|\mathbf{C}_{i,I-i}]$.

4. CONDITIONAL MSEP

In the last section we have provided the predictors for the ultimate claim given by

$$\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}} \quad \text{and} \quad \sum_{i=1}^{I}\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}}. \tag{4.1}$$

In this section we quantify the prediction uncertainty in terms of the second moment. This means our goal is to derive an estimate of the conditional MSEP for single accident years i, given $\mathcal{D}_I^N$, which is defined as

$$\begin{aligned}\mathrm{msep}_{\Sigma_n C_{i,J}^{(n)}|\mathcal{D}_I^N}\left(\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}}\right) &= \mathrm{E}\left[\left(\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}} - \sum_{n=1}^{N} C_{i,J}^{(n)}\right)^2\middle|\mathcal{D}_I^N\right] \\ &= \mathbf{1}' \cdot \mathrm{E}[(\widehat{\mathbf{C}_{i,J}} - \mathbf{C}_{i,J}) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathbf{C}_{i,J})'|\mathcal{D}_I^N] \cdot 1, \end{aligned}\tag{4.2}$$

as well as an estimate of the conditional MSEP for aggregated accident years given by

$$\mathrm{msep}_{\sum_{i,n} C_{i,J}^{(n)}|\mathcal{D}_I^N}\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{C_{i,J}^{(n)}}\right) = \mathrm{E}\left[\left(\sum_{i=1}^{I}\sum_{n=1}^{N}\widehat{C_{i,J}^{(n)}} - \sum_{i=1}^{I}\sum_{n=1}^{N}C_{i,J}^{(n)}\right)^2\middle|\mathcal{D}_I^N\right]. \tag{4.3}$$

In Subsection 4.1 we derive an estimator for the conditional MSEP for a single accident year i and in Subsection 4.2 for aggregated accident years.

4.1 Conditional MSEP for Single Accident Years

We choose $i \in \{1, \ldots, I\}$. The conditional MSEP (4.2) decouples into (a) conditional process variance and (b) conditional estimation error:

$$\mathrm{msep}_{\sum_n C_{i,J}^{(n)}|\mathcal{D}_I^N}\left(\sum_{n=1}^{N}\widehat{C_{i,J}^{(n)}}\right) = \underbrace{\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1}}_{\text{conditional process variance}}$$

$$+ \underbrace{\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}}_{\text{conditional estimation error}}. \tag{4.4}$$

The conditional process variance $\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1}$ originates from the stochastic movement of $\mathbf{C}_{i,J}$, whereas the conditional estimation error reflects the uncertainty in the prediction of the conditional expectation (mean value) $\mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]$. We derive estimates for both the conditional process variance (see Subsection 4.1.1) and the conditional estimation error (see 4.1.2) for N correlated runoff triangles. Note that for the decoupling $\widehat{\mathbf{C}_{i,J}}$ is known/observable at time I (i.e., $\widehat{\mathbf{C}_{i,J}}$ is $\mathcal{D}_I^N$-measurable).

We emphasize that we are interested into study of the conditional quantities (see also Mack 1993, Section 3). In the literature one also finds estimates for the unconditional MSEP.

4.1.1 Conditional Process Variance

If we identify the empty product by $\Pi_{l=k+1}^{k} \mathrm{D}(\mathbf{f}_l) = \mathrm{I}$, where I is the $N \times N$ identity matrix, we obtain the following result for the conditional process variance.

Lemma 4.1 (Process Variance for Single Accident Years)

Under Model Assumptions 3.1 the conditional process variance for the ultimate claim $\mathbf{C}_{i,J}$ of accident year $i \in \{1, \ldots, I\}$, given the observations $\mathcal{D}_I^N$, is given by

$$\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1} = \mathbf{1}' \cdot \left(\sum_{l=I-i}^{J-1}\prod_{k=l+1}^{J-1}\mathrm{D}(\mathbf{f}_k) \cdot \Sigma_{i,l}^{\mathbf{C}} \cdot \prod_{k=l}^{J-1}\mathrm{D}(\mathbf{f}_k)\right) \cdot \mathbf{1}, \tag{4.5}$$

where

$$\Sigma_{i,l}^{\mathbf{C}} = \mathrm{E}[\mathrm{D}(\mathbf{C}_{i,l})^{1/2} \cdot \Sigma_l \cdot \mathrm{D}(\mathbf{C}_{i,l})^{1/2}|\mathbf{C}_{i,I-i}]. \tag{4.6}$$

PROOF

We assume $i \in \{1, \ldots, I\}$. Using the Markov property and the independence of different accident years we have

$$\begin{aligned}\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1} &= \mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathbf{C}_{i,I-i}) \cdot \mathbf{1}\\ &= \mathbf{1}' \cdot \mathrm{E}[\mathrm{Var}(\mathbf{C}_{i,J}|\mathbf{C}_{i,J-1})|\mathbf{C}_{i,I-i}] \cdot \mathbf{1} + \mathbf{1}' \cdot \mathrm{Var}(\mathrm{E}[\mathbf{C}_{i,J}|\mathbf{C}_{i,J-1}]|\mathbf{C}_{i,I-i}) \cdot \mathbf{1}.\end{aligned} \tag{4.7}$$

For the second term on the right-hand side of (4.7) we obtain with (3.5)

$$\mathbf{1}' \cdot \mathrm{Var}(\mathrm{E}[\mathbf{C}_{i,J}|\mathbf{C}_{i,J-1}]|\mathbf{C}_{i,I-i}) \cdot \mathbf{1} = \mathbf{1}' \cdot \mathrm{D}(\mathbf{f}_{J-1}) \cdot \mathrm{Var}(\mathbf{C}_{i,J-1}|\mathbf{C}_{i,I-i}) \cdot \mathrm{D}(\mathbf{f}_{J-1}) \cdot \mathbf{1}. \tag{4.8}$$

Using (3.6) we obtain for the first term on the right-hand side of (4.7)

$$\mathbf{1}' \cdot \mathrm{E}[\mathrm{Var}(\mathbf{C}_{i,J}|\mathbf{C}_{i,J-1})|\mathbf{C}_{i,I-i}] \cdot \mathbf{1} = \mathbf{1}' \cdot \mathrm{E}[\mathrm{D}(\mathbf{C}_{i,J-1})^{1/2} \cdot \Sigma_{J-1} \cdot \mathrm{D}(\mathbf{C}_{i,J-1})^{1/2}|\mathbf{C}_{i,I-i}] \cdot \mathbf{1}. \quad (4.9)$$

Hence, we obtain the following recursive formula for the conditional process variance of accident year $i > I - J$

$$\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1} = \mathbf{1}' \cdot (\Sigma_{i,J-1}^{\mathbf{C}} + \mathrm{D}(\mathbf{f}_{J-1}) \cdot \mathrm{Var}(\mathbf{C}_{i,J-1}|\mathcal{D}_I^N) \cdot \mathrm{D}(\mathbf{f}_{J-1})) \cdot \mathbf{1}. \quad (4.10)$$

Iteration of this recursive formula finishes the proof of the lemma.

□

REMARKS 4.2

- Formula (4.10) gives a recursive formula for the conditional process variance of N correlated runoff triangles.
- Formula (4.5) is the conditional process variance for single accident years that was already derived in Braun (2004) and Merz and Wüthrich (2008) in matrix notation. For $N = 1$ this formula reduces to the conditional process variance for single accident years of a individual runoff triangle (cf. Mack 1993).

If we replace the parameters $\mathbf{f}_k$ and $\Sigma_{i,l-1}^{\mathbf{C}}$ in (4.5) by their estimates (cf. Section 5), we obtain an estimator of the conditional process variance for a single accident year.

4.1.2 Conditional Estimation Error

In this subsection we treat the second term

$$\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N])' \cdot \mathbf{1} \quad (4.11)$$

on the right-hand side of (4.4). This means that we want to determine the uncertainty in the estimation of $\mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]$ by the predictor $\widehat{\mathbf{C}_{i,J}}$. Using Lemma 3.3 and CL predictor (3.10) we have for the conditional estimation error of accident year $i \in \{1, \ldots, I\}$ the representation

$$\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}$$
$$= \mathbf{1}' \cdot \left(\prod_{j=I-i}^{J-1} \mathrm{D}(\widehat{\mathbf{f}}_j) - \prod_{j=I-i}^{J-1} \mathrm{D}(\mathbf{f}_j) \right) \cdot \mathbf{C}_{i,I-i} \cdot \mathbf{C}'_{i,I-i} \cdot \left(\prod_{j=I-i}^{J-1} \mathrm{D}(\widehat{\mathbf{f}}_j) - \prod_{j=I-i}^{J-1} \mathrm{D}(\mathbf{f}_j) \right) \cdot \mathbf{1}. \quad (4.12)$$

Moreover, since it holds for any N-dimensional vectors **a, b, c,** and **d,**

$$\mathrm{D}(\mathbf{b}) \cdot \mathbf{a} \cdot \mathbf{c}' \cdot \mathrm{D}(\mathbf{d}) = \mathrm{D}(\mathbf{a}) \cdot \mathbf{b} \cdot \mathbf{d}' \cdot \mathrm{D}(\mathbf{c}), \quad (4.13)$$

we obtain for the conditional estimation error

$$\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}$$
$$= \mathbf{1}' \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot (\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J}) \cdot (\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J})' \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot \mathbf{1}, \quad (4.14)$$

where the product of the multivariate CL estimators and factors $\hat{\mathbf{g}}_{i|j}$ and $\mathbf{g}_{i|j}$, respectively, are for $j = I - i + 1, \ldots, J$ defined by

$$\hat{\mathbf{g}}_{i|j} = \mathrm{D}(\widehat{\mathbf{f}}_{I-i}) \cdot \ldots \cdot \mathrm{D}(\widehat{\mathbf{f}}_{j-1}) \cdot \mathbf{1},$$
$$\mathbf{g}_{i|j} = \mathrm{D}(\mathbf{f}_{I-i}) \cdot \ldots \cdot \mathrm{D}(\mathbf{f}_{j-1}) \cdot \mathbf{1}. \quad (4.15)$$

This means that to determine the conditional estimation error we would like to calculate the right-hand side of (4.14). Observe that the realizations of the estimators $\widehat{\mathbf{f}}_{I-i}, \ldots, \widehat{\mathbf{f}}_{J-1}$ are known at time I, but the "true" CL factors $\mathbf{f}_{I-i}, \ldots, \mathbf{f}_{J-1}$ are unknown. Hence (4.14) cannot be calculated explicitly.

Therefore to determine the conditional estimation error we need to analyze how much the ''possible'' CL estimators $\widehat{\mathbf{f}}_j$ fluctuate around their ''true'' mean values $\mathbf{f}_j$. In the following we measure these volatilities of the estimators $\widehat{\mathbf{f}}_j$ around $\mathbf{f}_j$. This is done by means of ''resampled'' observations for $\widehat{\mathbf{f}}_j$. We use the terminology ''resampling'' to indicate that there are various different ways to generate new observations and to study these possible fluctuations. Here we consider conditional resampling in the multivariate CL model. We define a multivariate CL time series model, which is helpful to describe the conditional resampling approach (see Approach 3 in Buchwalder et al. 2006) to estimate the fluctuations of the estimators $\widehat{\mathbf{f}}_0, \ldots, \widehat{\mathbf{f}}_{J-1}$ around the true multivariate CL factors $\mathbf{f}_0, \ldots, \mathbf{f}_{J-1}$. This conditional resampling approach enables to give analytical estimators for the estimation error.

REMARK

There are various ways to resample these values, conditional and unconditional ones (see, e.g., Approaches 1–3 in Buchwalder et al. 2006). The question as to which approach should be chosen is not a mathematical one and has led to extensive discussions among actuaries (see Buchwalder et al. 2006; Mack, Quarg, and Braun 2006; Gisler 2006; Venter 2006; Wüthrich, Merz, and Bühlmann in press; Murphy 2007). It depends on the circumstances of the questions as to which approach should be used for a specific practical problem. However, for many practical problems the resulting numerical answers do not substantially differ (see Wüthrich, Merz, and Bühlmann in press).

Model Assumptions 4.3 (Multivariate Time Series Model)

- Cumulative claims $\mathbf{C}_{i,j}$ of different accident years i are independent.
- There exist N-dimensional constants

$$\mathbf{f}_j = (f_j^{(1)}, \ldots, f_j^{(N)})' \quad \text{and} \quad \boldsymbol{\sigma}_j = (\sigma_j^{(1)}, \ldots, \sigma_j^{(N)})'$$

with $f_j^{(n)} > 0$, $\sigma_j^{(n)} > 0$ and N-dimensional random variables

$$\boldsymbol{\varepsilon}_{i,j+1} = (\varepsilon_{i,j+1}^{(1)}, \ldots, \varepsilon_{i,j+1}^{(N)})',$$

such that for all $i \in \{0, \ldots, I\}$ and $j \in \{0, \ldots, J-1\}$ we have

$$\mathbf{C}_{i,j+1} = \mathrm{D}(\mathbf{f}_j) \cdot \mathbf{C}_{i,j} + \mathrm{D}(\mathbf{C}_{i,j})^{1/2} \cdot \mathrm{D}(\boldsymbol{\varepsilon}_{i,j+1}) \cdot \boldsymbol{\sigma}_j. \tag{4.16}$$

The random variables $\boldsymbol{\varepsilon}_{i,j+1}$ are independent with $\mathrm{E}[\boldsymbol{\varepsilon}_{i,j+1}] = \mathbf{0}$ and

$$\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}) = \mathrm{E}[\boldsymbol{\varepsilon}_{i,j+1} \cdot \boldsymbol{\varepsilon}'_{i,j+1}] = \begin{pmatrix} 1 & \rho_j^{(1,2)} & \cdots & \cdots & \rho_j^{(1,N)} \\ \rho_j^{(2,1)} & 1 & \cdots & \cdots & \rho_j^{(2,N)} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \rho_j^{(N,1)} & \rho_j^{(N,2)} & \cdots & \cdots & 1 \end{pmatrix}, \tag{4.17}$$

where $\rho_j^{(n,m)} \in (-1, 1)$ for $n, m \in \{1, \ldots, N\}$ and $n \neq m$. □

Observe, under Model Assumptions 4.3 and using (4.13), we have

$$\begin{aligned}\mathrm{Cov}(\mathbf{C}_{i,j}, \mathbf{C}_{i,j}|\mathbf{C}_{i,j-1}) &= \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2} \cdot \mathrm{E}[\mathrm{D}(\boldsymbol{\varepsilon}_{i,j}) \cdot \boldsymbol{\sigma}_{j-1} \cdot \boldsymbol{\sigma}'_{j-1} \cdot \mathrm{D}(\boldsymbol{\varepsilon}_{i,j})] \cdot \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2} \\ &= \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2} \cdot \mathrm{D}(\boldsymbol{\sigma}_{j-1}) \cdot \mathrm{E}[\boldsymbol{\varepsilon}_{i,j} \cdot \boldsymbol{\varepsilon}'_{i,j}] \cdot \mathrm{D}(\boldsymbol{\sigma}_{j-1}) \cdot \mathrm{D}(\mathbf{C}_{i,j-1})^{1/2}. \end{aligned} \tag{4.18}$$

REMARKS 4.4

- This is the multivariate CL time series model in Merz and Wüthrich (2008) in matrix notation. For the special case $N = 1$ we obtain the time series model in Buchwalder et al. (2006) for a single runoff triangle.
- The Time Series Model 4.3 defines an N-dimensional autoregressive process that is particulary useful for the derivation of the conditional estimation error because it explicitly describes the resampling mechanism (cf. [4.25] below).

- In Model Assumptions 4.3 we impose stronger conditions compared to Model Assumptions 3.1. These stronger assumptions are used to motivate the MSEP estimator. Hence, strictly speaking, the subsequent derivations hold true only in the time series model.
- It is easy to show that Time Series Model Assumptions 4.3 imply the Model Assumptions 3.1 of the multivariate CL model with

$$\Sigma_{j-1} = \mathrm{E}[\mathrm{D}(\boldsymbol{\varepsilon}_{i,j}) \cdot \boldsymbol{\sigma}_{j-1} \cdot \boldsymbol{\sigma}'_{j-1} \cdot \mathrm{D}(\boldsymbol{\varepsilon}_{i,j})] = \mathrm{D}(\boldsymbol{\sigma}_{j-1}) \cdot \mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j}, \boldsymbol{\varepsilon}_{i,j}) \cdot \mathrm{D}(\boldsymbol{\sigma}_{j-1})$$

$$= \begin{pmatrix} (\sigma_{j-1}^{(1)})^2 & \sigma_{j-1}^{(1)}\sigma_{j-1}^{(2)}\rho_{j-1}^{(1,2)} & \cdots & \sigma_{j-1}^{(1)}\sigma_{j-1}^{(N)}\rho_{j-1}^{(1,N)} \\ \sigma_{j-1}^{(2)}\sigma_{j-1}^{(1)}\rho_{j-1}^{(2,1)} & (\sigma_{j-1}^{(2)})^2 & \cdots & \sigma_{j-1}^{(2)}\sigma_{j-1}^{(N)}\rho_{j-1}^{(2,N)} \\ \vdots & \vdots & & \vdots \\ \sigma_{j-1}^{(N)}\sigma_{j-1}^{(1)}\rho_{j-1}^{(N,1)} & \sigma_{j-1}^{(N)}\sigma_{j-1}^{(2)}\rho_{j-1}^{(N,2)} & \cdots & (\sigma_{j-1}^{(N)})^2 \end{pmatrix}. \tag{4.19}$$

- Theoretically our process could have negative values for the cumulative claims $C_{i,j+1}^{(n)}$. To avoid this problem, we could reformulate the definition of $\varepsilon_{i,j+1}^{(n)}$ such that its distribution is conditionally, given $\mathcal{B}_0^N$, centered with variance 1 and such that $C_{i,j+1}^{(n)}$ is positive, $P(\cdot \mid \mathcal{B}_0^N)$-a.s. (for more details see Wüthrich, Merz, and Bühlmann in press).
- Within the CL framework Braun (2004) and Merz and Wüthrich (2008) proposed the development year–based correlations given by (4.17). Often correlations between different runoff triangles are attributed to claims inflation. Under this point of view it may seem more reasonable to allow for correlation between the cumulative claims of the same calender year (diagonals of the claims development triangles). This would introduce dependencies between accident years. However, mathematical theory at the moment is not able to treat such year-based correlations within the CL framework: that is, all calender year–based dependencies should be removed from the data before calculating the reserves with the CL method. However, after correcting the data for the calender year–based correlations, further direct and indirect sources for correlations between different runoff triangles of a portfolio exist and should be taken into account (cf. Houltram 2003). This is exactly what our model does.

We now describe the conditional resampling approach (Approach 3 in Buchwalder et al. 2006) in the multivariate setup. From the right-hand side of (4.14) we see that the main difficulty in the determination of the volatility in the estimates comes from the calculation of the term ($N \times N$-matrix)

$$\hat{\mathbf{g}}_{i|J} \cdot \hat{\mathbf{g}}'_{i|J} = \left(\prod_{j=I-i}^{J-1} \hat{f}_j^{(n)} \cdot \hat{f}_j^{(m)} \right)_{1 \le n,m \le N}. \tag{4.20}$$

Observe that the estimators for the CL factors are uncorrelated (cf. Lemma 3.5 c). However, they are not independent (see Mack, Quarg, and Braun 2006; Wüthrich, Merz, and Bühlmann in press). This means that we cannot easily calculate the expected value of (4.20) because in general

$$\mathrm{E}\left[\prod_{j=I-i}^{J-1} \hat{f}_j^{(n)} \cdot \hat{f}_j^{(m)} \right] \neq \prod_{j=I-i}^{J-1} \mathrm{E}[\hat{f}_j^{(n)} \cdot \hat{f}_j^{(m)}]. \tag{4.21}$$

Therefore we approximate the right-hand side of (4.14) with the conditional resampling approach of Buchwalder et al. (2006); that is, we calculate the values

$$\prod_{j=I-i}^{J-1} \mathrm{E}[\hat{f}_j^{(n)} \cdot \hat{f}_j^{(m)} | \mathcal{B}_j^N]. \tag{4.22}$$

Note that (4.22) is a conditional approach, whereas the left-hand side of (4.21) is unconditional. Observe that the estimates $\hat{\mathbf{f}}_j$ are functions of $(\mathbf{C}_{i,j+1})_{i=0,\dots,I-j-1}$ and $(\mathbf{C}_{i,j})_{i=0,\dots,I-j-1}$ (see (3.9)). However, in the conditional resampling approach only $(\mathbf{C}_{i,j+1})_{i=0,\dots,I-j-1}$ are resampled, whereas $(\mathbf{C}_{i,j})_{i=0,\dots,I-j-1}$ serve as fixed volume measures.

Because $(\mathbf{C}_{i,j})_{0\le j\le J}$ is a Markov chain, we can write its probability distribution (with the help of stochastic kernels K_j) as follows:

$$dP_i(\mathbf{Z}_0, \ldots, \mathbf{Z}_J) = K_0(d\mathbf{Z}_0) \cdot K_1(\mathbf{Z}_0, d\mathbf{Z}_1) \cdot K_2(\mathbf{Z}_1, d\mathbf{Z}_2) \cdot \ldots \cdot K_J(\mathbf{Z}_{J-1}, d\mathbf{Z}_J). \tag{4.23}$$

In the conditional resampling approach we always resample the next step in the time series and keep fixed the set of actual observations $(\mathbf{C}_{i,j})_{i=0,\ldots,I-j-1}$ as fixed volume measures. This means that, given $\mathcal{D}_I^N$, we consider the conditional measure on the upper triangle $(\tilde{\mathbf{C}}_{k,j})_{k,j} = (\tilde{\mathbf{C}}_{k,j})_{k=0,\ldots,I-1,j=1,\ldots,I-k}$ defined by

$$dP^*_{\mathcal{D}_I^N}((\tilde{\mathbf{C}}_{k,j})_{k,j}) = \prod_{k=0}^{I-1} K_1(\mathbf{C}_{k,0}, d\tilde{\mathbf{C}}_{k,1}) \cdot \ldots \cdot K_{I-k}(\mathbf{C}_{k,I-k-1}, d\tilde{\mathbf{C}}_{k,I-k}), \tag{4.24}$$

for the resampling. In other words this means that we generate for $i \in \{0, \ldots, I\}$ and $j \in \{0, \ldots, J-1\}$, given $\mathbf{C}_{i,j}$, a "new" observation $\tilde{\mathbf{C}}_{i,j+1}$ by resampling the next step in the time series:

$$\tilde{\mathbf{C}}_{i,j+1} = \mathrm{D}(\mathbf{f}_j) \cdot \mathbf{C}_{i,j} + \mathrm{D}(\mathbf{C}_{i,j})^{1/2} \cdot \mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j, \tag{4.25}$$

where

$$\tilde{\boldsymbol{\varepsilon}}_{i,j+1} \text{ and } \boldsymbol{\varepsilon}_{i,j+1} \text{ are independent and identically distributed copies.} \tag{4.26}$$

This means that $\mathbf{C}_{i,j}$ acts as a deterministic volume measure, and we resample successively the next observation $\tilde{\mathbf{C}}_{i,j+1} \overset{(d)}{=} \mathbf{C}_{i,j+1}$, given $\mathbf{C}_{i,j}$.

In the spirit of the conditional resampling approach this leads to the following resampled estimates of the multivariate development factors:

$$\begin{aligned}\hat{\mathbf{f}}_j &= \left(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{i,j})^{1/2}\right)^{-1} \cdot \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{i,j})^{1/2}\mathrm{D}(\mathbf{C}_{i,j})^{-1} \cdot \tilde{\mathbf{C}}_{i,j+1} \\ &= \mathbf{f}_j + \left(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{i,j})^{1/2}\right)^{-1} \cdot \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j.\end{aligned} \tag{4.27}$$

REMARKS 4.5

- In (4.27) and the following exposition we use the previous notation $\hat{\mathbf{f}}_j$ for the conditionally resampled estimates of the multivariate development factors $\mathbf{f}_j$ to avoid an overloaded notation.
- In the language of probability measures this means that the resampled estimates $\hat{\mathbf{f}}_j$ in (4.27) have distribution $P^*_{\mathcal{D}_I^N}$: The denominators are given by $\mathbf{C}_{i,j}$ (fixed volume measures), and the numerators are resampled according to $K_{j+1}(\mathbf{C}_{i,j}, d\tilde{\mathbf{C}}_{i,j+1})$.

In this way we obtain a product structure and the following lemma:

Lemma 4.6

Under Model Assumptions 4.3 and assumptions (4.25)–(4.26) *we have the following:*

a. *The estimators* $\hat{\mathbf{f}}_0, \ldots, \hat{\mathbf{f}}_{J-1}$ *are independent under the probability measure* $P^*_{\mathcal{D}_I^N}$

b. $\mathrm{E}^*_{\mathcal{D}_I^N}[\hat{\mathbf{f}}_j] = \mathbf{f}_j$ *for* $0 \le j \le J-1$ *and*

c. $\mathrm{E}^*_{\mathcal{D}_I^N}[\hat{f}_j^{(n)} \cdot \hat{f}_j^{(m)}] = f_j^{(n)} \cdot f_j^{(m)} + \sum_{i=0}^{I-j-1} \mathbf{a}^i_{n|j} \cdot \Sigma_j \cdot \mathbf{a}^{i\,\prime}_{m|j}$, *where* $\mathbf{a}^i_{n|j}$ *is the n-th row of the* $N \times N$ *matrix*

$$\mathrm{A}_j^i = \left(\sum_{k=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{k,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{k,j})^{1/2}\right)^{-1} \cdot \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}. \tag{4.28}$$

PROOF

a. Follows from (4.27) and the fact that $\tilde{\boldsymbol{\varepsilon}}_{i,j+1}$, $\tilde{\boldsymbol{\varepsilon}}_{i,k+1}$ are independent for $j \neq k$.

b. Follows from (4.27) and $\mathrm{E}^{*}_{\mathcal{D}_I^N}[\tilde{\boldsymbol{\varepsilon}}_{i,j+1}] = \mathbf{0}$.

c. We consider the $N \times N$ matrix A_j^i with rows $\mathbf{a}^i_{1|j}, \ldots, \mathbf{a}^i_{N|j}$ as in (4.28). Henceforth,

$$\widehat{\mathbf{f}}_j = \mathbf{f}_j + \sum_{i=0}^{I-j-1} \mathrm{A}_j^i \cdot \mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j \tag{4.29}$$

and

$$\widehat{f}_j^{(n)} = f_j^{(n)} + \sum_{i=0}^{I-j-1} \mathbf{a}^i_{n|j} \cdot \mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j. \tag{4.30}$$

Under measure $P^{*}_{\mathcal{D}_I^N}$ we obtain, using the independence between different accident years,

$$\begin{aligned}\mathrm{E}^{*}_{\mathcal{D}_I^N}[\widehat{f}_j^{(n)} \cdot \widehat{f}_j^{(m)}] &= f_j^{(n)} \cdot f_j^{(m)} + \mathrm{E}^{*}_{\mathcal{D}_I^N}\left[\sum_{i=0}^{I-j-1} (\mathbf{a}^i_{n|j} \cdot \mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j) \cdot (\mathbf{a}^i_{m|j} \cdot \mathrm{D}(\tilde{\boldsymbol{\varepsilon}}_{i,j+1}) \cdot \boldsymbol{\sigma}_j)'\right] \\ &= f_j^{(n)} \cdot f_j^{(m)} + \sum_{i=0}^{I-j-1} \mathbf{a}^i_{n|j} \cdot \mathrm{E}[\mathrm{D}(\boldsymbol{\varepsilon}_{i,j+1}) \cdot \boldsymbol{\sigma}_j \cdot \boldsymbol{\sigma}_j' \cdot \mathrm{D}(\boldsymbol{\varepsilon}_{i,j+1})] \cdot \mathbf{a}^i_{m|j}. \end{aligned} \tag{4.31}$$

Finally, using (4.19) completes the proof of Lemma 4.6. □

Now we are ready to derive an estimator for the conditional estimation error (the right-hand side of (4.14)). Using Lemma 4.6 a–c we obtain for the conditional estimation error (4.14) of accident year $i > I - J$ the following estimator:

$$\begin{aligned}&\mathbf{1}' \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot \mathrm{E}^{*}_{\mathcal{D}_I^N}[(\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J}) \cdot (\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J})'] \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot \mathbf{1} \\ &\quad = \mathbf{1}' \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot (\Delta_{i,J}^{(n,m)})_{1\le n,m\le N} \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot \mathbf{1}, \end{aligned} \tag{4.32}$$

where the $N \times N$ matrix $(\Delta_{i,J}^{(n,m)})_{1\le n,m\le N}$ is defined by

$$(\Delta_{i,J}^{(n,m)})_{1\le n,m\le N} = \mathrm{Var}_{P^{*}_{\mathcal{D}_I^N}}(\hat{\mathbf{g}}_{i|J}) = \mathrm{E}^{*}_{\mathcal{D}_I^N}[\hat{\mathbf{g}}_{i|J} \cdot \hat{\mathbf{g}}'_{i|J}] - \mathrm{E}^{*}_{\mathcal{D}_I^N}[\hat{\mathbf{g}}_{i|J}] \cdot \mathrm{E}^{*}_{\mathcal{D}_I^N}[\hat{\mathbf{g}}'_{i|J}]. \tag{4.33}$$

The single entries of this matrix are given by

$$\Delta_{i,J}^{(n,m)} = \prod_{l=I-i}^{J-1} \left(f_l^{(n)} \cdot f_l^{(m)} + \sum_{k=0}^{I-l-1} \mathbf{a}^k_{n|l} \cdot \Sigma_l \cdot \mathbf{a}^{k\,\prime}_{m|l} \right) - \prod_{l=I-i}^{J-1} f_l^{(n)} \cdot f_l^{(m)}. \tag{4.34}$$

REMARKS 4.7

- The components $\Delta_{i,J}^{(n,m)}$ of the $N \times N$ matrix $\mathrm{Var}P^{*}_{\mathcal{D}_I^N}(\hat{\mathbf{g}}_{i|J})$ can be rewritten in a recursive form:

$$\Delta_{i,j}^{(n,m)} = \Delta_{i,j-1}^{(n,m)} \cdot f_{j-1}^{(n)} \cdot f_{j-1}^{(m)} + \prod_{l=I-i}^{j-2} \left(f_l^{(n)} \cdot f_l^{(m)} + \sum_{k=0}^{I-l-1} \mathbf{a}^k_{n|l} \cdot \Sigma_l \cdot \mathbf{a}^{k\,\prime}_{m|l} \right) \cdot \sum_{k=0}^{I-j} \mathbf{a}^k_{n|j-1} \cdot \Sigma_j \cdot \mathbf{a}^{k\,\prime}_{m|j-1} \tag{4.35}$$

  with $\Delta_{i,I-i}^{(n,m)} = 0$ for all $i \in \{1, \ldots, I\}$ and $1 \le n, m \le N$.

- If we use the estimates $\widehat{\mathbf{f}}_j^{(0)}$ for the CL factors $\mathbf{f}_j$ instead of $\widehat{\mathbf{f}}_j$ in (3.10) and (4.4) (i.e., we disregard the correlation structure between different runoff subportfolios; cf. Remarks 3.4) term (4.34) for the conditional estimation error is given by

$$\Delta_{i,j}^{(n,m)} = \prod_{l=I-i}^{J-1}\left(f_l^{(n)}f_l^{(m)} + \frac{\rho_l^{(n,m)}\sigma_l^{(n)}\sigma_l^{(m)}}{S_l^{[I-l-1](n)}S_l^{[I-l-1](m)}}\sum_{k=0}^{I-l-1}\sqrt{C_{k,l}^{(n)}C_{k,l}^{(m)}}\right) - \prod_{l=I-i}^{J-1} f_l^{(n)}f_l^{(m)} \tag{4.36}$$

with $S_l^{[I-l-1](n)} = \sum_{i=0}^{I-l-1} C_{i,l}^{(n)}$. For a proof see Merz and Wüthrich (2008), Section 4.1.2. Note, however, that in general the estimation error for $\widehat{\mathbf{f}}_j^{(0)}$ becomes larger compared to $\widehat{\mathbf{f}}_j$. This is justified by Lemma 3.5f, which gives an optimality statement for $\widehat{\mathbf{f}}_j$.

If we replace the parameters in (4.5)–(4.6) and (4.34) by their estimates (see Section 5 for parameter estimates), we obtain the following estimator for the conditional MSEP in the multivariate CL model.

Result 4.8 (MSEP for Single Accident Years)

Under Model Assumptions 4.3 we have the following estimator for the conditional MSEP of the ultimate claim for a single accident year $i \in \{1, \ldots, I\}$:

$$\widehat{\mathrm{msep}}_{\sum_n C_{i,J}^{(n)} | \mathcal{D}_I^N}\left(\sum_{n=1}^{N}\widehat{C_{i,J}^{(n)}}\right) = \underbrace{\mathbf{1}' \cdot \left(\sum_{l=I-i+1}^{J}\prod_{k=l}^{J-1} \mathrm{D}(\widehat{\mathbf{f}}_k)\cdot\widehat{\Sigma}_{i,l-1}^{\mathrm{C}}\cdot\prod_{k=l}^{J-1}\mathrm{D}(\widehat{\mathbf{f}}_k)\right)\cdot\mathbf{1}}_{\text{estimator process variance}}$$

$$+ \underbrace{\mathbf{1}'\cdot \mathrm{D}(\mathbf{C}_{i,I-i})\cdot(\widehat{\Delta}_{i,J}^{(n,m)})_{1\le n,m\le N}\cdot\mathrm{D}(\mathbf{C}_{i,I-i})\cdot\mathbf{1}}_{\text{estimator estimation error}}, \tag{4.37}$$

with

$$\widehat{\Sigma}_{i,l-1}^{\mathrm{C}} = \mathrm{D}(\widehat{\mathbf{C}_{i,l-1}})^{1/2}\cdot\widehat{\Sigma}_{l-1}\cdot\mathrm{D}(\widehat{\mathbf{C}_{i,l-1}})^{1/2}, \tag{4.38}$$

and

$$\widehat{\Delta}_{i,J}^{(n,m)} = \prod_{l=I-i}^{J-1}\left(\widehat{f}_l^{(n)}\cdot\widehat{f}_l^{(m)} + \sum_{k=0}^{I-l-1}\widehat{\mathbf{a}}_{n|l}^{k}\cdot\widehat{\Sigma}_l\cdot(\widehat{\mathbf{a}}_{m|l}^{k})'\right) - \prod_{l=I-i}^{J-1}\widehat{\mathbf{f}}_l^{(n)}\cdot\widehat{\mathbf{f}}_l^{(m)}, \tag{4.39}$$

where $\widehat{\mathbf{a}}_{n|l}^k$ and $\widehat{\mathbf{a}}_{m|l}^k$ are the n-th and m-th row of the estimate $\widehat{\mathrm{A}}_l^k$ of the $N \times N$ matrix A_l^k (see (5.3) below), and $\widehat{\Sigma}_j$ is an estimator for Σ_j *(see (5.1) below).*

REMARKS 4.9

- For $N = 1$ estimator (4.37) reduces to the estimator of the MSEP for a single accident year and an individual runoff triangle given in Buchwalder et al. (2006).
- In practice, estimator (4.37) is used as an estimator for the conditional MSEP in both Model 4.3 and Model 3.1.
- If we replace the estimates $\widehat{\mathbf{f}}_j$ by $\widehat{\mathbf{f}}_j^{(0)}$ we need to replace (4.39) by

$$\widehat{\Delta}_{i,J}^{(n,m)} = \prod_{l=I-i}^{J-1}\left(\widehat{f}_l^{(n)}\cdot\widehat{f}_l^{(m)} + \frac{\widehat{\rho}_l^{(n,m)}\widehat{\sigma}_l^{(n)}\widehat{\sigma}_l^{(m)}}{S_l^{[I-l-1](n)}\cdot S_l^{[I-l-1](m)}}\sum_{k=0}^{I-l-1}\sqrt{C_{k,l}^{(n)}\cdot C_{k,l}^{(m)}}\right) - \prod_{l=I-i}^{J-1}\widehat{f}_l^{(n)}\cdot\widehat{f}_l^{(m)}, \tag{4.40}$$

  which is the result obtained by Merz and Wüthrich (2008), formula (3.12).

4.2 Conditional MSEP for Aggregated Accident Years

We consider two different accident years $1 \le i < k \le I$. From Model Assumptions 3.1 we know that the ultimate claims $\mathbf{C}_{i,J}$ and $\mathbf{C}_{k,J}$ are independent. However, as in the univariate case we have to be careful if we aggregate the estimators $\widehat{\mathbf{C}_{i,J}}$ and $\widehat{\mathbf{C}_{k,J}}$ because they use the same observations for estimating the CL factors $\mathbf{f}_j$ and therefore are no longer independent. We define the conditional MSEP of aggregated accident years i and k by

$$\text{msep}_{\Sigma_n C_{i,J}^{(n)} + \Sigma_n C_{k,J}^{(n)} |\mathcal{D}_I^N} \left(\sum_{n=1}^N \widehat{C_{i,J}^{(n)}} + \sum_{n=1}^N \widehat{C_{k,J}^{(n)}} \right) = \mathrm{E}\left[\left(\sum_{n=1}^N (\widehat{C_{i,J}^{(n)}} + \widehat{C_{k,J}^{(n)}}) - \sum_{n=1}^N (C_{i,J}^{(n)} + C_{k,J}^{(n)}) \right)^2 \middle| \mathcal{D}_I^N \right]. \quad (4.41)$$

As for a single accident year we have the decomposition

$$\text{msep}_{\Sigma_n C_{i,J}^{(n)} + \Sigma_n C_{k,j}^{(n)} |\mathcal{D}_I^N} \left(\sum_{n=1}^N \widehat{C_{i,J}^{(n)}} + \sum_{n=1}^N \widehat{C_{k,J}^{(n)}} \right) = \mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N) \cdot \mathbf{1}$$

$$+ \mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} + \widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} + \widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}. \quad (4.42)$$

Using the independence of different accident years, the conditional process variance can be decoupled as follows:

$$\mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N) \cdot \mathbf{1} = \mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{i,J}|\mathcal{D}_I^N) \cdot \mathbf{1} + \mathbf{1}' \cdot \mathrm{Var}(\mathbf{C}_{k,J}|\mathcal{D}_I^N) \cdot \mathbf{1}. \quad (4.43)$$

For the conditional estimation error we obtain

$$\begin{aligned}
&\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} + \widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} + \widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{i,J} + \mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1} \\
&\quad = \mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N])' \cdot \mathbf{1} \\
&\qquad + \mathbf{1}' \cdot (\widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{k,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1} \\
&\qquad + 2 \cdot \mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}. \quad (4.44)
\end{aligned}$$

Hence we have the following decomposition for the conditional MSEP of two aggregated accident years:

$$\text{msep}_{\Sigma_n C_{i,J}^{(n)} + \Sigma_n C_{k,J}^{(n)} |\mathcal{D}_I^N} \left(\sum_{n=1}^N \widehat{C_{i,J}^{(n)}} + \sum_{n=1}^N \widehat{C_{k,J}^{(n)}} \right) = \text{msep}_{\Sigma_n C_{i,J}^{(n)} |\mathcal{D}_I^N} \left(\sum_{n=1}^N \widehat{C_{i,J}^{(n)}} \right) + \text{msep}_{\Sigma_n C_{k,J}^{(n)} |\mathcal{D}_I^N} \left(\sum_{n=1}^N \widehat{C_{k,J}^{(n)}} \right)$$

$$+ 2 \cdot \mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1}. \quad (4.45)$$

This means that in addition to the conditional MSEP for single accident years we obtain twice the cross-products

$$\begin{aligned}
&\mathbf{1}' \cdot (\widehat{\mathbf{C}_{i,J}} - \mathrm{E}[\mathbf{C}_{i,J}|\mathcal{D}_I^N]) \cdot (\widehat{\mathbf{C}_{k,J}} - \mathrm{E}[\mathbf{C}_{k,J}|\mathcal{D}_I^N])' \cdot \mathbf{1} \\
&\quad = \mathbf{1}' \cdot \left(\prod_{j=I-i}^{J-1} \mathrm{D}(\hat{\mathbf{f}}_j) - \prod_{j=I-i}^{J-1} \mathrm{D}(\mathbf{f}_j) \right) \cdot \mathbf{C}_{i,I-i} \cdot \mathbf{C}_{k,I-k}' \cdot \left(\prod_{j=I-k}^{J-1} \mathrm{D}(\hat{\mathbf{f}}_j) - \prod_{j=I-k}^{J-1} \mathrm{D}(\mathbf{f}_j) \right) \cdot \mathbf{1} \\
&\quad = \mathbf{1}' \cdot \mathrm{D}(\mathbf{C}_{i,I-i}) \cdot (\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J}) \cdot (\hat{\mathbf{g}}_{k|J} - \mathbf{g}_{k|J})' \cdot \mathrm{D}(\mathbf{C}_{k,I-k}) \cdot \mathbf{1}. \quad (4.46)
\end{aligned}$$

Using the described conditional resampling approach we derive an estimator for the cross-products (4.46). As in the case of a single accident year we denote the conditional probability measure of the resampled multivariate CL estimates by $P^*_{\mathcal{D}_I^N}$. Then we can explicitly calculate the cross-products (4.46) under the measure $P^*_{\mathcal{D}_I^N}$ and obtain the estimator

$$\mathbf{1}' \cdot D(\mathbf{C}_{i,I-i}) \cdot E^{*}_{\mathcal{D}_I^N}[(\hat{\mathbf{g}}_{i|J} - \mathbf{g}_{i|J}) \cdot (\hat{\mathbf{g}}_{k|J} - \mathbf{g}_{k|J})'] \cdot D(\mathbf{C}_{k,I-k}) \cdot \mathbf{1}$$

$$= \mathbf{1}' \cdot D(\mathbf{C}_{i,I-i}) \cdot \mathrm{Cov}_{P^*_{\mathcal{D}_I^N}}(\hat{\mathbf{g}}_{i|J}, \hat{\mathbf{g}}_{k|J}) \cdot D(\mathbf{C}_{k,I-k}) \cdot \mathbf{1}$$

$$= \mathbf{1}' \cdot D(\mathbf{C}_{i,I-i}) \cdot \mathrm{Var}_{P^*_{\mathcal{D}_I^N}}(\hat{\mathbf{g}}_{i|J}) \cdot D(\mathbf{C}_{k,I-k}) \cdot \prod_{l=I-k}^{I-i-1} D(\mathbf{f}_l) \cdot \mathbf{1}$$

$$= \mathbf{1}' \cdot D(\mathbf{C}_{i,I-i}) \cdot (\Delta_{i,J}^{(n,m)})_{1 \le n,m \le N} \cdot D(\mathbf{C}_{k,I-k}) \cdot \prod_{l=I-k}^{I-i-1} D(\mathbf{f}_l) \cdot \mathbf{1}, \tag{4.47}$$

where $(\Delta_{i,J}^{(n,m)})_{1 \le n,m \le N}$ is defined in (4.33)–(4.34). Replacing the parameters with their estimators (see Section 5 for estimators) we obtain from (4.45) and (4.47) the following estimator for the conditional MSEP of aggregated years.

Result 4.10 (MSEP Aggregated Accident Years)

Under Model Assumptions 4.3 we have the following estimator for the conditional MSEP of the ultimate claim for aggregated accident years:

$$\widehat{\mathrm{msep}}_{\sum_i \sum_n C_{i,J}^{(n)} | \mathcal{D}_I^N}\left(\sum_{i=1}^{I} \sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}}\right) = \sum_{i=1}^{I} \widehat{\mathrm{msep}}_{\sum_n C_{i,J}^{(n)} | \mathcal{D}_I^N}\left(\sum_{n=1}^{N} \widehat{C_{i,J}^{(n)}}\right)$$

$$+ 2 \cdot \sum_{1 \le i < k \le I} \mathbf{1}' \cdot D(\mathbf{C}_{i,I-i}) \cdot (\widehat{\Delta}_{i,J}^{(n,m)})_{1 \le n,m \le N} \cdot D(\mathbf{C}_{k,I-k}) \cdot \prod_{l=I-k}^{I-i-1} D(\widehat{\mathbf{f}}_l) \cdot \mathbf{1}, \tag{4.48}$$

with $\widehat{\Delta}_{i,J}^{(n,m)}$ *given by* (4.39).

The same remarks apply as in Remarks 4.9.

5. PARAMETER ESTIMATION

In this section we give estimates $\widehat{\mathbf{f}}_j$, $\widehat{\boldsymbol{\sigma}}_j$, and $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ of the N-dimensional parameters $\mathbf{f}_j$, $\boldsymbol{\sigma}_j$ and of the $N \times N$–dimensional parameters $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ for $j = 0, \ldots, J-1$. Observe that the age-to-age factor estimates $\widehat{\mathbf{f}}_j$ can be calculated only when the covariance matrices Σ_j are known.

Motivated by (4.19), (4.6), and (4.28) we use these parameter estimates to define the following estimates for Σ_j, $\Sigma_{i,j}^{\mathrm{C}}$, and A_j^i, respectively:

$$\widehat{\Sigma}_j = D(\widehat{\boldsymbol{\sigma}}_j) \cdot \widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}) \cdot D(\widehat{\boldsymbol{\sigma}}_j), \tag{5.1}$$

$$\widehat{\Sigma_{i,j}^{\mathrm{C}}} = D(\widehat{\mathbf{C}_{i,j}})^{1/2} \cdot \widehat{\Sigma}_j \cdot D(\widehat{\mathbf{C}_{i,j}})^{1/2}, \tag{5.2}$$

and

$$\widehat{\mathrm{A}}_j^i = \left(\sum_{k=0}^{I-j-1} D(\mathbf{C}_{k,j})^{1/2} \cdot \widehat{\Sigma}_j^{-1} \cdot D(\mathbf{C}_{k,j})^{1/2}\right)^{-1} \cdot D(\widehat{\mathbf{C}_{i,j}})^{1/2} \cdot \widehat{\Sigma}_j^{-1}. \tag{5.3}$$

The estimation of the CL factor $\mathbf{f}_j$ is given in (3.9). However, in contrast to the CL method for a single runoff triangle this provides only an implicit expression because the multivariate CL estimators $\widehat{\mathbf{f}}_j$ depend on the parameters $\boldsymbol{\sigma}_j$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ (see (3.9) and (4.19)), which, on the other hand, are estimated by means of $\widehat{\mathbf{f}}_j$. Therefore we propose an iterative estimation procedure of these parameters.

5.1 Estimation of $\mathbf{f}_j$

As starting values we define $\widehat{\mathbf{f}}_j^{(0)}$ by (3.12). These are the "best" estimators if we neglect the covariance structure. The coordinates of $\widehat{\mathbf{f}}_j^{(0)}$ are the univariate CL estimators for the N individual runoff subportfolios. Estimator $\widehat{\mathbf{f}}_j^{(0)}$ is an unbiased optimal estimator for $\mathbf{f}_j$ if the N subportfolios are uncorrelated, but it is not optimal if the subportfolios are correlated (cf. Remarks 3.4 and Lemma 3.5 f). From $\widehat{\mathbf{f}}_j^{(0)}$ we derive estimates $\hat{\boldsymbol{\sigma}}_j^{(1)}$ and $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})^{(1)}$ of $\boldsymbol{\sigma}_j$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$, respectively (see (5.7) and (5.12)–(5.13) below). Then these estimates are used to determine $\widehat{\mathbf{f}}_j^{(1)}$: If we iterate this procedure, we obtain for $k \geq 1$

$$\widehat{\mathbf{f}}_j^{(k)} = (\widehat{f}_j^{(1)(k)}, \ldots, \widehat{f}_j^{(N)(k)})'$$
$$= \left(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2} (\widehat{\Sigma}_j^{(k)})^{-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2} \right)^{-1} \cdot \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2} (\widehat{\Sigma}_j^{(k)})^{-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2} \cdot \mathbf{F}_{i,j+1} \tag{5.4}$$

with

$$\widehat{\Sigma}_j^{(k)} = \mathrm{D}(\hat{\boldsymbol{\sigma}}_j^{(k)}) \cdot \widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})^{(k)} \cdot \mathrm{D}(\hat{\boldsymbol{\sigma}}_j^{(k)}). \tag{5.5}$$

5.2 Estimation of σ_j

The N-dimensional parameters $\boldsymbol{\sigma}_j$ are also estimated iteratively from the data. An unbiased estimator of $\boldsymbol{\sigma}_j^2 = ((\sigma_j^{(1)})^2, \ldots, (\sigma_j^{(N)})^2)'$ is given by

$$\widehat{\boldsymbol{\sigma}_j^2} = \frac{1}{I-j-1} \cdot \sum_{i=0}^{I-j-1} (\mathrm{D}(\mathbf{F}_{i,j+1}) - \mathrm{D}(\widehat{\mathbf{f}}_j^{(0)}))^2 \cdot \mathbf{C}_{i,j} \tag{5.6}$$

for $0 \leq j \leq J-1$. This estimator is used as initial value for the iteration, which is given by the estimators of $\boldsymbol{\sigma}_j$:

$$\hat{\boldsymbol{\sigma}}_j^{(k)} = \sqrt{\frac{1}{I-j-1} \cdot \sum_{i=0}^{I-j-1} (\mathrm{D}(\mathbf{F}_{i,j+1}) - \mathrm{D}(\widehat{\mathbf{f}}_j^{(k-1)}))^2 \cdot \mathbf{C}_{i,j}} \tag{5.7}$$

for $0 \leq j \leq J-1$ and $k \geq 1$. Note that $(\hat{\boldsymbol{\sigma}}_j^{(1)})^2 = \widehat{\boldsymbol{\sigma}_j^2}$.

5.3 Estimation of $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$

The $N \times N$ matrices $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ are estimated iteratively from the data too. If $\boldsymbol{\sigma}_j$ is known, then the Hadamard product (entrywise product)

$$\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}) = (\hat{\rho}_j^{(n,m)})_{1 \leq n,m \leq N} = \widehat{P_{j+1}} \odot Q_{j+1} \tag{5.8}$$

of the two matrices

$$\widehat{P_{j+1}} = \sum_{l=0}^{I-j-1} \mathrm{D}(\boldsymbol{\sigma}_j)^{-1} \mathrm{D}(\mathbf{C}_{l,j})^{1/2} (\mathbf{F}_{l,j+1} - \widehat{\mathbf{f}}_j^{(0)}) \cdot (\mathbf{F}_{l,j+1} - \widehat{\mathbf{f}}_j^{(0)})' \mathrm{D}(\mathbf{C}_{l,j})^{1/2} \mathrm{D}(\boldsymbol{\sigma}_j)^{-1} \tag{5.9}$$

and

$$Q_{j+1} = (q_{j+1}^{(n,m)})_{1 \leq n,m \leq N} = \left(\frac{1}{I-j-2+w_{j+1}^{(n,m)}} \right)_{1 \leq n,m \leq N} \tag{5.10}$$

with

$$w_{j+1}^{(n,m)} = \frac{\left(\sum_{l=0}^{I-j-1} \sqrt{C_{l,j}^{(n)}} \cdot \sqrt{C_{l,j}^{(m)}}\right)^2}{\sum_{l=0}^{I-j-1} C_{l,j}^{(n)} \cdot \sum_{l=0}^{I-j-1} C_{l,j}^{(m)}} \tag{5.11}$$

is a positive semidefinite unbiased estimator for the positive definite covariance matrix $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ for $j = 0, \ldots, J-1$ (see Appendix, Lemma 7.1). Note that this estimator applies for the last development year only if $J < I$, otherwise we need an extrapolation; see Remarks 5.1 below.

The estimator (5.8) leads to the initial value of the iteration when we replace the unknown $\boldsymbol{\sigma}_j$ with the estimator $\hat{\boldsymbol{\sigma}}_j^{(1)}$. Henceforth, the iteration for the estimation of the $N \times N$–dimensional parameter $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ is given by

$$\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})^{(k)} = (\hat{\rho}_j^{(n,m)(k)})_{1 \le n,m \le N} = \widehat{P_{j+1}}^{(k)} \odot Q_{j+1} \tag{5.12}$$

for $0 \le j \le J-1$ and $k \ge 1$, where

$$\widehat{P_{j+1}}^{(k)} = \sum_{l=0}^{I-j-1} \mathrm{D}(\hat{\boldsymbol{\sigma}}_j^{(k)})^{-1} \mathrm{D}(\mathbf{C}_{l,j})^{1/2} (\mathbf{F}_{l,j+1} - \hat{\mathbf{f}}_j^{(k-1)}) \cdot (\mathbf{F}_{l,j+1} - \hat{\mathbf{f}}_j^{(k-1)})' \mathrm{D}(\mathbf{C}_{l,j})^{1/2} \mathrm{D}(\hat{\boldsymbol{\sigma}}_j^{(k)})^{-1}. \tag{5.13}$$

REMARKS 5.1

- The initial values $\hat{\rho}_j^{(n,m)(1)}$ of $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})^{(1)}$ are the estimates used in Braun (2004) and Merz and Wüthrich (2008) to estimate the correlation coefficients $\rho_j^{(n,m)}$.
- If we have enough data (i.e., $I > J$), we are able to estimate iteratively the parameters $\boldsymbol{\sigma}_{J-1}$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,J}, \boldsymbol{\varepsilon}_{i,J})$ by (5.7) and (5.12)–(5.13), respectively. Otherwise, if $I = J$, we do not have enough data to estimate the last variance and covariance terms. In such cases one extrapolates — for $1 \le n < m \le N$ and $k \ge 1$ — the often exponentially decreasing series

$$\hat{\sigma}_0^{(n)(k)}, \ldots, \hat{\sigma}_{J-2}^{(n)(k)} \tag{5.14}$$

and

$$\left|\hat{\rho}_0^{(n,m)(k)} \cdot \hat{\sigma}_0^{(n)(k)} \cdot \hat{\sigma}_0^{(m)(k)}\right|, \ldots, \left|\hat{\rho}_{J-2}^{(n,m)(k)} \cdot \hat{\sigma}_{J-2}^{(n)(k)} \cdot \hat{\sigma}_{J-2}^{(m)(k)}\right| \tag{5.15}$$

by one additional member $\hat{\sigma}_{J-1}^{(n)(k)}$ and $\hat{\varphi}_{J-1}^{(n,m)(k)}$, respectively. From these estimates we obtain an estimate $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,J}, \boldsymbol{\varepsilon}_{i,J})^{(k)}$ for the covariance matrix $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,J}, \boldsymbol{\varepsilon}_{i,J})$ by

$$\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,J}, \boldsymbol{\varepsilon}_{i,J})^{(k)} = (\hat{\rho}_{J-1}^{(n,m)(k)})_{1 \le n,m \le N} = \left(\frac{\hat{\varphi}_{J-1}^{(n,m)(k)}}{\hat{\sigma}_{J-1}^{(n)(k)} \cdot \hat{\sigma}_{J-1}^{(m)(k)}}\right)_{1 \le n,m \le N} \tag{5.16}$$

for $1 \le n < m \le N$. However, one needs to check carefully that these etimated covariance matrices are positive definite. In the higher dimensional cases this is often nontrivial, and, in fact, many choices made are nonpositive definite, which asks for additional adjustments.

- Observe, that the $N \times N$–dimensional estimate $\widehat{P_{j+1}}^{(k)}$ is singular when $j \ge I - N + 1$ because in this case the dimension of the linear space generated by any realizations of the $(I - j)$ N-dimensional random vectors

$$\mathrm{D}(\hat{\boldsymbol{\sigma}}_j^{(k)})^{-1} \mathrm{D}(\mathbf{C}_{l,j})^{1/2} (\mathbf{F}_{l,j+1} - \hat{\mathbf{f}}_j^{(k-1)}) \quad \text{with } l \in \{0, \ldots, I-j-1\} \tag{5.17}$$

is at most $I - j \leq I - (I - N + 1) = N - 1$. Furthermore, the realizations of (5.17) may be linearly dependent for some $j < I - N + 1$, which implies that the corresponding realization of the random matrix $\widehat{P_{j+1}}^{(k)}$ is singular. In practical applications, it is thus important to verify whether the estimates $\widehat{P_{j+1}}^{(k)}$ are invertible or not and to modify those estimates (e.g., by extrapolation like in the example below) that are not invertible.

6. EXAMPLE

We use the data set from Braun (2004). We consider two portfolios A and B (i.e., $N = 2$), which contain data of general liability and auto liability business, respectively. The data sets are given in Tables 1 and 2. The positive correlation between these two lines of business is justified by the fact that both portfolios contain liability business. Hence, certain events may influence both portfolios, and we are able to learn from the observations from one portfolio for the behavior of the other portfolio.

It holds that $I = J = 13$. This means that we do not have enough data to apply the estimators (5.7) and (5.12)–(5.13) to derive estimates $\hat{\sigma}_{12}^{(1)(k)}$, $\hat{\sigma}_{12}^{(2)(k)}$, and $\hat{\rho}_{12}^{(1,2)(k)}$ for $\sigma_{12}^{(1)(k)}$, $\sigma_{12}^{(2)(k)}$, and $\rho^{(1,2)(k)}$, respectively (cf. Remarks 5.1). That is, we have only one observation $\mathbf{C}_{0,13}$, and therefore we cannot estimate the second moments described above. Therefore, in the sequel we apply the extrapolation used by Mack (1993) to obtain estimates for these last terms:

$$\hat{\sigma}_{12}^{(1)(k)} = \min\{\hat{\sigma}_{11}^{(1)(k)}, \hat{\sigma}_{10}^{(1)(k)}, (\hat{\sigma}_{11}^{(1)(k)})^2/\hat{\sigma}_{10}^{(1)(k)}\}, \tag{6.1}$$

$$\hat{\sigma}_{12}^{(2)(k)} = \min\{\hat{\sigma}_{11}^{(2)(k)}, \hat{\sigma}_{10}^{(2)(k)}, (\hat{\sigma}_{11}^{(2)(k)})^2/\hat{\sigma}_{10}^{(2)(k)}\}, \tag{6.2}$$

$$\hat{\rho}_{12}^{(1,2)(k)} = \frac{\hat{\varphi}_{12}^{(1,2)(k)}}{\hat{\sigma}_{12}^{(1)(k)} \cdot \hat{\sigma}_{12}^{(2)(k)}}, \tag{6.3}$$

with

$$\hat{\varphi}_{12}^{(1,2)(k)} = \min\left\{ |\hat{\rho}_{11}^{(1,2)(k)}\hat{\sigma}_{11}^{(1)(k)}\hat{\sigma}_{11}^{(2)(k)}|,\ |\hat{\rho}_{10}^{(1,2)(k)}\hat{\sigma}_{10}^{(1)(k)}\hat{\sigma}_{10}^{(2)(k)}|,\ \frac{|\hat{\rho}_{11}^{(1,2)(k)}\hat{\sigma}_{11}^{(1)(k)}\hat{\sigma}_{11}^{(2)(k)}|^2}{|\hat{\rho}_{10}^{(1,2)(k)}\hat{\sigma}_{10}^{(1)(k)}\hat{\sigma}_{10}^{(2)(k)}|} \right\}. \tag{6.4}$$

As (5.12)–(5.13) would lead to the estimate $\rho_{11}^{(1,2)(1)} = 1{,}001$, which is based only on two observations, we have also extrapolated $\rho_{11}^{(1,2)(k)}$ by

Table 1

Observed Cumulative Claims $C_{i,j}^{(1)}$ in Portfolio A (millions $)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	59,966	163,152	254,512	349,524	433,265	475,778	513,660	520,309	527,978	539,039	537,301	540,873	547,696	549,589
1	49,685	153,344	272,936	383,349	458,791	503,358	532,615	551,437	555,792	556,671	560,844	563,571	562,795	
2	51,914	170,048	319,204	425,029	503,999	544,769	559,475	577,425	588,342	590,985	601,296	602,710		
3	84,937	273,183	407,318	547,288	621,738	687,139	736,304	757,440	758,036	782,084	784,632			
4	98,921	278,329	448,530	561,691	641,332	721,696	742,110	752,434	768,638	768,373				
5	71,708	245,587	416,882	560,958	654,652	726,813	768,358	793,603	811,100					
6	92,350	285,507	466,214	620,030	741,226	827,979	873,526	896,728						
7	95,731	313,144	553,702	755,978	857,859	962,825	1,022,241							
8	97,518	343,218	575,441	769,017	934,103	1,019,303								
9	173,686	459,416	722,336	955,335	1,141,750									
10	139,821	436,958	809,926	1,174,196										
11	154,965	528,080	1,032,684											
12	196,124	772,971												
13	204,325													

Table 2
Observed Cumulative Claims $C_{i,j}^{(2)}$ in Portfolio B (millions $)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	114,423	247,961	312,982	344,340	371,479	371,102	380,991	385,468	385,152	392,260	391,225	391,328	391,537	391,428
1	152,296	305,175	376,613	418,299	440,308	465,623	473,584	478,427	478,314	479,907	480,755	485,138	483,974	
2	144,325	307,244	413,609	464,041	519,265	527,216	535,450	536,859	538,920	539,589	539,765	540,742		
3	145,904	307,636	387,094	433,736	463,120	478,931	482,529	488,056	485,572	486,034	485,016			
4	170,333	341,501	434,102	470,329	482,201	500,961	504,141	507,679	508,627	507,752				
5	189,643	361,123	446,857	508,083	526,562	540,118	547,641	549,605	549,693					
6	179,022	396,224	497,304	553,487	581,849	611,640	622,884	635,452						
7	205,908	416,047	520,444	565,721	600,609	630,802	648,365							
8	210,951	426,429	525,047	587,893	640,328	663,152								
9	213,426	509,222	649,433	731,692	790,901									
10	249,508	580,010	722,136	844,159										
11	258,425	686,012	915,109											
12	368,762	909,066												
13	394,997													

$$\widehat{\rho}_{11}^{(1,2)(k)} = \frac{\widehat{\varphi}_{11}^{(1,2)(k)}}{\widehat{\sigma}_{11}^{(1)(k)} \cdot \widehat{\sigma}_{11}^{(2)(k)}}, \tag{6.5}$$

with

$$\widehat{\varphi}_{11}^{(1,2)(k)} = \min\left\{ \left|\widehat{\rho}_{10}^{(1,2)(k)}\widehat{\sigma}_{10}^{(1)(k)}\widehat{\sigma}_{10}^{(2)(k)}\right|,\ \left|\widehat{\rho}_{9}^{(1,2)(k)}\widehat{\sigma}_{9}^{(1)(k)}\widehat{\sigma}_{9}^{(2)(k)}\right|,\ \frac{\left|\widehat{\rho}_{10}^{(1,2)(k)}\widehat{\sigma}_{10}^{(1)(k)}\widehat{\sigma}_{10}^{(2)(k)}\right|^2}{\left|\widehat{\rho}_{9}^{(1,2)(k)}\widehat{\sigma}_{9}^{(1)(k)}\widehat{\sigma}_{9}^{(2)(k)}\right|} \right\}. \tag{6.6}$$

Table 3 shows the parameter estimates in the first three iterations $k = 1, 2, 3$. We observe fast convergence of the two-dimensional estimates $\widehat{\mathbf{f}}_j^{(k-1)}$, $\widehat{\boldsymbol{\sigma}}_j^{(k)}$ and the one-dimensional estimates $\widehat{\rho}_j^{(1,2)(k)}$ ($k = 1, 2, 3$) in the sense that there are only small changes in the estimates after three iterations. Except for development years 6 and 10 we observe positive estimates $\widehat{\rho}_j^{(1,2)(k)}$ for the correlation coefficients. The two negative estimates should not be overstated because the estimates $\widehat{\rho}_6^{(1,2)(k)}$ and $\widehat{\rho}_{10}^{(1,2)(k)}$ are based only on seven and three observations, respectively. Observe that the estimates for the correlation coefficients are quite stable for the first six development years.

Table 3
Estimates $\widehat{\mathbf{f}}_j^{(k-1)}$, $\widehat{\boldsymbol{\sigma}}_j^{(k)}$, and $\widehat{\rho}_j^{(1,2)(k)}$ for Parameters $\mathbf{f}_j$, $\boldsymbol{\sigma}_j$, and $\rho_j^{(1,2)}$ in the First Three Iterations (k = 1, 2, 3)

Portfolio A/B	0	1	2	3	4	5	6	7	8	9	10	11	12
$\widehat{\mathbf{f}}_j^{(0)}$	3.23473	1.72048	1.35361	1.17889	1.10650	1.05466	1.02610	1.01448	1.01199	1.00619	1.00454	1.00548	1.00346
	2.22582	1.26945	1.12036	1.06676	1.03542	1.01677	1.00968	1.00006	1.00374	0.99946	1.00387	0.99891	0.99972
$\widehat{\boldsymbol{\sigma}}_j^{(1)}$	132.83	83.83	37.85	26.18	12.01	14.49	7.13	7.21	11.70	6.95	1.63	7.35	1.63
	105.38	24.64	17.94	19.07	12.50	5.55	4.52	2.13	5.14	1.40	3.21	1.37	0.58
$\widehat{\rho}_j^{(1,2)(1)}$	0.24537	0.49513	0.68236	0.44649	0.48686	0.45062	−0.17157	0.80209	0.33660	0.68744	−0.00379	0.00001	0.00000
$\widehat{\mathbf{f}}_j^{(1)}$	3.22696	1.71949	1.35247	1.17885	1.10644	1.05471	1.02612	1.01512	1.01208	1.00642	1.00454	1.00548	1.00346
	2.22236	1.26881	1.12002	1.06652	1.03563	1.01684	1.00970	1.00022	1.00383	0.99943	1.00387	0.99891	0.99972
$\widehat{\boldsymbol{\sigma}}_j^{(2)}$	132.85	83.83	37.86	26.18	12.01	14.49	7.13	7.24	11.70	6.59	1.63	7.35	1.63
	105.39	24.64	17.94	19.07	12.51	5.55	4.52	2.13	5.14	1.40	3.21	1.37	0.58
$\widehat{\rho}_j^{(1,2)(2)}$	0.24754	0.49562	0.68281	0.44666	0.48724	0.45081	−0.17176	0.80563	0.33718	0.68938	−0.00380	0.00001	0.00000
$\widehat{\mathbf{f}}_j^{(2)}$	3.22687	1.71949	1.35247	1.17885	1.10644	1.05471	1.02612	1.01514	1.01208	1.00642	1.00454	1.00548	1.00346
	2.22232	1.26881	1.12002	1.06652	1.03563	1.01684	1.00970	1.00022	1.00383	0.99943	1.00387	0.99891	0.99972
$\widehat{\boldsymbol{\sigma}}_j^{(3)}$	132.85	83.83	37.86	26.18	12.01	14.49	7.13	7.24	11.70	6.59	1.63	7.35	1.63
	105.39	24.64	17.94	19.07	12.51	5.55	4.52	2.13	5.14	1.40	3.21	1.37	0.58
$\widehat{\rho}_j^{(1,2)(3)}$	0.24757	0.49563	0.68281	0.44666	0.48724	0.45081	−0.17176	0.80573	0.33718	0.68939	−0.00380	0.00001	0.00000

Hence, using the estimators $\widehat{\mathbf{f}}_j^{(k-1)}$, $\widehat{\boldsymbol{\sigma}}_j^{(k)}$, and $\widehat{\rho}_j^{(1,2)(k)}$ ($k = 1, 2, 3$) we find an estimate (and the corresponding error) for the aggregate reserve of the portfolio consisting of the two runoff subportfolios A and B

$$\sum_{i=1}^{13} R_i = \sum_{i=1}^{13} ((C_{i,13}^{(1)} - C_{i,13-i}^{(1)}) + (C_{i,13}^{(2)} - C_{i,13-i}^{(2)})). \tag{6.7}$$

Table 4 shows the estimates for the aggregate reserve, conditional process standard deviation, (conditional estimation error)$^{1/2}$, and conditional standard error of prediction for the aggregated ultimate claim over all accident years calculated with the formulas in Buchwalder et al. (2006), Braun (2004), Merz and Wüthrich (2008), and the formulas in this paper, respectively.

We see that the univariate approaches of Braun (2004) and Merz and Wüthrich (2008) lead to the same estimates for reserves and process standard deviations as the multivariate approach with parameter estimates $\widehat{\mathbf{f}}_j^{(0)}$, $\widehat{\boldsymbol{\sigma}}_j^{(1)}$, and $\widehat{\rho}_j^{(1,2)(1)}$ (i.e., iteration $k = 1$) does. However, the results for the conditional estimation error are not the same. As expected the multivariate approach leads to a lower result for the conditional estimation error and conditional standard error of prediction than the formulas given in Braun (2004) and Merz and Wüthrich (2008) do. The use of the estimates $\widehat{\mathbf{f}}_j^{(k-1)}$, $\widehat{\boldsymbol{\sigma}}_j^{(k)}$, and $\widehat{\rho}_j^{(1,2)(k)}$ for $k = 2, 3$ (i.e., estimates of the second and third iteration, respectively) leads to a further decrease of results for the conditional estimation error and reserves. These estimates lead to a total reserve that is \$3,500 million less than the one based on $\widehat{\mathbf{f}}_j^{(0)}$.

Observe that our results for the methods of Braun (2004) and Merz and Wüthrich (2008) for the conditional process standard deviation and (conditional estimation error)$^{1/2}$ slightly differ from the numerical values given in these earlier papers because we use different estimates $\widehat{\rho}_{J-2}^{(1,2)} = 0{,}00001$ and $\widehat{\rho}_{J-1}^{(1,2)} = 0{,}00000$ for $\rho_{J-2}^{(1,2)}$ and $\rho_{J-1}^{(1,2)}$.

REMARK

In this example the use of optimal multivariate CL estimates $\widehat{\mathbf{f}}_j$ leads merely to a small effect of diversification and a slightly lower result for the conditional estimation error than the (univariate) approaches of Braun (2004) and Merz and Wüthrich (2008) do. Furthermore, the impact of the use of the optimal multivariate CL estimates $\widehat{\mathbf{f}}_j$ on the conditional estimation error is larger than the difference between the formula in Merz and Wüthrich (2008) for the conditional estimation error and its linear approximation in Braun (2004).

Table 4

Results for Individual Runoff Subportfolios A and B and Whole Runoff Portfolio for Aggregated Accident Years Derived by the Different Formulas

	Runoff Subportfolio Buchwalder et al. (2006)		Runoff Portfolio		Runoff Portfolio		
Formula	A	B	Braun (2004)	Merz and Wüthrich (2008)	Result 4.10 Iteration $k = 1$	Result 4.10 Iteration $k = 2$	Result 4.10 Iteration $k = 3$
Estimated reserves	6,155,261	2,063,612	8,218,874	8,218,874	8,218,874	8,215,227	8,215,350
Process std. deviation	330,485	134,676	396,731	396,731	396,731	396,799	396,805
$\sqrt{\text{Estimation error}}$	270,878	91,599	313,718	313,751	313,122	313,071	313,074
Prediction std. error	427,311	162,874	505,781	505,802	505,412	505,433	505,440

APPENDIX

PROOF OF LEMMA 3.5

a. Using $\mathbf{F}_{i,j} = \mathrm{D}(\mathbf{C}_{i,j-1})^{-1}\cdot \mathbf{C}_{i,j}$ and (3.5) we obtain

$$\mathrm{E}[\hat{\mathbf{f}}_j|\mathcal{B}_j^N] = \left(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{i,j})^{1/2}\right)^{-1} \cdot \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{i,j})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{i,j})^{-1/2} \cdot \mathrm{E}[\mathbf{C}_{i,j+1}|\mathcal{B}_j^N] = \mathbf{f}_j.$$

b. Follows immediately from a.

c. Using a and b we have for $j < k$

$$\mathrm{E}[\hat{\mathbf{f}}_j \cdot \hat{\mathbf{f}}_k'] = \mathrm{E}[\mathrm{E}[\hat{\mathbf{f}}_j \cdot \hat{\mathbf{f}}_k'|\mathcal{B}_k^N]] = \mathrm{E}[\hat{\mathbf{f}}_j \cdot \mathrm{E}[\hat{\mathbf{f}}_k'|\mathcal{B}_k^N]] = \mathrm{E}[\hat{\mathbf{f}}_j] \cdot \mathbf{f}_k' = \mathrm{E}[\hat{\mathbf{f}}_j] \cdot \mathrm{E}[\hat{\mathbf{f}}_k]'. \tag{A.1}$$

d. Using a we obtain

$$\begin{aligned}
\mathrm{E}[\widehat{\mathbf{C}_{i,J}}|\mathbf{C}_{i,I-i}] &= \mathrm{E}[\mathrm{D}(\hat{\mathbf{f}}_{J-1}) \cdot \mathrm{D}(\hat{\mathbf{f}}_{J-2}) \cdot \ldots \cdot \mathrm{D}(\hat{\mathbf{f}}_{I-i}) \cdot \mathbf{C}_{i,I-i}|\mathbf{C}_{i,I-i}] \\
&= \mathrm{E}[\mathrm{E}[\mathrm{D}(\hat{\mathbf{f}}_{J-1})|\mathcal{B}_{J-1}^N] \cdot \mathrm{D}(\hat{\mathbf{f}}_{J-2}) \cdot \ldots \cdot \mathrm{D}(\hat{\mathbf{f}}_{I-i}) \cdot \mathbf{C}_{i,I-i}|\mathbf{C}_{i,I-i}] \\
&= \mathrm{D}(\mathbf{f}_{J-1}) \cdot \mathrm{E}[\widehat{\mathbf{C}_{i,J-1}}|\mathbf{C}_{i,I-i}].
\end{aligned}$$

Iteration of this procedure and Lemma 3.3 leads to

$$\mathrm{E}[\widehat{\mathbf{C}_{i,J}}|\mathbf{C}_{i,I-i}] = \mathrm{D}(\mathbf{f}_{J-1}) \cdot \ldots \cdot \mathrm{D}(\mathbf{f}_{I-i}) \cdot \mathbf{C}_{i,I-i} = \mathrm{E}[\mathbf{C}_{i,J}|\mathbf{C}_{i,I-i}]. \tag{A.2}$$

e. Follows immediately from d.

f. See proof of Theorem 3.1 in Pröhl and Schmidt (2005). □

Lemma A.1

Under Model Assumptions 4.3 we have

a. *Given* $\mathcal{B}_j^N$, $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ *is an unbiased estimator for* $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$, *i.e,*

$$\mathrm{E}[\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})|\mathcal{B}_j^N] = \mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}).$$

b. $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ *is an (unconditionally) unbiased estimator for* $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$, *i.e.,*

$$\mathrm{E}[\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})] = \mathrm{Cov}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}).$$

PROOF

Here b easily follows from a. Hence we prove only a again. We consider component

$$\hat{\rho}_j^{(n,m)} = q_{j+1}^{(n,m)} \cdot \sum_{l=0}^{I-j-1} \frac{\sqrt{C_{l,j}^{(n)}} \cdot \sqrt{C_{l,j}^{(m)}}}{\sigma_j^{(n)} \cdot \sigma_j^{(m)}} \cdot \left(\frac{C_{l,j+1}^{(n)}}{C_{l,j}^{(n)}} - \hat{f}_j^{(n)(0)}\right) \cdot \left(\frac{C_{l,j+1}^{(m)}}{C_{l,j}^{(m)}} - \hat{f}_j^{(m)(0)}\right) \tag{A.3}$$

of the estimator $\widehat{\mathrm{Cov}}(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1})$ for $j \in \{0, \ldots, J-1\}$. Using the conditional unbiasedness of the individual CL factors we have

$$f_j^{(n)} = \mathrm{E}[F_{l,j+1}^{(n)}|\mathcal{B}_j^N] = \mathrm{E}\left[\left.\frac{C_{l,j+1}^{(n)}}{C_{l,j}^{(n)}}\right| \mathcal{B}_j^N\right] = \mathrm{E}[\hat{f}_j^{(n)(0)}|\mathcal{B}_j^N] \tag{A.4}$$

for $l = 0, \ldots, I - j - 1$, and $n = 1, \ldots, N$. This implies

$$E[\hat{\rho}_j^{(n,m)}|\mathcal{B}_j^N] = q_{j+1}^{(n,m)} \cdot \sum_{l=0}^{I-j-1} \frac{\sqrt{C_{l,j}^{(n)} \cdot C_{l,j}^{(m)}}}{\sigma_j^{(n)} \cdot \sigma_j^{(m)}} \cdot \mathrm{Cov}(F_{l,j+1}^{(n)} - \hat{f}_j^{(n)(0)}, F_{l,j+1}^{(m)} - \hat{f}_j^{(m)(0)}|\mathcal{B}_j^N).$$

For the covariance term on the right-hand side we have

$$\begin{aligned}
&\mathrm{Cov}(F_{l,j+1}^{(n)} - \hat{f}_j^{(n)(0)}, F_{l,j+1}^{(m)} - \hat{f}_j^{(m)(0)}|\mathcal{B}_j^N) \\
&\quad = \mathrm{Cov}(F_{l,j+1}^{(n)}, F_{l,j+1}^{(m)}|\mathcal{B}_j^N) - \mathrm{Cov}(\hat{f}_j^{(n)(0)}, F_{l,j+1}^{(m)}|\mathcal{B}_j^N) \\
&\qquad - \mathrm{Cov}(F_{l,j+1}^{(n)}, \hat{f}_j^{(m)(0)}|\mathcal{B}_j^N) + \mathrm{Cov}(\hat{f}_j^{(n)(0)}, \hat{f}_j^{(m)(0)}|\mathcal{B}_j^N) \\
&\quad = \rho_j^{(n,m)}\sigma_j^{(n)}\sigma_j^{(m)} \cdot (C_{l,j}^{(n)} \cdot C_{l,j}^{(m)})^{-1/2} \cdot \left(1 - \frac{C_{l,j}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} - \frac{C_{l,j}^{(m)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(m)}}\right) \\
&\qquad + \rho_j^{(n,m)}\sigma_j^{(n)}\sigma_j^{(m)} \cdot \frac{\sum_{i=0}^{I-j-1} \sqrt{C_{i,j}^{(n)} \cdot C_{i,j}^{(m)}}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)} \cdot \sum_{i=0}^{I-j-1} C_{i,j}^{(m)}}. \qquad \text{(A.5)}
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
E[\hat{\rho}_j^{(n,m)}|\mathcal{B}_j^N] &= q_{j+1}^{(n,m)} \cdot \sum_{l=0}^{I-j-1} \rho_j^{(n,m)} \cdot \left(1 - \frac{C_{l,j}^{(n)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)}} - \frac{C_{l,j}^{(m)}}{\sum_{i=0}^{I-j-1} C_{i,j}^{(m)}}\right) \\
&\quad + q_{j+1}^{(n,m)} \cdot \rho_j^{(n,m)} \cdot \frac{\left(\sum_{i=0}^{I-j-1} \sqrt{C_{i,j}^{(n)} \cdot C_{i,j}^{(m)}}\right)^2}{\sum_{i=0}^{I-j-1} C_{i,j}^{(n)} \cdot \sum_{i=0}^{I-j-1} C_{i,j}^{(m)}} \\
&= q_{j+1}^{(n,m)} \cdot \rho_j^{(n,m)} \cdot (I - j - 2 + w_{j+1}^{(n,m)}) \\
&= \rho_j^{(n,m)}, \qquad \text{(A.6)}
\end{aligned}$$

which completes the proof of claim a. □

REFERENCES

AJNE, B. 1994. Additivity of Chain-Ladder Projections. *ASTIN Bulletin* 24(2): 313–18.

BRAUN, C. 2004. The Prediction Error of the Chain Ladder Method Applied to Correlated Run Off Triangles. *ASTIN Bulletin* 34(2): 399–423.

BREHM, P. 2002. Correlation and the Aggregation of Unpaid Loss Distributions. *CAS Forum* (fall): 1–24.

BUCHWALDER, M., H. BÜHLMANN, M. MERZ, AND M. V. WÜTHRICH. 2006. The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited). *ASTIN Bulletin* 36(2): 521–42.

CASUALTY ACTUARIAL SOCIETY (CAS). 1990. *Foundations of Casualty Actuarial Science.* 4th ed. City: CAS.

GISLER, A. 2006. The Estimation Error in the Chain-Ladder Reserving Method: A Bayesian Approach. *ASTIN Bulletin* 36(2): 554–65.

HOULTRAM, A. 2003. Reserving Judgement: Considerations Relevant to Insurance Liability Assessment under GPS210. Institute of Actuaries of Australia, XIV General Insurance Seminar, November 9–12.

KIRSCHNER, G., C. KERLEY, AND B. ISAACS. 2002. Two Approaches to Calculating Correlated Reserve Indications across Multiple Lines of Business. *CAS Forum* (fall): 211–46.

KLEMMT, H. J. 2004. Trennung von Schadenarten und Additivität bei Chain Ladder Prognosen. Paper presented at the fall meeting of the German ASTIN Group, Munich.

MACK, T. 1993. Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin* 23(2): 213–25.

MACK, T., G. QUARG, AND C. BRAUN. 2006. The Mean Square Error of Prediction in the Chain Ladder Reserving Method—A Comment. *ASTIN Bulletin* 36(2): 543–52.

MERZ, M., AND M. V. WÜTHRICH. 2008. Prediction Error of the Chain Ladder Reserving Method Applied to Correlated Run Off Trapezoids. *Annals of Actuarial Science.* 2(1): 25–50.

MURPHY, D. M. 2007. Chain Ladder Reserve Risk Estimators. *CAS E-Forum* (summer): 1–14.

PRÖHL, C., AND K. D. SCHMIDT. 2005. Multivariate Chain-Ladder. *Dresdner Schriften zur Versicherungsmathematik.*

ROGERS, L. C. G., AND D. WILLIAMS. 1994. *Diffusions, Markov Processes and Martingales.* Vol. 1: *Foundations.* 2nd ed. Chichester: Wiley.

SCHMIDT, K. D. 2006. Optimal and Additive Loss Reserving for Dependent Lines of Business. *CAS Forum* (fall): 319–51.

TAYLOR, G. 2000. *Loss Reserving: An Actuarial Perspective.* Boston: Kluwer Academic.

TAYLOR, G., AND G. MCGUIRE. 2005. Synchronous Bootstrapping of Seemingly Unrelated Regressions. Paper presented at the 36th International ASTIN Colloquium, September 4–7, ETH, Zurich.

———. 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.

TEUGELS, J. L., AND B. SUNDT. 2004. *Encyclopedia of Actuarial Science.* Vol. 1. Chichester: Wiley.

VENTER, G. G. 2006. Discussion of Mean Square Error of Prediction in the Chain-Ladder Reserving Method. *ASTIN Bulletin* 36(2): 566–72.

WÜTHRICH, M. V., M. MERZ, AND H. BÜHLMANN. In press. Bounds on the Estimation Error in the Chain Ladder Method. *Scandinavian Actuarial Journal.*

Discussions on this paper can be submitted until October 1, 2008. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.