Uncertainty of the Claims Development Result in the Chain Ladder Method

Mario V. Wüthrich*, Michael Merz§, Natalia Lysenko*

Abstract

Using the distribution-free chain ladder method we estimate the total ultimate claim amounts at time I and after updating the information at time $I+1$. The observable claims development result at time $I+1$ for accounting year $(I, I+1]$ is then defined to be the difference between these two successive best estimate predictions for the ultimate claim. We analyze the uncertainty of this observable claims development result.

1 Introduction

In this paper we consider the problem of quantifying the uncertainty associated with the development of loss reserves for prior accident years in non-life insurance companies. For concreteness, let us assume that we are currently at time $t = I$ and we consider the next accounting year $(I, I+1]$. Then the planned profit and loss statement at time I (budget values) and the earnings statement at time $I+1$ in general look as in Table 1.

Positions a) and b) correspond to the premium income and its associated claims (generated by the premium liability). Position d) corresponds to expenses such as acquisition expenses, head office expenses, etc. Position e) corresponds to the financial returns generated on the balance sheet/assets. All these positions are typically well-understood and

*ETH Zurich, Department of Mathematics, CH-8092 Zurich, Switzerland

§University Tübingen, Faculty of Economics, D-72074 Tübingen, Germany

		budget values at Jan. 1, year I	earnings statement at Dec. 31, year I
a)	premiums earned	4'000'000	4'020'000
b)	claims incurred current accident year	-3'200'000	-3'240'000
c)	loss experience prior accident years	0	-40'000
d)	underwriting and other expenses	-1'000'000	-990'000
e)	investment income	600'000	610'000
	income before taxes	400'000	360'000

Table 1: Income statement, in Euro 1'000

risk managers, actuaries, investment officers and accountants usually believe to know how to quantify the uncertainty of these positions in the budget values at Jan. 1, year I.

Position c), "loss experience prior accident years", is, however, much less understood. It corresponds to the difference between the claims reserves at time $t = I$ and at time $t = I + 1$ adjusted for the payments during accounting year $(I, I + 1]$ for claims with accident years prior to accounting year I. To the best of our knowledge, there exists no standard terminology for this difference. In different annual reports, one may find terms such as loss experience previous years, incurred losses prior years, claims incurred prior years, run-off profit and loss, profit and loss on prior years, development of reserves for losses in prior years, etc. In our discussion, we refer to it as the *claims development result (CDR)*. Our aim in the present work is to discuss this position c), the CDR, within the framework of the distribution-free chain ladder method.

Claims reserves and uncertainties

Assuming that the assumptions of the chain ladder method are fulfilled and that claims reserves are calculated by the widely used (distribution-free) chain ladder method (see Mack [4]) we provide estimates which quantify the uncertainty of the CDR. The two components contributing to the uncertainty of the CDR are 1) a (conditional) process variance term which comes from the fact that we predict future payments (random variables); 2) a (conditional) parameter estimation error term which corresponds to the fact that the true model parameters are unknown, and are therefore estimated based on the information available at time I and time $I+1$. For example, if we consider position c) in the income statement given above, we would like to know whether the negative CDR of Euro -40'000 was only caused by an error due to the estimation of the unknown model parameters or whether there really was a shortfall in the claims development.

Short-term vs. long-term view

In the literature, it is usually the total uncertainty in the claims reserves until the claims are finally settled that has been studied. For the distribution-free chain ladder method this has first been done by Mack [4]. The study of the total uncertainty of the full run-off is a long-term view. This classical view in claims reserving is very important for solving solvency questions. In the present work we concentrate on a second important view, the short-term view. The short-term view is important for a variety of reasons:

- If the short-term behaviour is not adequate, the company may simply not get to the "long-term".
- A short-term view is relevant for management decisions, as actions need to be taken on a regular basis.
- Reflected through the annual financial statements and reports, the short-term performance of the company is of interest and importance to regulators, clients, investors, stock-markets, etc. Its consistency will ultimately have an impact on the reputation of the company in the insurance market.

Organization

The paper is organized as follows. In Section 2, we introduce our notation and methodology. We then define the true and the observable CDR for single accident years and aggregated accident years. Section 3 presents the main results of the paper, which include a measure of the variability of the CDR with estimated chain ladder factors (observable CDR) around the CDR with known chain ladder factors (true CDR) for a single accident year (Result 3.1) and over all prior accident years (Result 3.2). To illustrate the results, a numerical example with concluding remarks is provided in Section 4. The interested reader will find the proofs of the results in Section 5.

2 Methodology

2.1 Notation

Let $C_{i,j}$ denote the cumulative payments for accident year $i \in \{0, \dots, I\}$ until development year $j \in \{0, \dots, J\}$, where the accident year is referred to as the year in which an event triggering insurance claims occurs. For simplicity, we assume that $I = J$. However, all the results we present also hold for the case when the last accident year for which data is available is greater than the last development year, i.e., $I > J$.

Then the outstanding loss liabilities or future claims payments for accident year $i \in \{0, \dots, I\}$ are given by

$$R_i^I = C_{i,J} - C_{i,I-i} \tag{2.1}$$

and

$$R_i^{I+1} = C_{i,J} - C_{i,I-i+1} \tag{2.2}$$

at times $t = I$ and $t = I + 1$, respectively.

Because of the structure of the data available on insurance claims, it is customary to visualize the development of the claims using the so-called loss development trapezoids or triangles. Table 2 below gives an example of a typical loss development triangle for the case when $I = J$.

Let

$$\mathcal{D}_I = \{C_{i,j};\ i + j \leq I \text{ and } i \leq I\} \tag{2.3}$$

denote the claims data available at time $t = I$ and

$$\mathcal{D}_{I+1} = \{C_{i,j};\ i + j \leq I + 1 \text{ and } i \leq I\} = \mathcal{D}_I \cup \{C_{i,I-i+1};\ i \leq I\} \tag{2.4}$$

denote the claims data available one period later, at time $t = I + 1$.

accident year i	development year j							
	0	1	2	3	...	j	...	J
0								
1				realizations of r.v. $C_{i,j} \in \mathcal{D}_I$				
$\vdots$				(observations)				
$\vdots$								
$I-j$								
$\vdots$								
$\vdots$								
$I-2$						to be predicted r.v. $C_{i,j}$		
$I-1$								
I								

Table 2: Loss development triangle for $I = J$

If we go one step ahead in time, we obtain new observations $\{C_{i,I-i+1};\ i \leq I\}$ on the new diagonal (cf. Table 3). More formally, this means that we get an enlargement of the σ-algebra generated by the observations $\mathcal{D}_I$ to the σ-algebra generated by the observations $\mathcal{D}_{I+1}$, i.e. $\sigma(\mathcal{D}_I) \rightarrow \sigma(\mathcal{D}_{I+1})$.

accident year i	development year j				
	0	...	j	...	J
0					
$\vdots$		$\mathcal{D}_I$			
$I-j$					
$\vdots$					
I					

accident year i	development year j				
	0	...	j	...	J
0					
$\vdots$			$\mathcal{D}_{I+1}$		
$I-j$					
$\vdots$					
I					

Table 3: Loss development triangle at time $t = I$ and $t = I + 1$

2.2 Distribution-free chain ladder method

We study the claims reserving process and the CDR within the framework of the distribution-free chain ladder method. We therefore define the following time series model.

Model Assumptions 2.1

- The cumulative payments $C_{i,j}$ in different accident years $i \in \{0, \dots, I\}$ are independent.

- There exist constants $f_l > 0$, $\sigma_l > 0$ ($l = 0, \dots, J-1$) such that for all $1 \le j \le J$ and $0 \le i \le I$ we have

$$C_{i,j} = f_{j-1} \cdot C_{i,j-1} + \sigma_{j-1} \cdot \sqrt{C_{i,j-1}} \cdot \varepsilon_{i,j}, \tag{2.5}$$

  where conditionally, given $\mathcal{B}_0 = \{C_{0,0}, \dots, C_{I,0}\}$ with $C_{i,0} > 0$, $\varepsilon_{i,j}$ are independent random variables with $E\left[\varepsilon_{i,j} | \mathcal{B}_0\right] = 0$, $E\left[\varepsilon_{i,j}^2 | \mathcal{B}_0\right] = 1$ and $P\left[C_{i,j} > 0 | \mathcal{B}_0\right] = 1$ for all $i = 0, \dots, I$ and $j = 1, \dots, J$.

$\square$

Remarks 2.2

- The random variables $\varepsilon_{i,j}$ are defined conditionally, given $\mathcal{B}_0$ with $C_{i,0} > 0$, in order to ensure that all cumulative claims $C_{i,j}$ stay positive, $P\left[\cdot | \mathcal{B}_0\right]$-a.s. This is done to avoid difficulties in the time series development (2.5) and implies that all following derivations and statements are done under the conditional probability measure $P\left[\cdot | \mathcal{B}_0\right]$. In order to not overload the notation this conditional probability measure is abbreviated by P.

- Observe that Model Assumptions 2.1 define independent Markov chains $(C_{i,j})_{j \ge 0}$ for $i = 0, \dots, I$ under the probability measure P. These model assumptions satisfy the assumptions of the distribution-free chain ladder model:

$$E[C_{i,j} | C_{i,j-1}] = f_{j-1} \cdot C_{i,j-1}, \tag{2.6}$$

$$\mathrm{Var}(C_{i,j} | C_{i,j-1}) = \sigma_{j-1}^2 \cdot C_{i,j-1}. \tag{2.7}$$

  Moreover, note that we require stronger assumptions than the ones used in the original work of Mack [4]. We will need these assumptions in a similar fashion as in Buchwalder et al. [2] (see also Wüthrich-Merz-Bühlmann [6]).

Model Assumptions 2.1 imply

$$E\left[C_{i,J}\middle|\,\mathcal{D}_I\right] = C_{i,I-i} \cdot \prod_{j=I-i}^{J-1} f_j \quad \text{and} \quad E\left[C_{i,J}\middle|\,\mathcal{D}_{I+1}\right] = C_{i,I-i+1} \cdot \prod_{j=I-i+1}^{J-1} f_j. \tag{2.8}$$

This means that once the chain ladder factors f_j are known we are able to calculate the conditionally expected ultimate claim $C_{i,J}$ given the information $\mathcal{D}_I$ and $\mathcal{D}_{I+1}$, respectively.

Of course, in general, the chain ladder factors f_j are not known and need to be estimated. Within the framework of the chain ladder method this is done as follows:

1. At time $t = I$, given information $\mathcal{D}_I$, the chain ladder factors f_j are estimated by

$$\widehat{f}_j^I = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{S_j^I}, \qquad \text{where} \quad S_j^I = \sum_{i=0}^{I-j-1} C_{i,j}. \tag{2.9}$$

2. At time $t = I+1$, given information $\mathcal{D}_{I+1}$, the chain ladder factors f_j are estimated by

$$\widehat{f}_j^{I+1} = \frac{\sum_{i=0}^{I-j} C_{i,j+1}}{S_j^{I+1}}, \qquad \text{where} \quad S_j^{I+1} = \sum_{i=0}^{I-j} C_{i,j}. \tag{2.10}$$

This means the chain ladder estimates $\widehat{f}_j^{I+1}$ at time $I+1$ use the increase in information about the claims development process in the new observed accounting year $(I, I+1]$ and are therefore based on the new observation $C_{I-j,j+1}$.

Mack [4] proved that these are unbiased estimators for f_j and, moreover, that $\widehat{f}_j^m$ and $\widehat{f}_l^m$ ($m = I$ or $I+1$) are uncorrelated random variables for $j \neq l$ (see Theorem 2 in Mack [4]). This immediately implies that, given $C_{i,I-i}$,

$$\widehat{C}_{i,j}^I = C_{i,I-i} \cdot \widehat{f}_{I-i}^I \cdot \ldots \cdot \widehat{f}_{j-2}^I \cdot \widehat{f}_{j-1}^I \tag{2.11}$$

is an unbiased estimator for $E\left[C_{i,j}\middle|\,\mathcal{D}_I\right]$ with $j \geq I-i$ and, given $C_{i,I-i+1}$,

$$\widehat{C}_{i,j}^{I+1} = C_{i,I-i+1} \cdot \widehat{f}_{I-i+1}^{I+1} \cdot \ldots \cdot \widehat{f}_{j-2}^{I+1} \cdot \widehat{f}_{j-1}^{I+1} \tag{2.12}$$

is an unbiased estimator for $E\left[C_{i,j}\middle|\,\mathcal{D}_{I+1}\right]$ with $j \geq I-i+1$.

Remarks 2.3

- The realizations of the estimators $\widehat{f}_0^I, \dots, \widehat{f}_{J-1}^I$ are known at time $t = I$, but the realizations of $\widehat{f}_0^{I+1}, \dots, \widehat{f}_{J-1}^{I+1}$ are unknown since $C_{I,1}, \dots, C_{I-J+1,J}$ are unknown.

- When indices of accident and development years are such that there are no factor products in (2.11) or (2.12), an empty product is replaced by 1. For example, $\widehat{C}_{i,I-i}^I = C_{i,I-i}$ and $\widehat{C}_{i,I-i+1}^{I+1} = C_{i,I-i+1}$.

- The estimators $\widehat{C}_{i,j}^{I+1}$ are based on the chain ladder estimators at time $I+1$ and therefore use the increase in information given by the new observed accounting year from I and $I+1$.

2.3 Claims development result (CDR)

We ignore any prudential margin and assume that loss reserves are set equal to the expected outstanding liabilities conditional on the available information. Then the claims development result (CDR) for the accounting year $(I, I+1]$ is defined as follows.

Definition 2.4 (True CDR for a single accident year)
The **true CDR** *for accident year* $i \in \{1, \dots, I\}$ *in accounting year* $(I, I+1]$ *is given by*

$$\mathrm{CDR}_i(I+1) = E\left[\left. R_i^I \right| \mathcal{D}_I\right] - \left(X_{i,I-i+1} + E\left[\left. R_i^{I+1} \right| \mathcal{D}_{I+1}\right]\right), \tag{2.13}$$

where $X_{i,I-i+1} = C_{i,I-i+1} - C_{i,I-i}$ *denotes the incremental payments. Furthermore, the true aggregate* CDR *is given by*

$$\sum_{i=1}^{I} \mathrm{CDR}_i(I+1). \tag{2.14}$$

This means that $E\left[\left. R_i^I \right| \mathcal{D}_I\right]$ is used to predict R_i^I at time I, and $E\left[\left. R_i^{I+1} \right| \mathcal{D}_{I+1}\right]$ is used to predict R_i^{I+1} at time $I+1$
Observe that the true CDR (2.13) is a $\mathcal{D}_{I+1}$-measurable random variable which can be rewritten as

$$\mathrm{CDR}_i(I+1) \quad = \quad E\left[\left. C_{i,J} \right| \mathcal{D}_I\right] - E\left[\left. C_{i,J} \right| \mathcal{D}_{I+1}\right]. \tag{2.15}$$

Using the martingale property of $\left(E\left[C_{i,J}\right|\mathcal{D}_t]\right)_{t\in\mathbb{N}_0}$ it is easy to see that

$$E\left[\mathrm{CDR}_i(I+1)\right|\mathcal{D}_I] = 0. \tag{2.16}$$

This means that for known chain ladder factors f_j the expected CDR (viewed from time I) is equal to zero in the chain ladder method. Hence, for known chain ladder factors we refer to $\mathrm{CDR}_i(I+1)$ as the *true* CDR. This justifies the fact that in the budget values of the income statement the position c) is predicted by Euro 0 (see position c) in Table 1).

In general the chain ladder factors f_j are not known (i.e. the true CDR is not observable) and therefore estimated by $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$, respectively. Hence, replacing the expected ultimate claims $E\left[C_{i,J}\right|\mathcal{D}_I]$ and $E\left[C_{i,J}\right|\mathcal{D}_{I+1}]$ with their estimates $\widehat{C}_{i,J}^I$ and $\widehat{C}_{i,J}^{I+1}$, respectively, the true CDR for accident year i $(1 \le i \le I)$ in accounting year $(I, I+1]$ is estimated in the chain ladder method by:

Estimator 2.5 (Observable CDR, estimator for true CDR)
The **observable CDR** *for accident year* $i \in \{1,\ldots,I\}$ *in accounting year* $(I, I+1]$ *in the chain ladder method is given by*

$$\widehat{\mathrm{CDR}}_i(I+1) = \widehat{R}_i^{\mathcal{D}_I} - \left(X_{i,I-i+1} + \widehat{R}_i^{\mathcal{D}_{I+1}}\right) = \widehat{C}_{i,J}^I - \widehat{C}_{i,J}^{I+1}, \tag{2.17}$$

where $\widehat{R}_i^{\mathcal{D}_I}$ *and* $\widehat{R}_i^{\mathcal{D}_{I+1}}$ *are defined by* (2.19) *and* (2.20), *respectively. Furthermore, the observable aggregate* CDR *is given by*

$$\sum_{i=1}^{I} \widehat{\mathrm{CDR}}_i(I+1). \tag{2.18}$$

Note that under the Model Assumptions 2.1, given $C_{i,I-i}$,

$$\widehat{R}_i^{\mathcal{D}_I} = \widehat{C}_{i,J}^I - C_{i,I-i} \qquad (1 \le i \le I), \tag{2.19}$$

is an unbiased estimator for $E\left[R_i^I\right|\mathcal{D}_I]$ and, given $C_{i,I-i+1}$,

$$\widehat{R}_i^{\mathcal{D}_{I+1}} = \widehat{C}_{i,J}^{I+1} - C_{i,I-i+1} \qquad (1 \le i \le I), \tag{2.20}$$

is an unbiased estimator for $E\left[R_i^{I+1}\right|\mathcal{D}_{I+1}]$.

Remarks 2.6

- We should point out that, although we refer to $\widehat{\mathrm{CDR}}_i(I+1)$ as being a predictor of $\mathrm{CDR}_i(I+1)$, it is the predictor only in a sense that the expected values of the outstanding loss liabilities at time I and $I+1$ are replaced by their chain ladder estimators $\widehat{R}_i^{\mathcal{D}_I}$ and $\widehat{R}_i^{\mathcal{D}_{I+1}}$. However, given claims data up to time $t=I$, we are not able to evaluate $\widehat{\mathrm{CDR}}_i(I+1)$; its realization is known only at the end of accounting year $(I, I+1]$ when $\mathcal{D}_{I+1}$ becomes available. In other words the unobservable true CDR is replaced by an observable CDR.

- At time $I+1$, $\widehat{\mathrm{CDR}}_i(I+1)$ refers to the observed Euro -40'000 in position c) in Table 1. There remains the two questions 1) whether the observable CDR at time $I+1$ differs significantly from expected CDR at time I and 2) whether we could have a positive CDR at time $I+1$ if we knew the true chain ladder factors f_j. These two questions are analyzed in the next sections.

3 MSE of the claims development result estimator

Our ultimate goal is to quantify the degree of variability of the observable aggregate CDR at time $I+1$ for all the previous accident years (see (2.18)) around the true aggregate CDR at time $I+1$ (see (2.14)).

We first consider the (conditional) variance of the true aggregate CDR. Using the independence between different accident years and (5.5) below, it is given by

$$\mathrm{Var}\left(\sum_{i=1}^{I} \mathrm{CDR}_i(I+1) \,\middle|\, \mathcal{D}_I\right) = \sum_{i=1}^{I} E\left[C_{i,J}\middle|\,\mathcal{D}_I\right]^2 \cdot \frac{\sigma_{I-i}^2/f_{I-i}^2}{C_{i,I-i}}. \tag{3.1}$$

At time $t=I$ this can be estimated by

$$\widehat{\mathrm{Var}}\left(\sum_{i=1}^{I} \mathrm{CDR}_i(I+1) \,\middle|\, \mathcal{D}_I\right) = \sum_{i=1}^{I} \left(\widehat{C}_{i,J}^I\right)^2 \cdot \frac{\left(\widehat{\sigma}_{I-i}^I\right)^2 / \left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}}, \tag{3.2}$$

where an unbiased estimator at time $t=I$ for the variances σ_{j-1}^2, $j \geq 1$, is given below in (3.11). This gives an estimate for the volatility of the true aggregate CDR (with the true chain ladder factors f_j). The estimator (3.2) can also be seen as an estimator of the pure (conditional) process variance of our stochastic process.

However, since the true chain ladder factors f_j are not known, we will not be able to observe the true aggregate CDR, we only have the observable aggregate CDR (with f_j estimated by their best estimators $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ at time I and $I+1$, respectively).

We measure the variability between true aggregate CDR and observable aggregate CDR by considering the conditional mean square error (MSE) given by

$$\mathrm{MSE}_{\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\right) = E\left[\left(\sum_{i=1}^{I}\mathrm{CDR}_i(I+1) - \sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\right)^2 \middle| \mathcal{D}_I\right]. \tag{3.3}$$

Definition (3.3) represents the "distance" between the true aggregate CDR (for known chain ladder factors f_j), $\sum_{i=1}^{I}\mathrm{CDR}_i(I+1)$, and its estimator, the observable aggregate CDR, $\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)$, given the information $\mathcal{D}_I$ (in the L^2-sense).

To tackle this problem, we first consider the conditional mean square error of the CDR for a single accident year.

3.1 Single accident year

For a single accident year $i \in \{1, \dots, I\}$, the conditional mean square error, given information $\mathcal{D}_I$, is defined as

$$\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right) = E\left[\left(\mathrm{CDR}_i(I+1) - \widehat{\mathrm{CDR}}_i(I+1)\right)^2 \middle| \mathcal{D}_I\right]. \tag{3.4}$$

Using (2.16), we obtain the following decomposition:

$$\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right) = \Phi_{i,J}^I + \left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle| \mathcal{D}_I\right]\right)^2, \tag{3.5}$$

where

$$\begin{aligned}\Phi_{i,J}^I = \mathrm{Var}\left(\mathrm{CDR}_i(I+1) - \widehat{\mathrm{CDR}}_i(I+1)\middle| \mathcal{D}_I\right) &= \mathrm{Var}\left(\mathrm{CDR}_i(I+1)\middle| \mathcal{D}_I\right) \\ &+ \mathrm{Var}\left(\widehat{\mathrm{CDR}}_i(I+1)\middle| \mathcal{D}_I\right) - 2\cdot \mathrm{Cov}\left(\widehat{\mathrm{CDR}}_i(I+1), \mathrm{CDR}_i(I+1)\middle| \mathcal{D}_I\right).\end{aligned} \tag{3.6}$$

The term $\Phi_{i,J}^I$ corresponds to the uncertainty implied by the stochastic process. It contains three ingredients: 1) the term $\mathrm{Var}\left(\mathrm{CDR}_i(I+1)\middle| \mathcal{D}_I\right)$ is the (conditional) process variance of the true CDR, 2) the term $\mathrm{Var}\left(\widehat{\mathrm{CDR}}_i(I+1)\middle| \mathcal{D}_I\right)$ can be interpreted as the (conditional) process variance of the observable CDR (due to the fact that the observable CDR at time $I+1$ is a random variable at time $t=I$), and 3) the covariance term tells us how the true CDR, $\mathrm{CDR}_i(I+1)$, is correlated with the observable CDR, $\widehat{\mathrm{CDR}}_i(I+1)$. The second term in (3.5),

$$\left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right)^2, \tag{3.7}$$

is related to the bias of the observable CDR as estimator of the expected CDR (2.16), which arises due to the fact that we need to estimate the true chain ladder factors.

The following result provides a formula for estimating the conditional mean square error (3.4) (for details of its derivation, see Section 5, formulas (5.12) and (5.14) below).

Result 3.1 (Conditional MSE estimator for a single accident year) *We estimate the conditional MSE of the CDR estimator at time $t=I+1$ for accounting year $(I,I+1]$ and any single accident year $i\in\{1,\ldots,I\}$ by*

$$\widehat{\mathrm{MSE}}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right)=\widehat{\Phi}^I_{i,J}+\left(\widehat{C}^I_{i,J}\right)^2\cdot\widehat{\Delta}^I_{i,J}, \tag{3.8}$$

where $\widehat{\Phi}^I_{1,J}=0$ and for $i>1$

$$\widehat{\Phi}^I_{i,J}=\left(\widehat{C}^I_{i,J}\right)^2\cdot\left[1+\frac{\left(\widehat{\sigma}^I_{I-i}\right)^2/\left(\widehat{f}^I_{I-i}\right)^2}{C_{i,I-i}}\right]\cdot\left(\prod_{l=I-i+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}^I_l\right)^2/\left(\widehat{f}^I_l\right)^2}{\left(S^{I+1}_l\right)^2}\cdot C_{I-l,l}\right)-1\right), \tag{3.9}$$

$$\widehat{\Delta}^I_{i,J}=\frac{\left(\widehat{\sigma}^I_{I-i}\right)^2/\left(\widehat{f}^I_{I-i}\right)^2}{S^I_{I-i}}+\sum_{j=I-i+1}^{J-1}\left(\frac{C_{I-j,j}}{S^{I+1}_j}\right)^2\cdot\frac{\left(\widehat{\sigma}^I_j\right)^2/\left(\widehat{f}^I_j\right)^2}{S^I_j} \tag{3.10}$$

and

$$\left(\widehat{\sigma}^I_{j-1}\right)^2=\frac{1}{I-j}\cdot\sum_{i=0}^{I-j}C_{i,j-1}\cdot\left(\frac{C_{i,j}}{C_{i,j-1}}-\widehat{f}^I_{j-1}\right)^2\quad \textit{for } j=1,\ldots,J. \tag{3.11}$$

The term $\widehat{\Phi}^I_{i,J}$ estimates the variance term $\Phi^I_{i,J}$ (given in (3.6)). The term $\left(\widehat{C}^I_{i,J}\right)^2\cdot\widehat{\Delta}^I_{i,J}$ is an estimate for the bias term (3.7). Observe that $\widehat{\Phi}^I_{i,J}>0$ if $i>1$; this can easily be seen using a linear approximation (from below) for the last factor in (3.9) (see also Buchwalder et al. [2], formula (4.29)).

3.2 Aggregation over prior accident years

When aggregating over prior accident years, one has to take into account the correlations between different accident years, since the same observations are used to estimate the chain ladder factors. Based on the definition of the conditional mean square error for the aggregate CDR estimator stated in (3.3), the following result summarizes how one can do the estimation under the chain ladder framework (refer to Section 5, formulas (5.26) and (5.27) below for the derivation).

Result 3.2 (Conditional MSE estimator for aggregated accident years) *We estimate the conditional MSE of the aggregate CDR estimator at time $t = I+1$ for accounting year $(I, I+1]$ by*

$$\widehat{\mathrm{MSE}}_{\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\right) \tag{3.12}$$

$$= \sum_{i=1}^{I}\widehat{\mathrm{MSE}}\left(\widehat{\mathrm{CDR}}_i(I+1)\right) + 2\cdot\sum_{i>k>0}\left(\widehat{\Psi}_{i,k}^{I} + \widehat{C}_{i,J}^{I}\cdot\widehat{C}_{k,J}^{I}\cdot\widehat{\Delta}_{k,J}^{I}\right),$$

where for $i > k > 1$

$$\widehat{\Psi}_{i,k}^{I} = \frac{\widehat{C}_{i,J}^{I}}{\widehat{C}_{k,J}^{I}}\cdot\left(1+\frac{\left(\widehat{\sigma}_{I-k}^{I}\right)^2/\left(\widehat{f}_{I-k}^{I}\right)^2}{S_{I-k}^{I+1}}\right)\cdot\left(1+\frac{\left(\widehat{\sigma}_{I-k}^{I}\right)^2/\left(\widehat{f}_{I-k}^{I}\right)^2}{C_{k,I-k}}\right)^{-1}\cdot\widehat{\Phi}_{k,J}^{I} \tag{3.13}$$

and $\widehat{\Psi}_{i,1}^{I} = 0$ for $i > 1$.

3.3 Variance of the observable CDR

In this subsection we provide an estimator for the variance of the observable aggregate CDR. Notice that, for known chain ladder factors f_j, the variance of the true aggregate CDR is given by (3.1). For estimated chain ladder factors, the uncertainty in position c) in the income statement (see Table 1) has two ingredients, namely a possible bias which is given by

$$(\text{u-bias})^2 = \widehat{E}_{\mathcal{D}_I}\left[\left(E\left[\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\,\middle|\,\mathcal{D}_I\right]\right)^2\right]$$
$$= \sum_{i=1}^{I}\left(\widehat{C}_{i,J}^{I}\right)^2\cdot\widehat{\Delta}_{i,J}^{I} + 2\cdot\sum_{i>k>0}\widehat{C}_{i,J}^{I}\cdot\widehat{C}_{k,J}^{I}\cdot\widehat{\Delta}_{k,J}^{I}, \tag{3.14}$$

(cf. (5.14) and (5.27) below) and the variance of the observable aggregate CDR, which is given in the next result (for its derivation see (5.29) and (5.30) below).

That is the uncertainty of prediction 0 in CDR position of the budget values (cf. position c) in Figure 1) is estimated by

$$E\left[\left(\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)-0\right)^2\,\middle|\,\mathcal{D}_I\right]=(\text{u-bias})^2+\widehat{\mathrm{Var}}\left(\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\,\middle|\,\mathcal{D}_I\right),\quad(3.15)$$

where the terms on the right-hand side of (3.15) are given by (3.14) and (3.16).

Result 3.3 (Conditional variance estimator for observable aggregate CDR) *We estimate the variance of the observable aggregate CDR for accounting year $(I, I+1]$ by*

$$\widehat{\mathrm{Var}}\left(\sum_{i=1}^{I}\widehat{\mathrm{CDR}}_i(I+1)\,\middle|\,\mathcal{D}_I\right)=\sum_{i=1}^{I}\widehat{\Gamma}_{i,J}^I+2\cdot\sum_{i>k>0}\widehat{\Upsilon}_{i,k}^I,\quad(3.16)$$

where for $i\geq 1$

$$\widehat{\Gamma}_{i,J}^I=\widehat{\mathrm{Var}}\left(\widehat{\mathrm{CDR}}_i(I+1)\,\middle|\,\mathcal{D}_I\right)$$
$$=\left(\widehat{C}_{i,J}^I\right)^2\cdot\left\{\left(\left[1+\frac{\left(\widehat{\sigma}_{I-i}^I\right)^2/\left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}}\right]\cdot\prod_{l=I-i+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}_{l}^I\right)^2/\left(\widehat{f}_{l}^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)\right)-1\right\},\quad(3.17)$$

and for $i>k>0$

$$\widehat{\Upsilon}_{i,k}^I=\widehat{\mathrm{Cov}}\left(\widehat{\mathrm{CDR}}_i(I+1),\widehat{\mathrm{CDR}}_k(I+1)\,\middle|\,\mathcal{D}_I\right)$$
$$=\widehat{C}_{i,J}^I\cdot\widehat{C}_{k,J}^I\cdot\left\{\left(\left[1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{S_{I-k}^{I+1}}\right]\prod_{l=I-k+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}_{l}^I\right)^2/\left(\widehat{f}_{l}^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)\right)-1\right\}.\quad(3.18)$$

Observe that

$$\widehat{\Gamma}_{i,J}^I=\widehat{\mathrm{Var}}\left(\widehat{\mathrm{CDR}}_i(I+1)\,\middle|\,\mathcal{D}_I\right)\geq\widehat{\mathrm{Var}}\left(\mathrm{CDR}_i(I+1)\,\middle|\,\mathcal{D}_I\right)\quad(3.19)$$

(cf. (3.2) and (3.17)).

4 Numerical example and conclusions

For our numerical example we use the dataset given in Table 4. The table contains cumulative payments for accident years $i \in \{0, 1, \dots, 8\}$ at time $I = 8$ and time $I + 1 = 9$.

	$j =$ 0	1	2	3	4	5	6	7	8
$i = 0$	2'202'584	3'210'449	3'468'122	3'545'070	3'621'627	3'644'636	3'669'012	3'674'511	3'678'633
$i = 1$	2'350'650	3'553'023	3'783'846	3'840'067	3'865'187	3'878'744	3'898'281	3'902'425	3'906'738
$i = 2$	2'321'885	3'424'190	3'700'876	3'798'198	3'854'755	3'878'993	3'898'825	3'902'130	
$i = 3$	2'171'487	3'165'274	3'395'841	3'466'453	3'515'703	3'548'422	3'564'470		
$i = 4$	2'140'328	3'157'079	3'399'262	3'500'520	3'585'812	3'624'784			
$i = 5$	2'290'664	3'338'197	3'550'332	3'641'036	3'679'909				
$i = 6$	2'148'216	3'219'775	3'428'335	3'511'860					
$i = 7$	2'143'728	3'158'581	3'376'375						
$i = 8$	2'144'738	3'218'196							
$\widehat{f}_j^I$	1.4759	1.0719	1.0232	1.0161	1.0063	1.0056	1.0013	1.0011	
$\widehat{f}_j^{I+1}$	1.4786	1.0715	1.0233	1.0152	1.0072	1.0053	1.0011	1.0011	
$\left(\widehat{\sigma}_j^I\right)^2$	911.43	189.82	97.81	178.75	20.64	3.23	0.36	0.04	

Table 4: Run-off triangle (cumulative payments, in Euro 1'000) for time $I = 8$ and $I = 9$

Table 4 summarizes the chain ladder estimates $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ of the age-to-age factors f_j as well as the variance estimates $\left(\widehat{\sigma}_j^I\right)^2$ for $j = 0, \dots, 7$. Since we do not have enough data to estimate σ_7^2 (recall $I = J$) we use the approximation given in Mack [4]

$$\left(\widehat{\sigma}_7^I\right)^2 = \min\left\{\left(\widehat{\sigma}_6^I\right)^2, \left(\widehat{\sigma}_5^I\right)^2, \left(\widehat{\sigma}_6^I\right)^4 / \left(\widehat{\sigma}_5^I\right)^2\right\}. \tag{4.1}$$

Using the estimates $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ we calculate the reserves $\widehat{R}_i^{\mathcal{D}_I}$ for the outstanding loss liabilities R_i^I at time $t = I$ and $X_{i,I-i+1} + \widehat{R}_i^{\mathcal{D}_{I+1}}$ for $X_{i,I-i+1} + R_i^I$ at time $t = I + 1$, respectively. This then gives realizations of the observable CDR for single accident years and for aggregated accident years (see Table 5). Observe that we have a negative observable aggregate CDR at time $I + 1$ of about Euro $-40'000$ (which corresponds to position c) in Table 1).

The question which we now have is whether the true aggregate CDR could also be positive if we had the true chain ladder factors f_j at time $t = I$. We therefore perform the variance and MSE analysis using the results of Section 3. Table 6 provides the estimates – for single and aggregated accident years.

	$\widehat{R}_i^{\mathcal{D}_I}$	$X_{i,I-i+1} + \widehat{R}_i^{\mathcal{D}_{I+1}}$	$\widehat{\mathrm{CDR}}_i(I+1)$
0	0	0	0
1	4'378	4'313	65
2	9'348	7'649	1'698
3	28'392	24'046	4'347
4	51'444	66'494	-15'050
5	111'811	93'451	18'360
6	187'084	189'851	-2'767
7	411'864	401'134	10'731
8	1'433'505	1'490'962	-57'458
Total	2'237'826	2'277'900	-40'075

Table 5: Realization of the observable CDR at time $t = I + 1$, in Euro $1'000$

We observe that the estimated standard deviation of the true aggregate CDR is equal to Euro $65'412$, which means that it is not unlikely to have the true aggregate CDR in the range of about Euro $\pm 40'000$. Moreover, we see that the square root of the estimate for the MSE of the observable aggregate CDR is of size Euro $47'909$ (see Table 6), i.e., it has about the same size as the realization of the observable aggregate CDR. This means that we can not exclude that the negative loss experience for prior accident years was only caused by the fact that we didn't know the true chain ladder factors.

Also observe that the estimate for a possible bias of Euro $45'123$ is quite large, which indicates a large uncertainty in the parameter estimates. The estimate for the standard deviation of the observable aggregate CDR (relative to its mean) is given by Euro $75'412$. As expected, it is substantially larger compared to the process standard deviation of the true aggregate CDR. The main reason for this fact is that there is a relatively large covariance term (Euro $36'304$) between the accident years in the observable aggregate CDR (they use the same observations). The other reason is given by inequality (3.19). The prediction uncertainty in the budget values for the observable CDR by 0 is given by 87'881. Hence our observation of $-40'000$ is in a reasonable range and therefore not an outlier.

Finally, these results are compared to the classical Mack formula [4] for the estimate of the conditional mean square error of prediction of the ultimate claim $\widehat{C}_{i,J}^I$ and $\sum_{i=1}^I \widehat{C}_{i,J}^I$ for a single accident year and aggregated accident years, respectively, in the distribution-free

	$\widehat{R}_i^{\mathcal{D}_I}$	$\widehat{\mathrm{Var}}^{1/2}$	$\widehat{\mathrm{MSE}}_{\mathcal{D}_I}^{1/2}$	$\left(\widehat{\Phi}_{i,J}^I\right)^{1/2}$	u-bias	$\left(\widehat{\Gamma}_{i,J}^I\right)^{1/2}$	$\left(\widehat{\Gamma}_{i,J}^I + \text{u-bias}^2\right)^{1/2}$	$\mathrm{msep}_{Mack}^{1/2}$
0	0							
1	4'378	395	407	0	407	395	567	567
2	9'348	1'185	900	204	876	1'203	1'488	1'566
3	28'392	3'395	1'966	413	1'922	3'420	3'923	4'157
4	51'444	8'673	4'395	921	4'298	8'721	9'723	10'536
5	111'811	25'877	11'804	1'982	11'636	25'953	28'443	30'319
6	187'084	18'875	9'100	4'582	7'863	19'423	20'954	35'967
7	411'864	25'822	11'131	5'211	9'836	26'343	28'119	45'090
8	1'433'505	49'978	18'581	6'083	17'558	50'347	53'320	69'552
$\mathrm{cov}^{1/2}$		0	39'746	12'999	37'560	36'304	52'237	50'361
Total	2'237'826	65'412	47'909	16'097	45'123	75'412	87'881	108'401

Table 6: Volatilities of the estimates in Euro $1'000$ with:

$\widehat{R}_i^{\mathcal{D}_I}$	estimated reserves at time $t = I$, cf. (2.19)
$\widehat{\mathrm{Var}}^{1/2} = \widehat{\mathrm{Var}}\left(\mathrm{CDR}_i(I+1)\,\Big|\,\mathcal{D}_I\right)^{1/2}$	estimated std. dev. of the true CDR, cf. (3.2)
$\widehat{\mathrm{MSE}}_{\mathcal{D}_I}^{1/2} = \widehat{\mathrm{MSE}}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right)^{1/2}$	estimated $\mathrm{mse}^{1/2}$ between true and observable CDR, cf. (3.4) and (3.8)
$\left(\widehat{\Phi}_{i,J}^I\right)^{1/2}$	process std. dev. of the difference betw. true and observable CDR, cf. formulas (3.6) and (3.9)
u-bias	uncertainty in the bias term, cf. (3.5), (3.10) and (3.14)
$\left(\widehat{\Gamma}_{i,J}^I\right)^{1/2} = \widehat{\mathrm{Var}}\left(\widehat{\mathrm{CDR}}_i(I+1)\Big|\,\mathcal{D}_I\right)^{1/2}$	std. dev. of the observable CDR, cf. (3.17)
$\left(\widehat{\Gamma}_{i,J}^I + \text{u-bias}^2\right)^{1/2}$	prediction std. dev. of $\widehat{\mathrm{CDR}}_i(I+1)$ compared to 0, cf. (3.15)
$\mathrm{msep}_{Mack}^{1/2}$	mean square error of prediction$^{1/2}$ of the ultimate claim, cf. Mack [4]

chain ladder model:

$$\mathrm{msep}_{Mack}\left(\widehat{C}_{i,J}^{I}\right) = E\left[\left(C_{i,J} - \widehat{C}_{i,J}^{I}\right)^2 \middle| \mathcal{D}_I\right] \tag{4.2}$$

and

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

Notice that the information in the next accounting year (diagonal $I+1$) contributes substantially to the total uncertainty of the ultimate claims over prior accident years, whose prediction standard deviation is estimated by Euro $108'401$ (this is also confirmed by the field study AISAM-ACME [1]).

Conclusions

In the example considered (see Table 1 and 5) we have a fairly large uncertainty coming from the parameter estimates (see Table 6). Due to this uncertainty we can not exclude that we could also have a positive loss experience for the prior years, if we had put our reserves at $E\left[R_i^I \middle| \mathcal{D}_I\right]$ instead of $\widehat{R}_i^{\mathcal{D}_I}$, at time $t = I$.

Of course, this is only a simplified example compared to practical problems. In practice, the total business is usually split into different (hopefully homogeneous) subclasses. Then for each subclass the reserves are determined by an appropriate method (possibly different from the chain ladder method on cumulative payments). It remains an open challenging problem to do a similar analysis for different methods and then aggregate the uncertainties to a total uncertainty of the whole portfolio. However, we believe that our results will serve as a first step in addressing this problem and, moreover, provide a range for possible numerical results.

For future research, we propose to study the claims reserving process over the whole time horizon, i.e., to study the process $\widehat{R}_i^{\mathcal{D}_{I+k}}$ for $k \geq 0$. This then represents the run-off situation of the claims reserving process $E\left[R_i^{I+k} \middle| \mathcal{D}_{I+k}\right]$, $k \geq 0$, and may lead to the construction of the non-life run-off valuation portfolio, see Buchwalder et al. [2].

5 Proofs and derivations

In this section we present key steps in our derivation of the estimators given in Results 3.1, 3.2 and 3.3.

5.1 Preliminary results

The following lemma will be useful for our derivations.

Lemma 5.1 *Under Model Assumptions 2.1 we have*

a) $C_{i,I-i+1}, \widehat{f}_{I-i+1}^{I+1}, \ldots, \widehat{f}_{J-1}^{I+1}$ *are conditionally independent w.r.t.* $\mathcal{D}_I$*;*

b) $E\left[\widehat{f}_l^{I+1}\middle|\mathcal{D}_I\right] = \frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}} + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}} = \frac{S_l^I}{S_l^{I+1}} \cdot \widehat{f}_l^I + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}}$*;*

c) $E\left[\widehat{C}_{i,j}^{I+1}\middle|\mathcal{D}_I\right] = C_{i,I-i} \cdot f_{I-i} \cdot \prod_{l=I-i+1}^{j-1} E\left[\widehat{f}_l^{I+1}\middle|\mathcal{D}_I\right]$*;*

d) $E\left[C_{i,I-i+1}^2\middle|\mathcal{D}_I\right] = f_{I-i}^2 \cdot C_{i,I-i}^2 + \sigma_{I-i}^2 \cdot C_{i,I-i}$*;*

e) $E\left[\left(\widehat{f}_l^{I+1}\right)^2\middle|\mathcal{D}_I\right] = \left(\frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}} + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}}\right)^2 + \frac{\sigma_l^2}{\left(S_l^{I+1}\right)^2} \cdot C_{I-l,l}$*;*

f) $E\left[C_{i,I-i+1} \cdot \widehat{f}_{I-i}^{I+1}\middle|\mathcal{D}_I\right] = \frac{1}{S_{I-i}^{I+1}} \cdot \left(f_{I-i}^2 \cdot C_{i,I-i}^2 + \sigma_{I-i}^2 \cdot C_{i,I-i} + S_{I-i+1}^{I+1} \cdot C_{i,I-i} \cdot f_{I-i}\right)$.

Proof of Lemma 5.1.

a)-c) See proof of Lemma 2.4 in Merz-Wüthrich [3].

d) Follows immediately from Model Assumptions 2.1.

e) From (2.10) we get

$$E\left[\left(\widehat{f}_l^{I+1}\right)^2\middle|\mathcal{D}_I\right] = \left(\frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}}\right)^2 + 2 \cdot \frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{\left(S_l^{I+1}\right)^2} \cdot E\left[C_{I-l,l+1}\middle|\mathcal{D}_I\right] + \frac{E\left[C_{I-l,l+1}^2\middle|\mathcal{D}_I\right]}{\left(S_l^{I+1}\right)^2} \tag{5.1}$$

and together with part d) we obtain claim e).

f) Using (2.10), part d) and (2.8) we obtain

$$
\begin{aligned}
E\left[C_{i,I-i+1} \cdot \widehat{f}_{I-i}^{I+1} \middle| \mathcal{D}_I \right] &= \frac{1}{S_{I-i}^{I+1}} \cdot E\left[C_{i,I-i+1} \cdot \sum_{k=0}^{i} C_{k,I-i+1} \middle| \mathcal{D}_I \right] \\
&= \frac{1}{S_{I-i}^{I+1}} \cdot E\left[C_{i,I-i+1}^2 \middle| \mathcal{D}_I \right] + \frac{1}{S_{I-i}^{I+1}} \cdot \sum_{k=0}^{i-1} C_{k,I-i+1} \cdot E\left[C_{i,I-i+1} \middle| \mathcal{D}_I \right] \\
&= \frac{1}{S_{I-i}^{I+1}} \cdot \left(f_{I-i}^2 \cdot C_{i,I-i}^2 + \sigma_{I-i}^2 \cdot C_{i,I-i} + S_{I-i+1}^{I+1} \cdot C_{i,I-i} \cdot f_{I-i} \right).
\end{aligned}
\tag{5.2}
$$

This completes the proof of Lemma 5.1.

□

5.2 Derivation of Result 3.1 for a single accident year

To derive Result 3.1, we use the decomposition given in (3.5) and estimate each of its terms separately.

Lemma 5.2 *Under Model Assumptions 2.1, we have*

$$
\begin{aligned}
\Phi_{i,J}^I = E\left[C_{i,J} \middle| \mathcal{D}_I \right]^2 \cdot \frac{\sigma_{I-i}^2 / f_{I-i}^2}{C_{i,I-i}} + f_{I-i}^2 \cdot C_{i,I-i}^2 \cdot \mathrm{Var}\left(\prod_{l=I-i+1}^{J-1} \widehat{f}_l^{I+1} \middle| \mathcal{D}_I \right) \\
+ \sigma_{I-i}^2 \cdot C_{i,I-i} \cdot \left(\prod_{l=I-i+1}^{J-1} E\left[\left(\widehat{f}_l^{I+1} \right)^2 \middle| \mathcal{D}_I \right] - 2 \prod_{j=I-i+1}^{J-1} E\left[\widehat{f}_j^{I+1} \middle| \mathcal{D}_I \right] \cdot \prod_{j=I-i+1}^{J-1} f_j \right)
\end{aligned}
\tag{5.3}
$$

for $i \in \{2, \ldots, I\}$ *and* $\Phi_{1,J}^I = 0$.

Proof. For $i = 1$, we have (recall $I = J$)

$$
\begin{aligned}
\Phi_{1,J}^I &= \mathrm{Var}\left(\mathrm{CDR}_1(I+1) - \widehat{\mathrm{CDR}}_1(I+1) \middle| \mathcal{D}_I \right) \\
&= \mathrm{Var}\left(E\left[C_{1,J} \middle| \mathcal{D}_{I+1} \right] - \widehat{C}_{1,J}^{I+1} \middle| \mathcal{D}_I \right) = \mathrm{Var}\left(C_{1,J} - C_{1,J} \middle| \mathcal{D}_I \right) = 0.
\end{aligned}
\tag{5.4}
$$

Consider $i > 1$. For $\Phi_{i,J}^I$ we use the decomposition given in (3.6). For the first term on the right-hand side of (3.6), the process variance of the true CDR, using Model Assumptions

2.1, (2.15) and (2.8), we obtain

$$\begin{aligned}\mathrm{Var}\big(\mathrm{CDR}_i(I+1)\,\big|\,\mathcal{D}_I\big) &= \mathrm{Var}\big(E\left[C_{i,J}\right|\mathcal{D}_I] - E\left[C_{i,J}\right|\mathcal{D}_{I+1}]\,\big|\,\mathcal{D}_I\big) \\ &= \mathrm{Var}\big(E\left[C_{i,J}\right|\mathcal{D}_{I+1}]\,\big|\,\mathcal{D}_I\big) \\ &= \prod_{j=I-i+1}^{J-1} f_j^2 \cdot \mathrm{Var}(C_{i,I-i+1}|\mathcal{D}_I) \\ &= C_{i,I-i} \cdot \prod_{j=I-i+1}^{J-1} f_j^2 \cdot \sigma_{I-i}^2 \\ &= E\left[C_{i,J}\right|\mathcal{D}_I]^2 \cdot \frac{\sigma_{I-i}^2/f_{I-i}^2}{C_{i,I-i}}.\end{aligned} \tag{5.5}$$

To estimate the last two terms in the decomposition (3.6), observe that, using (2.15) and (2.17),

$$\begin{aligned}&\mathrm{Var}\left(\widehat{\mathrm{CDR}}_i(I+1)\Big|\,\mathcal{D}_I\right) - 2\cdot\mathrm{Cov}\left(\widehat{\mathrm{CDR}}_i(I+1), \mathrm{CDR}_i(I+1)\Big|\,\mathcal{D}_I\right) \\ &\quad = \mathrm{Var}\left(\widehat{C}_{i,J}^{I+1}\Big|\,\mathcal{D}_I\right) - 2\cdot\mathrm{Cov}\left(\widehat{C}_{i,J}^{I+1}, E\left[C_{i,J}\right|\mathcal{D}_{I+1}]\Big|\,\mathcal{D}_I\right).\end{aligned} \tag{5.6}$$

Using Lemma 5.1 a) and d) the first term on the right-hand side of (5.6) is

$$\begin{aligned}\mathrm{Var}\left(\widehat{C}_{i,J}^{I+1}\Big|\,\mathcal{D}_I\right) &= E\left[\left(\widehat{C}_{i,J}^{I+1}\right)^2\Big|\,\mathcal{D}_I\right] - \left(E\left[\widehat{C}_{i,J}^{I+1}\Big|\,\mathcal{D}_I\right]\right)^2 \\ &= E\left[C_{i,I-i+1}^2\right|\mathcal{D}_I] \cdot \prod_{l=I-i+1}^{J-1} E\left[\left(\widehat{f}_l^{I+1}\right)^2\Big|\,\mathcal{D}_I\right] \\ &\qquad - \left(E\left[C_{i,I-i+1}\right|\mathcal{D}_I]\right)^2 \cdot \prod_{l=I-i+1}^{J-1}\left(E\left[\widehat{f}_l^{I+1}\Big|\,\mathcal{D}_I\right]\right)^2 \\ &= \left(f_{I-i}^2 \cdot C_{i,I-i}^2 + \sigma_{I-i}^2 \cdot C_{i,I-i}\right) \cdot \prod_{l=I-i+1}^{J-1} E\left[\left(\widehat{f}_l^{I+1}\right)^2\Big|\,\mathcal{D}_I\right] \\ &\qquad - f_{I-i}^2 \cdot C_{i,I-i}^2 \cdot \prod_{l=I-i+1}^{J-1}\left(E\left[\widehat{f}_l^{I+1}\Big|\,\mathcal{D}_I\right]\right)^2 \\ &= f_{I-i}^2 \cdot C_{i,I-i}^2 \cdot \mathrm{Var}\left(\prod_{I-i+1}^{J-1} \widehat{f}_l^{I+1}\,\Bigg|\,\mathcal{D}_I\right) \\ &\qquad + \sigma_{I-i}^2 \cdot C_{i,I-i} \cdot \prod_{l=I-i+1}^{J-1} E\left[\left(\widehat{f}_l^{I+1}\right)^2\Big|\,\mathcal{D}_I\right].\end{aligned} \tag{5.7}$$

Applying Lemma 5.1 a), c) and d), the covariance term can be rewritten as

$$\begin{aligned}
\operatorname{Cov}\left(\widehat{C}_{i,J}^{I+1}, E\left[C_{i,J}\middle|\mathcal{D}_{I+1}\right]\middle|\mathcal{D}_I\right) &= \operatorname{Cov}\left(C_{i,I-i+1}\cdot\prod_{j=I-i+1}^{J-1}\widehat{f}_j^{I+1}, C_{i,I-i+1}\cdot\prod_{j=I-i+1}^{J-1}f_j\middle|\mathcal{D}_I\right)\\
&= \operatorname{Var}\left(C_{i,I-i+1}\middle|\mathcal{D}_I\right)\cdot\prod_{j=I-i+1}^{J-1}E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]\cdot\prod_{j=I-i+1}^{J-1}f_j\\
&= \sigma_{I-i}^2\cdot C_{i,I-i}\cdot\prod_{j=I-i+1}^{J-1}E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]\cdot\prod_{j=I-i+1}^{J-1}f_j.
\end{aligned}\tag{5.8}$$

Hence, from (5.5) -(5.8) we obtain the expression for $\Phi_{i,J}^I$. This completes the proof.

□

An estimator $\widehat{\Phi}_{i,J}^I$ for $\Phi_{i,J}^I$ is obtained by replacing the unknown parameters on the right-hand side of (5.3) by their estimators at time $t = I$. Then, for the first term in (5.3) we have (see also (5.5))

$$\widehat{\operatorname{Var}}\big(\mathrm{CDR}_i(I+1)\mid\mathcal{D}_I\big) = \left(\widehat{C}_{i,J}^I\right)^2\cdot\frac{\left(\widehat{\sigma}_{I-i}^I\right)^2/\left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}},\tag{5.9}$$

where $\left(\widehat{\sigma}_{I-i}^I\right)^2$ was given in (3.11). Observe that $\left(\widehat{\sigma}_{I-i}^I\right)^2$ is an unbiased estimator for σ_{I-i}^2. Also note that the conditional moments on the right-hand side of (5.3) are functions of the unknown parameters f_l and σ_l^2 (cf. Lemma 5.1 a), b) and e)). If these are replaced by their best estimates $\widehat{f}_l^I$ and $\left(\widehat{\sigma}_l^I\right)^2$ at time $t = I$, we obtain the following estimators:

$$\begin{aligned}
\widehat{E}\left[\widehat{f}_l^{I+1}\middle|\mathcal{D}_I\right] &= \widehat{f}_l^I,\\
\widehat{E}\left[\left(\widehat{f}_l^{I+1}\right)^2\middle|\mathcal{D}_I\right] &= \left(\widehat{f}_l^I\right)^2 + \frac{\left(\widehat{\sigma}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}.
\end{aligned}\tag{5.10}$$

Using (5.10) and replacing the other parameters f_{I-i} and σ_{I-i}^2 by their best estimates at time $t = I$ gives for the second and third term on the right-hand side of (5.3)

$$\begin{aligned}
&\left(\left(\widehat{f}_{I-i}^I\right)^2\cdot C_{i,I-i}^2 + \left(\widehat{\sigma}_{I-i}^I\right)^2\cdot C_{i,I-i}\right)\cdot\prod_{l=I-i+1}^{J-1}\left(\left(\widehat{f}_l^I\right)^2 + \frac{\left(\widehat{\sigma}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)\\
&\quad - \left(\widehat{C}_{i,J}^I\right)^2\cdot\left[1 + 2\cdot\frac{\left(\widehat{\sigma}_{I-i}^I\right)^2/\left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}}\right].
\end{aligned}\tag{5.11}$$

This implies that the process variances of the true and the observable CDR together with the covariance term (cf. (3.6)) are estimated by

$$\begin{aligned}\widehat{\Phi}^I_{i,J} = &\left(\left(\widehat{f}^I_{I-i}\right)^2 \cdot C^2_{i,I-i} + \left(\widehat{\sigma}^I_{I-i}\right)^2 \cdot C_{i,I-i}\right) \cdot \prod_{l=I-i+1}^{J-1} \left(\left(\widehat{f}^I_l\right)^2 + \frac{\left(\widehat{\sigma}^I_l\right)^2}{\left(S^{I+1}_l\right)^2} \cdot C_{I-l,l}\right) \\ &- \left(\widehat{C}^I_{i,J}\right)^2 \cdot \left[1 + \frac{\left(\widehat{\sigma}^I_{I-i}\right)^2 / \left(\widehat{f}^I_{I-i}\right)^2}{C_{i,I-i}}\right] \\ = &\left(\widehat{C}^I_{i,J}\right)^2 \cdot \left[1 + \frac{\left(\widehat{\sigma}^I_{I-i}\right)^2 / \left(\widehat{f}^I_{I-i}\right)^2}{C_{i,I-i}}\right] \cdot \left(\prod_{l=I-i+1}^{J-1} \left(1 + \frac{\left(\widehat{\sigma}^I_l\right)^2 / \left(\widehat{f}^I_l\right)^2}{\left(S^{I+1}_l\right)^2} \cdot C_{I-l,l}\right) - 1\right).\end{aligned} \tag{5.12}$$

This is the first part of Result 3.1.

Now we come to the bias term

$$\left(E\left[\widehat{\mathrm{CDR}}_i(I+1) \middle| \mathcal{D}_I\right]\right)^2. \tag{5.13}$$

Observe that it can not be calculated explicitly because the true chain ladder factors f_j are not known. Hence we need another technique to estimate this term. There are different approaches in the literature (unconditional and conditional ones, see Mack [4], Buchwalder et al. [2] and Mack et al. [5]) for the estimation of this term. Using a conditional resampling technique (Approach 3 in Buchwalder et al. [2]), the bias term (5.13) is estimated by

$$\widehat{E}_{\mathcal{D}_I}\left[\left(E\left[\widehat{\mathrm{CDR}}_i(I+1) \;\middle|\; \mathcal{D}_I\right]\right)^2\right] = \left(\widehat{C}^I_{i,J}\right)^2 \cdot \widehat{\Delta}^I_{i,J}, \tag{5.14}$$

where

$$\widehat{\Delta}^I_{i,J} = \frac{\left(\widehat{\sigma}^I_{I-i}\right)^2 / \left(\widehat{f}^I_{I-i}\right)^2}{S^I_{I-i}} + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{S^{I+1}_j}\right)^2 \cdot \frac{\left(\widehat{\sigma}^I_j\right)^2 / \left(\widehat{f}^I_j\right)^2}{S^I_j}. \tag{5.15}$$

For details on the derivation of (5.14), we refer to Result 4.4 in Merz-Wüthrich [3]. Result 3.1 is obtained by adding up the above estimators (5.12) and (5.14) for the variance term and the bias, respectively.

5.3 Derivation of Result 3.2 - aggregation over accident years

Consider two accident years i and k with $i > k$. Using the fact that claims amounts in different accident years are independent and (2.16), we obtain

$$\begin{aligned} &\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)+\widehat{\mathrm{CDR}}_k(I+1)\right) \\ &\qquad = \mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right)+\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_k(I+1)\right) \\ &\qquad\quad - 2\cdot E\left[\left.\widehat{\mathrm{CDR}}_i(I+1)\cdot \mathrm{CDR}_k(I+1)\right|\mathcal{D}_I\right] \\ &\qquad\quad - 2\cdot E\left[\left.\widehat{\mathrm{CDR}}_k(I+1)\cdot \mathrm{CDR}_i(I+1)\right|\mathcal{D}_I\right] \\ &\qquad\quad + 2\cdot E\left[\left.\widehat{\mathrm{CDR}}_i(I+1)\cdot \widehat{\mathrm{CDR}}_k(I+1)\right|\mathcal{D}_I\right]. \end{aligned} \tag{5.16}$$

Observe that the random variable $\widehat{\mathrm{CDR}}_k(I+1)$ does not depend on the observations of accident year $i > k$. Hence the fourth term on the right-hand side of (5.16) disappears, because of the independence between different accident years and (2.16). We define

$$\Psi^I_{i,k} = \mathrm{Cov}\left(\left.\widehat{\mathrm{CDR}}_i(I+1),\widehat{\mathrm{CDR}}_k(I+1)\right|\mathcal{D}_I\right) - \mathrm{Cov}\left(\left.\widehat{\mathrm{CDR}}_i(I+1),\mathrm{CDR}_k(I+1)\right|\mathcal{D}_I\right). \tag{5.17}$$

Then we can rewrite (5.16) as follows

$$\begin{aligned} &\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)+\widehat{\mathrm{CDR}}_k(I+1)\right) \\ &\qquad = \mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_i(I+1)\right)+\mathrm{MSE}_{\mathcal{D}_I}\left(\widehat{\mathrm{CDR}}_k(I+1)\right) \\ &\qquad\quad + 2\cdot\left(\Psi^I_{i,k} + E\left[\left.\widehat{\mathrm{CDR}}_i(I+1)\right|\mathcal{D}_I\right]\cdot E\left[\left.\widehat{\mathrm{CDR}}_k(I+1)\right|\mathcal{D}_I\right]\right). \end{aligned} \tag{5.18}$$

This means that in addition to the MSE for single accident years we also have a covariance term $\Psi^I_{i,k}$, as well as a term

$$E\left[\left.\widehat{\mathrm{CDR}}_i(I+1)\right|\mathcal{D}_I\right]\cdot E\left[\left.\widehat{\mathrm{CDR}}_k(I+1)\right|\mathcal{D}_I\right] \tag{5.19}$$

for the bias. This is similar to (3.5). The reason for these additional terms is that we use the same observations to estimate the chain ladder factors which are then applied to different accident years i and k.

Lemma 5.3 *Under Model Assumptions 2.1 we have for $i > k > 1$ that $\Psi^I_{i,k}$ is given by* (5.24)*, and $\Psi^I_{i,1} = 0$ for $i > 1$.*

Proof. The proof of $\Psi_{i,1}^I = 0$ is similar to the proof of $\Phi_{1,J}^I = 0$ (see proof of Lemma 5.2 above). Consider $k > 1$. We have

$$\begin{aligned}
\Psi_{i,k}^I &= \operatorname{Cov}\left(\widehat{\mathrm{CDR}}_i(I+1), \widehat{\mathrm{CDR}}_k(I+1) - \mathrm{CDR}_k(I+1) \Big| \mathcal{D}_I \right) \\
&= \operatorname{Cov}\left(\widehat{C}_{i,J}^{I+1}, \widehat{C}_{k,J}^{I+1} - E\left[C_{k,J} \middle| \mathcal{D}_{I+1} \right] \Big| \mathcal{D}_I \right) \\
&= E\left[\widehat{C}_{i,J}^{I+1} \cdot \left(\widehat{C}_{k,J}^{I+1} - E\left[C_{k,J} \middle| \mathcal{D}_{I+1} \right] \right) \Big| \mathcal{D}_I \right] \\
&\qquad - E\left[\widehat{C}_{i,J}^{I+1} \Big| \mathcal{D}_I \right] \cdot E\left[\widehat{C}_{k,J}^{I+1} - E\left[C_{k,J} \middle| \mathcal{D}_{I+1} \right] \Big| \mathcal{D}_I \right].
\end{aligned} \tag{5.20}$$

The last term on the right-hand side of (5.20) can easily be calculated using Lemma 5.1 a) and c), hence

$$\begin{aligned}
& E\left[\widehat{C}_{i,J}^{I+1} \Big| \mathcal{D}_I \right] \cdot E\left[\widehat{C}_{k,J}^{I+1} - E\left[C_{k,J} \middle| \mathcal{D}_{I+1} \right] \Big| \mathcal{D}_I \right] \\
&\quad = E\left[C_{i,I-i+1} \middle| \mathcal{D}_I \right] \cdot \prod_{j=I-i+1}^{J-1} E\left[\widehat{f}_j^{I+1} \Big| \mathcal{D}_I \right] \\
&\qquad \cdot C_{k,I-k} \cdot f_{I-k} \cdot \left(\prod_{j=I-k+1}^{J-1} E\left[\widehat{f}_j^{I+1} \Big| \mathcal{D}_I \right] - \prod_{j=I-k+1}^{J-1} f_j \right).
\end{aligned} \tag{5.21}$$

Now there remains to calculate the first term on the right-hand side of (5.20). Using Lemma 5.1 a) we get

$$\begin{aligned}
E\left[\widehat{C}_{i,J}^{I+1} \cdot \widehat{C}_{k,J}^{I+1} \Big| \mathcal{D}_I \right] &= E\left[C_{i,I-i+1} \middle| \mathcal{D}_I \right] \cdot \prod_{j=I-i+1}^{I-k-1} E\left[\widehat{f}_j^{I+1} \Big| \mathcal{D}_I \right] \\
&\quad \cdot E\left[C_{k,I-k+1} \cdot \widehat{f}_{I-k}^{I+1} \Big| \mathcal{D}_I \right] \cdot \prod_{j=I-k+1}^{J-1} E\left[\left(\widehat{f}_j^{I+1} \right)^2 \Big| \mathcal{D}_I \right],
\end{aligned} \tag{5.22}$$

and, using Lemma 5.1 a) once more and the independence between different accident years,

$$\begin{aligned}
& E\left[\widehat{C}_{i,J}^{I+1} \cdot E\left[C_{k,J} \middle| \mathcal{D}_{I+1} \right] \Big| \mathcal{D}_I \right] \\
&\quad = E\left[C_{i,I-i+1} \middle| \mathcal{D}_I \right] \cdot E\left[C_{k,I-k+1} \cdot \prod_{j=I-i+1}^{J-1} \widehat{f}_j^{I+1} \Bigg| \mathcal{D}_I \right] \cdot \prod_{j=I-k+1}^{J-1} f_j \\
&\quad = E\left[C_{i,I-i+1} \middle| \mathcal{D}_I \right] \cdot \prod_{\substack{j=I-i+1 \\ j \neq I-k}}^{J-1} E\left[\widehat{f}_j^{I+1} \Big| \mathcal{D}_I \right] \cdot \prod_{j=I-k+1}^{J-1} f_j \cdot E\left[C_{k,I-k+1} \cdot \widehat{f}_{I-k}^{I+1} \Big| \mathcal{D}_I \right].
\end{aligned} \tag{5.23}$$

If we put the above together, we find

$$
\begin{aligned}
\Psi_{i,k}^I &= E\left[C_{i,I-i+1}\middle|\mathcal{D}_I\right]\cdot\prod_{j=I-i+1}^{I-k-1}E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right] \qquad (5.24)\\
&\cdot\Bigg[E\left[C_{k,I-k+1}\cdot\widehat{f}_{I-k}^{I+1}\middle|\mathcal{D}_I\right]\cdot\left(\prod_{j=I-k+1}^{J-1}E\left[\left(\widehat{f}_j^{I+1}\right)^2\middle|\mathcal{D}_I\right]-\prod_{j=I-k+1}^{J-1}f_j\cdot E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]\right)\\
&-E\left[C_{k,I-k+1}\middle|\mathcal{D}_I\right]\cdot E\left[\widehat{f}_{I-k}^{I+1}\middle|\mathcal{D}_I\right]\left(\prod_{j=I-k+1}^{J-1}E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]^2-\prod_{j=I-k+1}^{J-1}f_j\cdot E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]\right)\Bigg].
\end{aligned}
$$

This completes the proof.

□

The next step is as in (5.12), i.e., we replace the unknown parameters on the right-hand side of (5.24) by their estimators at time $t = I$. Using Lemma 5.1 f), (2.9) and (2.10), we obtain for the conditional expectation $E\left[C_{k,I-k+1}\cdot\widehat{f}_{I-k}^{I+1}\middle|\mathcal{D}_I\right]$ on the right-hand side of (5.24) the following estimator at time I

$$
\begin{aligned}
&\widehat{E}\left[C_{k,I-k+1}\cdot\widehat{f}_{I-k}^{I+1}\middle|\mathcal{D}_I\right] \qquad (5.25)\\
&\quad=\frac{1}{S_{I-k}^{I+1}}\cdot\left(\left(\widehat{f}_{I-k}^I\right)^2\cdot C_{k,I-k}^2+\left(\widehat{\sigma}_{I-k}^I\right)^2\cdot C_{k,I-k}+S_{I-k+1}^{I+1}\cdot C_{k,I-k}\cdot\widehat{f}_{I-k}^I\right)\\
&\quad=\left(\widehat{f}_{I-k}^I\right)^2\cdot C_{k,I-k}\cdot\left(1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{S_{I-k}^{I+1}}\right).
\end{aligned}
$$

Since f_j and $E\left[\widehat{f}_j^{I+1}\middle|\mathcal{D}_I\right]$ are estimated by $\widehat{f}_j^I$, the last term on the right-hand side of (5.24) disappears (cf. (5.10); the bias is considered by (5.19)). This gives the following estimator for $\Psi_{i,k}^I$:

$$
\begin{aligned}
\widehat{\Psi}_{i,k}^I &= \widehat{C}_{i,I-k}^I\cdot\widehat{E}\left[C_{k,I-k+1}\cdot\widehat{f}_{I-k}^{I+1}\middle|\mathcal{D}_I\right] \qquad (5.26)\\
&\quad\cdot\left(\prod_{l=I-k+1}^{J-1}\left(\left(\widehat{f}_l^I\right)^2+\frac{\left(\widehat{\sigma}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)-\prod_{l=I-k+1}^{J-1}\left(\widehat{f}_l^I\right)^2\right)\\
&=\widehat{C}_{i,J}^I\cdot\widehat{C}_{k,J}^I\left(1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{S_{I-k}^{I+1}}\right)\cdot\left(\prod_{l=I-k+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}_l^I\right)^2/\left(\widehat{f}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)-1\right)\\
&=\frac{\widehat{C}_{i,J}^I}{\widehat{C}_{k,J}^I}\cdot\frac{1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{S_{I-k}^{I+1}}}{1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{C_{k,I-k}}}\cdot\widehat{\Phi}_{k,J}^I.
\end{aligned}
$$

Finally, the bias term (5.19) is again estimated using the conditional resampling approach (see Approach 3 in Buchwalder et al. [2]). This has already been done in Merz-Wüthrich [3], formula (4.27), hence the term (5.19) for $i > k$ is estimated by (see also (5.14))

$$\widehat{E}_{\mathcal{D}_I}\left[E\left[\widehat{\mathrm{CDR}}_i(I+1)\,\Big|\,\mathcal{D}_I\right]\cdot E\left[\widehat{\mathrm{CDR}}_k(I+1)\,\Big|\,\mathcal{D}_I\right]\right]=\widehat{C}_{i,J}^I\cdot\widehat{C}_{k,J}^I\cdot\widehat{\Delta}_{k,J}^I. \tag{5.27}$$

We have now derived an estimator of the conditional MSE for two different accident years $i > k$. It is then straightforward to estimate the conditional MSE for arbitrary sums over different accident years. This leads to Result 3.2.

5.4 Derivation of Result 3.3 - observable CDR

Observe that for $i, k > 0$

$$\mathrm{Cov}\left(\widehat{\mathrm{CDR}}_i(I+1),\widehat{\mathrm{CDR}}_k(I+1)\Big|\,\mathcal{D}_I\right)=\mathrm{Cov}\left(\widehat{C}_{i,J}^{I+1},\widehat{C}_{k,J}^{I+1}\Big|\,\mathcal{D}_I\right). \tag{5.28}$$

For $i = k$ we obtain the following estimator, see (5.7) and (5.10),

$$\begin{aligned}\widehat{\Gamma}_{i,J}^I &= \widehat{\mathrm{Var}}\left(\widehat{C}_{i,J}^{I+1}\Big|\,\mathcal{D}_I\right)\\ &=\left(\widehat{C}_{i,J}^I\right)^2\cdot\left\{\left(\left[1+\frac{\left(\widehat{\sigma}_{I-i}^I\right)^2/\left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}}\right]\cdot\prod_{l=I-i+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}_l^I\right)^2/\left(\widehat{f}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)\right)-1\right\}.\end{aligned} \tag{5.29}$$

For $i > k$ the estimator is (see (5.21), (5.22), (5.10) and (5.25)),

$$\begin{aligned}\widehat{\Upsilon}_{i,k}^I &= \widehat{\mathrm{Cov}}\left(\widehat{C}_{i,J}^{I+1},\widehat{C}_{k,J}^{I+1}\Big|\,\mathcal{D}_I\right)\\ &=\widehat{C}_{i,J}^I\cdot\widehat{C}_{k,J}^I\cdot\left\{\left(\left[1+\frac{\left(\widehat{\sigma}_{I-k}^I\right)^2/\left(\widehat{f}_{I-k}^I\right)^2}{S_{I-k}^{I+1}}\right]\prod_{l=I-k+1}^{J-1}\left(1+\frac{\left(\widehat{\sigma}_l^I\right)^2/\left(\widehat{f}_l^I\right)^2}{\left(S_l^{I+1}\right)^2}\cdot C_{I-l,l}\right)\right)-1\right\}.\end{aligned} \tag{5.30}$$

This gives Result 3.3.

References

[1] AISAM-ACME (2007). Study on non-life long tail liabilities: Reserve risk and risk margin assessment under solvency II. Version 17 October 2007.

[2] Buchwalder, M., Bühlmann H., Merz, M., Wüthrich, M.V. (2006). The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). Astin Bulletin 36, 2, 521–542.

[3] Merz, M., Wüthrich, M.V. (2007). Prediction error of the expected claims development result in the chain ladder method. Bulletin of Swiss Association of Actuaries, 1, 117-137.

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

[5] Mack, T., Quarg, G., Braun, C. (2006). The mean square error of prediction in the chain ladder reserving method - a comment. Astin Bulletin, 36, 2, 543–552.

[6] Wüthrich, M.V., Merz, M., Bühlmann, H. (2007). Bounds on the estimation error in the chain ladder method. Accepted for publication in Scandinanvian Actuarial Journal.