The Chain Ladder Reserve Uncertainties Revisited

by Alois Gisler

Paper to be presented at the ASTIN Colloquium 2016 in Lisbon

Abstract:

Chain ladder (CL) is still one of the most popular and most used reserving method for the insurance practice. In 1993 Mack [7] presented the distribution-free CL-model and derived a formula for the uncertainty of the CL-reserves , which refers to the *ultimate prediction uncertainty* measured by the *mean square error of prediction.* For calculating the reserve risk and the cost-of-capital loading in solvency (SST and solvency II) one also needs estimators for the *one-year prediction uncertainty* of all future accounting years until final development.

In a recent paper [11] Merz and Wüthrich considered the different prediction uncertainties, that is the ultimate prediction uncertainty as well as the one year prediction uncertainties for all future accounting years until final development within the framework of a specific Bayesian-CL model. Taking a non-informative prior and after a first Taylor approximation they received the already existing result of Mack for the ultimate prediction uncertainty and formulas for the one-year run-off uncertainties for all future accounting years until final settlement of the run-off. However the Bayesian-CL model and the distribution free CL-model of Mack are two different pairs of shoes. Thus the results derived in [11] are results with regard to a different model and we do not know, whether they are also appropriate in the classical chain ladder model of Mack.

In this paper we derive the different kinds of prediction uncertainties strictly within the framework of the distribution-free CL model of Mack. By doing so, we gain more insight into the differences between the two model approaches and find the following main results: a) the formulas for the one-year prediction uncertainty in the classical Mack-model are different to the Merz-Wüthrich formulas, b) the Merz-Wüthrich formulas are obtained by a first order Taylor expansion, c) the Mack formula as well as the Merz-Wüthrich formulas for the total over all accident years can be written in a simpler way, d) we can see "behind the formulas", as they can be interpreted in an intuitive and understandable way.

Keywords: claims reserving, distribution free chain-ladder model, Bayesian chain-ladder model, conditional mean square error of prediction, ultimate run-off uncertainty, one-year run-off uncertainties, Mack's formula, Wuethrich-Merz formulas , cost of capital loading, market value margin.

1 Introduction

Accurate claims reserves are essential for an insurance company. It is by far the most important item on the liability side of the balance sheet and has a big impact on the profit and loss (P&C) account. A change of the reserves by a small percentage might well turn a positive year result into a negative one and vice versa. When non-life insurance companies went bankrupt, insufficient reserves were mostly one of the main reasons.

Chain Ladder (CL) and Bornhuetter Ferguson (BF) are still the most used and most popular reserving methods in the insurance practice. In this paper we concentrate on the CL-method.

CL has been used for decades for calculating reserves. In its origin it is a pure pragmatic method without an underlying mathematical model. As long as one is only interested in the reserve estimate, there is no need for a mathematical model. But as soon as one is also interested in the accuracy of the CL-reserves, one needs an underlying stochastic model.

Under *run-off risk* we understand the *risk of an adverse claims development*. It can be defined as *minus the claims development result* (CDR). As common in the actuarial literature (see for instance [9]), we will take *the conditional mean square error of prediction (msep)* as a measure of the reserve uncertainties. For best estimate reserves this msep is equal to the conditional expectation of the square of the CDR. Thereby we distinguish between *the ultimate run-off risk* referring to the ultimate claims development result (CDR) and the *one-year run-off risks* referring to the CDR resulting in one accounting year.

It was only in 1993 when Mack [7] presented a stochastic CL model and derived a formula for estimating the msep of the ultimate run-off risk (see Theorem A.1 in appendix A).

With the emergence of the new solvency regulation (Swiss solvency test and solvency II), there arose the need to assess another kind of reserve risk. The risk considered is the change of risk bearing capital within the next year (one-year time horizon). Hence the reserve risk relevant for solvency purposes is the one-year run-off risk in the next accounting year instead of the ultimate run-off risk considered in Mack. This risk is reflected under the position "claims development result" or "loss experience previous years" in next year's P&L account. A formula for estimating this one year uncertainty of the next accounting year was first published in a paper by Merz and Wüthrich [10] in 2008.

As the new solvency regulations are based on a market consistent valuation, the best estimate reserves have to be complemented by a market value margin corresponding to the discounted costs of capital needed for the entire run-off. For this purpose one also needs estimators of the uncertainty of the one-year run-off risk in later accounting years until the end of the claims development.

In a recent paper [11] of end 2014 Merz and Wüthrich reconsidered all the different CL reserve uncertainties within a specific Bayesian CL model (with Gamma-priors for the CL factors and with conditionally Gamma distributed observations (observed individual CL-factors)), which is very similar to a model which was earlier considered in [4]. In this model they derived formulas for all the different kind of CL-uncertainties inclusive the one-year run-off uncertainties in future accounting years. For the distribution-free case, the authors derive the formulas for the reserve uncertainties obtained by taking a non-informative prior followed by a first order Taylor approximation. These formulas can be found in appendix A.

However the Bayesian CL-model is different to the distribution free CL model of Mack. Thus we do not know whether the results derived in [11] are also appropriate for the classical Mack model. For instance, in the model considered in [11], the mean square error of prediction (msep) does only exist, if the observed triangle fulfils certain properties (see Theorem 3.8 and condition (4.3) in [11]). But in the Mack-model the msep does always exist.

In this paper we derive the different kinds of prediction uncertainties strictly within the framework of the distribution-free CL model of Mack. By doing so, we gain more insight into the differences between the two model approaches and find the following main results:

a) The formulas derived for the one-year prediction uncertainty in the classical Mack-model are different to the Merz-Wüthrich formulas.

b) The Merz-Wüthrich formulas are obtained by a first order Taylor expansion.

c) The Mack formula as well as the Merz-Wüthrich formulas for the total over all accident years can be written in a much simpler way.

d) We can see "behind the formulas", as they can be interpreted in an intuitive and understandable way.

e) The derivation of the formulas is straightforward. The mathematics is simple and easy to understand and makes use of two basic tools, the telescope formula 3.5 and the estimation principle 3.6.

In this connection it is also worthwhile to remember a discussion in 2006 with regard to the estimation error in the Mack formula. In [2] the authors suggested a different estimator, which we call the BBMW estimator. There was quite some discussion about these estimators (see [2], [8], [4]). Reconsidering this discussion we come to the conclusion that the Mack formula for the estimation error is the appropriate one, whereas the BBMW estimator is the adequate formula in the CL-Bayes model, but not in the Mack-model. More details about this side result of the paper can be found in appendix B

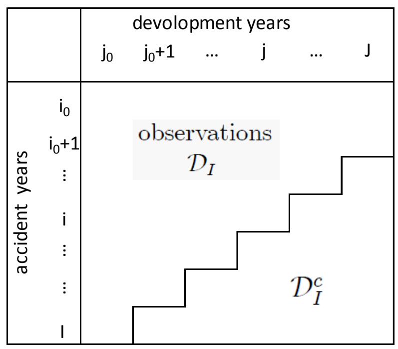
At this point we should also mention the most recent paper [12] of Ancus Röhr. He starts with a first order Taylor approximation of the claims development result (CDR) and calculates the prediction uncertainty of this modified CDR. He then also obtains the Mack as well as the Merz-Wüthrich results.

Organisation of the paper: In section 2 we introduce some notation and the data structure. In section 3 we review the CL-reserving method and the stochastic CL-model of Mack. At the end of this section the telescope formula and the estimation principle are presented. In section 4 we derive the formula for the ultimate run-off uncertainty (Mack-formula). In section 5 we consider the run-off uncertainty of the next accounting year, whereas in section 6 the formulas for the one-year uncertainties for all future accounting years until final claims development are derived. These results can be compared with the Wüthrich-Merz formulas. Finally, a numerical example is presented in section 7.

2 Notation and Data Structure

The starting point of claims reserving are observations $\mathcal{D}_I$ from accident periods $i = i_0, \dots, I$ and development periods $j = j_0, \dots, J$ arranged in a table with i on the vertical axes and j on the horizontal axes. In the following we will call $\mathcal{D}_I$ a data triangle, also in the case, where the shape is a trapezoid.

The data in the triangle are denoted by $C_{i,j} > 0$ and represent the cumulative claim figures (usually claim payments or incurred losses) of accident year $i \in \{i_0, \dots, I\}$ at the end of development year $j \in \{j_0, \dots, J)$. We further assume that the number of accident years is bigger or equal to the number of development years, that is $J - j_0 > I - i_0$. The index j_0 is introduced because in the actuarial literature the first development year is sometimes denoted by zero and sometimes by 1, hence j_0 is either zero or one.



claims development triangle

At time I, the data $C_{i,j} \in \mathcal{D}_I$ in the upper left part are known, whereas the data $C_{i,j} \in \mathcal{D}_I^c$ in the lower right part are future observations we want to predict. We assume that all claims are settled after development year J and that therefore $C_{i,J}$ denotes the ultimate claim of accident year i.

Some notations:

- diagonal functions
  To simplify notation and as already done by Ancus Röhr in [12] it is convenient to introduce the *diagonal functions*

  Definition 2.1 *diagonal functions*

$$j_i := \max\{j \text{ such that } C_{i,j} \in \mathcal{D}_I\}, \qquad (1)$$
$$i_j := \max\{i \text{ such that } C_{i,j} \in \mathcal{D}_I\}. \qquad (2)$$

  Note that C_{i,j_i} is the diagonal element in row i and that $C_{i_j,j}$ is the diagonal element in column j.

- The set of observations at time I is given by

$$\mathcal{D}_I = \{C_{i,j} : i = i_0, \dots, I,\ j \le j_i\}.$$

- Later we will also need the set of observations known at the end development year j, which we denote by

$$\mathcal{B}_j := \{C_{i,k} : C_{i,k} \in \mathcal{D}_I,\ k \leq j\}. \tag{3}$$

- coefficient of variation

  We denote the coefficient of variation of a random variable (r.v.) X by

$$\mathrm{CoV}\,(X) := \frac{\sqrt{\mathrm{Var}\,(X)}}{E\,[X]}$$

- In this paper empty sums and empty products are defined by

Definition 2.2 (empty sum and empty product)

$$\begin{aligned} \sum\nolimits_{l}^{u} x_k &:= 0 \ \ \textit{if}\ u < l, \\ \prod\nolimits_{k=l}^{u} x_k &:= 1 \ \ \textit{if}\ u < l. \end{aligned}$$

- We will later in the paper use the following weights:

Definition 2.3 *The weights $w_{i,j}$ are defined by*

$$w_{i,j} := \begin{cases} C_{i,j}, & if \ \ C_{i,j} \in \mathcal{D}_I \\ \widehat{C}_{i,j}^{CL}, & if \ \ C_{i,j} \in \mathcal{D}_I^c \end{cases}. \tag{4}$$

3 The chain ladder method

3.1 The chain ladder method

The CL method is a pragmatic method which has been used for decades for estimating reserves. The basic assumption behind CL is that, up to random fluctuation, the columns in the development triangle are proportional to each other, i. e. there exist constants f_j, $j = j_0, \ldots, J-1$, such that

$$C_{i,j+1} \approx f_j C_{i,j}. \tag{5}$$

The constants f_j are called claims-development factors, CL factors or age to age factors. Given the information $\mathcal{D}_I$ it is natural to estimate these unknown constants by

$$\widehat{f}_j^{CL} = \frac{\sum_{i=i_0}^{i_j-1} C_{i,j+1}}{\sum_{i=i_0}^{i_j-1} C_{i,j}}. \tag{6}$$

Due to (5), the ultimate claim of accident year i is predicted by

$$\widehat{C}_{i,J}^{CL} = C_{i,j_i} \prod_{j=j_i}^{J-1} \widehat{f}_j^{CL} \tag{7}$$

and the $C_{i,j}$ in the lower right part of the triangle are estimated by

$$\widehat{C}_{i,j}^{CL} = C_{i,j_i} \prod_{k=j_i}^{j-1} \widehat{f}_k^{CL} \text{ for } i = i_0, \ldots, I, \quad j = j_i + 1, \ldots, J. \tag{8}$$

(8) are called the CL-predictions. It is also useful and meaningful to define

$$\widehat{C}_{i,j}^{CL} := C_{i,j} \text{ for } C_{i,j} \in \mathcal{D}_I. \tag{9}$$

The CL reserve $\widehat{R}_i^{CL}$ of accident year i is an estimate of the *outstanding liabilities*

$$R_i = \sum_{j=j_i+1}^{J} C_{i,j} \tag{10}$$

and is the difference between the CL prediction $\widehat{C}_{i,J}^{CL}$ of the ultimate claim and the cumulative claim C_{i,j_i} known at time I, i.e.

$$\widehat{R}_i^{CL} = \widehat{C}_{i,J}^{CL} - C_{i,j_i} \tag{11}$$

Finally

$$\widehat{R}_{tot}^{CL} = \sum_{i=i_0}^{I} \widehat{R}_i^{CL} \tag{12}$$

is the total reserve over all accident years.

3.2 The stochastic CL-model of Mack

The following distribution-free stochastic model underlying the CL method was introduced in [7] by Mack .

Model Assumptions 3.1 (Mack-model)

i) Cumulative claims $C_{i,j}$ of different accident years are independent.

ii) There exist positive parameters $f_{j_0}, \ldots, f_{J-1}$ and $\sigma_{j_0}^2, \ldots, \sigma_{J-1}^2$ such that for all $i = i_0, \ldots, I$, and all $j = j_0, \ldots, J-1$

$$E[C_{i,j+1} | C_{i,j_0}, C_{i,j_0+1}, \ldots, C_{i,j}] = f_j C_{i,j}, \tag{13}$$
$$Var(C_{i,j+1} | C_{i,j_0}, C_{i,j_0+1}, \ldots, C_{i,j}) = \sigma_j^2 C_{i,j} \tag{14}$$

It is useful to introduce the *individual CL ratios*

$$F_{i,j} := \frac{C_{i,j+1}}{C_{i,j}}. \tag{15}$$

Because of model assumptions 3.1 $F_{i,j}$ belonging to different accident years are independent and it holds that

$$E[F_{i,j} | \mathcal{B}_j] = f_j, \tag{16}$$
$$\text{Var}(F_{i,j} | \mathcal{B}_j) = \frac{\sigma_j^2}{w_{i,j}} \text{ with } w_{i,j} = C_{i,j}. \tag{17}$$

We have deliberately written (17) in this more complicated way, because conditional on $C_{i,j}$, the denominator $C_{i,j}$ in $F_{i,j}$ is no longer a r.v., but a constant playing the role of a weight $w_{i,j} = C_{i,j}$.

The following properties with regard to the estimators $\widehat{f}_j^{CL}$ are well known and can easily be verified:

Properties 3.2

i) The estimators (6) *can be written as a weighted mean*

$$\widehat{f}_j^{CL} = \sum_{i=i_0}^{i_j-1} \frac{w_{i,j}}{w_{\bullet,j}} F_{i,j}, \quad \textit{where } w_{\bullet,j} = \sum_{i=i_0}^{i_j-1} w_{i,j}, \tag{18}$$

ii) Conditional on $\mathcal{B}_j$, *it holds that*

$$E\left[\left.\widehat{f}_j^{CL}\right| \mathcal{B}_j\right] = f_j, \tag{19}$$

$$\mathrm{Var}\left(\left.\widehat{f}_j^{CL}\right| \mathcal{B}_j\right) = \frac{\sigma_j^2}{w_{\bullet,j}}, \quad \textit{where } w_{\bullet,j} = \sum_{i=i_0}^{i_j-1} w_{i,j}, \tag{20}$$

and (19) *has minimum variance among all linear unbiased estimators.*

iii) σ_j^2 *can be estimated by*

$$\widehat{\sigma}_j^2 = \frac{1}{i_j - i_0} \sum_{i=i_0}^{i_j-1} w_{i,j} \left(F_{i,j} - \widehat{f}_j^{CL}\right)^2 \tag{21}$$

and it holds that

$$E\left[\left.\widehat{\sigma}_j^2\right| \mathcal{B}_j\right] = \sigma_j^2.$$

Most of the following properties with regard to the $F_{i,j}$ are well known. For completeness a proof is given in appendix C.

Properties 3.3 *It holds that*

i) Conditional expectation:

$$E\left[\left. F_{i,k}\right| \mathcal{B}_j\right] = f_k \quad \textit{for } j \le k \le J-1, \tag{22}$$

ii) Conditional dependence structure:

$$\left\{\left. F_{i,k}\right| \mathcal{B}_j : k = j, \ldots, J-1\right\} \ \textit{are uncorrelated,} \tag{23}$$

iii) Conditional coefficient of variation:

$$\mathrm{CoV}\left(\left. F_{i,k}\right| \mathcal{B}_j\right) = \frac{\sigma_k}{f_k} \sqrt{E\left[\left.\frac{1}{C_{i,k}}\right| \mathcal{B}_j\right]} \quad \textit{for } j \le k \le J-1. \tag{24}$$

3.3 Mean square error of prediction, telescope formula and estimation principle

As measure for the CL reserve uncertainty we use the conditional mean square error of prediction (msep).

Definition 3.4 *The conditional mean square error of prediction (msep) of the CL prediction* $\widehat{C}_{i,J}^{CL}$ *is defined by*

$$msep_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) := E\left[\left.\left(\widehat{C}_{i,J}^{CL} - C_{i,J}\right)^2\right|\mathcal{D}_I\right]$$

The conditional msep of other predictors and estimators are analogously defined. Note that

$$msep_{R_i|\mathcal{D}_I}\left(\widehat{R}_i^{CL}\right) = msep_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right),$$

since $\widehat{R}_i^{CL} = \widehat{C}_{i,J}^{CL} - C_{i,j_i}$ and since C_{i,j_i} is known at time I.

To derive estimators of this msep we will make extensive use of the following telescope formula.

Lemma 3.5 (Telescope Formula) *For any real numbers* x_j *and* y_j, $j = 1, 2, \ldots, J$, *it holds that*

$$\prod_{j=1}^{J} x_i - \prod_{j=1}^{J} y_i = \sum_{j=1}^{J}\left(\prod_{k=1}^{j-1} x_k\right)(x_j - y_j)\left(\prod_{m=i+1}^{I} y_m\right). \tag{25}$$

Proof:
This result is well known. We show it for a product with $J = 3$. The extension to any number J is self evident.

$$\begin{aligned} x_1x_2x_3 - y_1y_2y_3 &= x_1x_2x_3 - x_1x_2y_3 + x_1x_2y_3 - x_1y_2y_3 + x_1y_2y_3 - y_1y_2y_3 \\ &= (x_1 - y_1)\,y_2y_3 + x_1(x_2 - y_2)\,y_3 + x_1x_2(x_3 - y_3). \end{aligned}$$

□

In the formulas of the conditional msep there appear the unknown CL factors f_j. To find an estimator for the msep we have to estimate these unknown constants. As a general principle we replace them by the estimators $\widehat{f}_j$. But we can't do it for the quadratic difference terms $\left(\widehat{f}_j^{CL} - f_j\right)^2$, because this would give an estimator of zero, which is not meaningful. To find an estimator we have to study the fluctuation of $\widehat{f}_j^{CL}$ around f_j. For this purpose we take into account as many observations of $\mathcal{D}_I$ as possible and consider the conditional r.v. $\left.\widehat{f}_j^{CL}\right|\mathcal{B}_j$ with moments

$$E\left[\left.\widehat{f}_j^{CL}\right|\mathcal{B}_j\right] = f_j, \tag{26}$$

$$E\left[\left.\left(\widehat{f}_j^{CL} - f_j\right)^2\right|\mathcal{B}_j\right] = \frac{\sigma_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}. \tag{27}$$

Based on (27) we will use the following estimation principle:

Estimation Principle 3.6

a)

$$\text{estimator of } \left(\widehat{f}_j^{CL} - f_j\right)^2 := \widehat{E}\left[\left.(F_j - f_j)^2\right| \mathcal{B}_j\right] = \frac{\widehat{\sigma}_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}. \tag{28}$$

b) Other functions of f_j such as $\prod_{j=j_i}^{J} f_j^2$ are estimated by replacing the unknown f_j by $\widehat{f}_j^{CL}$.

Remarks:

- Note that

$$\begin{aligned} \frac{\sigma_j^2}{f_j^2} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}} &= \left(\mathrm{CoV}\left(\left.\widehat{f}_j^{CL}\right| \mathcal{B}_j\right)\right)^2 \quad \text{and hence} \\ \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^{CL}} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}} &= \left(\widehat{\mathrm{CoV}}\left(\left.\widehat{f}_j^{CL}\right| \mathcal{B}_j\right)\right)^2. \end{aligned} \tag{29}$$

4 The ultimate run-off prediction uncertainty

The ultimate claims development results (CDR) are defined by

$$\begin{aligned} CDR_{i,ultimate} &= \widehat{C}_{i,J}^{CL} - C_{i,J} \text{ for single accident years } i = 1, \ldots, I, \\ CDR_{tot,ultimate} &= \widehat{C}_{tot,J}^{CL} - C_{tot,J} \text{ for the total over all accident years,} \end{aligned}$$

where

$$\widehat{C}_{tot,J}^{CL} = \sum_{i=i_0}^{I} \widehat{C}_{i,J}^{CL}, \qquad C_{tot,J} = \sum_{i=i_0}^{I} C_{i,J}.$$

In this section we derive the ultimate run-off uncertainties defined as the $msep$ of the CDR in a slightly different way than in Mack (see [7]).

a) **single accident year** i

$$\begin{aligned} msep_{C_{i,J}|\mathcal{D}_I}\left(C_{i,J}^{CL}\right) &= E\left[\left.\left(C_{i,J} - C_{i,J}^{CL}\right)^2\right| \mathcal{D}_I\right] \\ &= \underbrace{E\left[\left.\underbrace{\left(C_{i,J} - E\left[C_{i,J}|\mathcal{D}_I\right]\right)^2}_{A_i}\right| \mathcal{D}_I\right]}_{\text{process variance } PV_i} + \underbrace{E\left[\left.\underbrace{\left(E\left[C_{i,J}|\mathcal{D}_I\right] - \widehat{C_{i,J}}^{CL}\right)^2}_{B_i}\right| \mathcal{D}_I\right]}_{\text{estimation error } EE_i}. \end{aligned} \tag{30}$$

Process Variance PV_i :

$$A_i = C_{i,j_i}\left(\prod_{j=j_i}^{J-1} F_{i,j} - \prod_{j=j_i}^{J-1} f_j\right),$$

where the diagonal element C_{i,j_i} is a multiplicative constant playing the role of a weight w_{i,j_i}.

By applying the telescope formula (25) we obtain

$$\begin{aligned} A_i &= w_{i,j_i}\left\{(F_{i,j_i} - f_{j_i})\prod_{j=j_i+1}^{J-1} f_j + \ldots + \prod_{k=j_i}^{j-1} F_{i,j}\left(F_{i,j} - f_j\right)\prod_{l=k+1}^{J-1} f_l \right. \\ &\qquad \left. \ldots + \prod_{k=j_i}^{J-2} F_{i,k}\left(F_{i,J-1} - f_{J-1}\right)\right\}. \\ &= \sum_{j=j_i}^{J-1} C_{i,j}\left(F_{i,j} - f_j\right)\prod_{k=j+1}^{J-1} f_k. \end{aligned}$$

Hence

$$PV_i = E\left[\left.A_i^2\right|\mathcal{D}_I\right] = E\left[\left(\sum_{j=j_i}^{J-1}\left(C_{i,j}\left(F_{i,j} - f_j\right)\prod\nolimits_{k=j+1}^{J-1} f_k\right)\right)^2 \mathcal{D}_I\right]. \tag{31}$$

The cross terms vanish since

$$\begin{aligned} &E\left[\left.C_{i,j}C_{i,j+k}\left(F_{i,j} - f_j\right)\left(F_{i,j+k} - f_{j+k}\right)\right|\mathcal{D}_I\right] = \\ &\quad = E\left[\left.E\left[\left.C_{i,j}C_{i,j+k}\left(F_{i,j} - f_j\right)\left(F_{i,j+k} - f_{j+k}\right)\right|\mathcal{B}_{j+k}\right]\right|\mathcal{D}_I\right] = 0 \text{ for } k > 0, \end{aligned} \tag{32}$$

where we have used that $\{F_{i,j+k} : k > 0\}$ are conditionally unbiased given $\mathcal{B}_{j+k}$. Hence

$$\begin{aligned} PV_i &= \sum_{j=j_i}^{J-1} E\left[C_{i,j}\left(F_{i,j} - f_j\right)^2\left(\prod\nolimits_{k=j+1}^{J-1} f_k\right)^2\middle|\mathcal{D}_I\right] \\ &= \sum_{j=j_i}^{J-1} E\left[E\left[\left(C_{i,j}\left(F_{i,j} - f_j\right)\prod_{k=j+1}^{J-1} f_k\right)^2\middle|\mathcal{B}_j\right]\mathcal{D}_I\right] \\ &= \sum_{j=j_i}^{J-1}\left(E\left[\left.C_{i,j}\right|\mathcal{D}_I\right]\prod_{k=j+1}^{J-1} f_k^2\right)\sigma_j^2 \qquad (33) \\ &= \sum_{j=j_i}^{J-1} E\left[\left.C_{i,J}\right|\mathcal{D}_I\right]^2\frac{\sigma_j^2}{f_j^2}\frac{1}{E\left[\left.C_{i,j}\right|\mathcal{D}_I\right]}. \qquad (34) \end{aligned}$$

By applying the estimation principle 3.6 we obtain

$$\widehat{PV}_i = \left(\widehat{C}_{i,J}^{CL}\right)^2\left(\sum_{j=j_i}^{J-1}\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\frac{1}{w_{i,j}}\right), \tag{35}$$

where $w_{i,j}$ are the weights as defined in (4).

Remarks:

- (35) is the same formula as in Mack [7].
- From (24) follows that

$$\mathrm{CoV}\left(F_{i,j} \middle| \mathcal{D}_I \right)^2 = \frac{\sigma_j^2}{f_j^2} E\left[\frac{1}{C_{i,j}} \middle| \mathcal{D}_I \right].$$

  Hence the summands in (35) have the following intuitively accessible interpretation

$$\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{w_{i,j}} = \widehat{\mathrm{CoV}}\left(F_{i,j} \middle| \mathcal{D}_I \right)^2. \tag{36}$$

Estimation error EE_i **:**
The *estimation error* EE_i as defined in (30) is given by

$$EE_i = E\left[B_i^2 \middle| \mathcal{D}_I \right], \tag{37}$$

where

$$\begin{aligned} B_i &= \left(E\left[C_{i,J} \middle| \mathcal{D}_I \right] - \widehat{C_{i,J}}^{CL} \right) \\ &= w_{i,j_i} \left(\prod_{j=j_i}^{J-1} f_j - \prod_{j=j_i}^{J-1} \widehat{f}_j^{CL} \right). \end{aligned} \tag{38}$$

By use of the telescope formula (25) we obtain

$$B_i = \sum_{j=j_i}^{J-1} E\left[C_{i,j} \middle| \mathcal{D}_I \right] \left(f_j - \widehat{f}_j^{CL} \right) \prod_{k=j+1}^{J-1} \widehat{f}_k^{CL}. \tag{39}$$

Conditional on $\mathcal{D}_I$, B_i is an unknown constant and by applying the estimation principle 3.6 we obtain

$$\begin{aligned} \widehat{EE}_i &= \widehat{B}_i^2 \\ &= \left(\widehat{C}_{i,J}^{CL} \right)^2 \left\{ \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}} \right\}. \end{aligned} \tag{40}$$

Remarks:

- (40) is the same estimator as in Mack [7].

– In (41) we have seen hat

$$\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}} = \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL}\middle|\mathcal{B}_j\right)^2, \tag{41}$$

which is an intuitively accessible interpretation of the summands in (40).

b) **Total over all accident years**

Process Variance PV_{tot}
Since different accident years are independent it follows that

$$\begin{aligned}\widehat{PV}_{tot} &= \sum_{i=i_0}^{I} \widehat{PV}_i = \sum_{i=i_J+1}^{I} \widehat{PV}_i \\ &= \sum_{i=i_J+1}^{I} \left(\widehat{C}_{i,J}^{CL}\right)^2 \left(\sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{w_{i,j}}\right).\end{aligned}$$

By changing the order of summation between i and j we obtain

$$\widehat{PV}_{tot} = \sum_{j=j_0}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \sum_{i=i_j}^{I} \frac{1}{w_{i,j}} \left(\widehat{C}_{i,J}^{CL}\right)^2. \tag{42}$$

Estimation Error $\boldsymbol{EE}_{tot}$

$$EE_{tot} = E\left[B_{tot}^2\middle|\mathcal{D}_I\right],$$

where

$$\begin{aligned}B_{tot} &= \sum_{i=i_0}^{I} B_i = \sum_{i=i_J+1}^{I} B_i \\ &= \sum_{i=i_J+1}^{I} \left(\sum_{j=j_i}^{J-1} E\left[C_{i,j}\middle|\mathcal{D}_I\right] \prod_{k=j+1}^{J-1} \widehat{f}_k^{CL}\right).\end{aligned}$$

Changing the order of summation between i and j yields

$$B_{tot} = \sum_{j=j_0}^{J-1} \left(\sum_{i=i_j}^{I} E\left[C_{i,j}\middle|\mathcal{D}_I\right] \prod_{k=j+1}^{J-1} \widehat{f}_k^{CL}\right) \left(f_j - \widehat{f}_j^{CL}\right).$$

By by applying the estimation principle 3.6 we find

$$\widehat{EE}_{tot} = \sum_{j=i_0}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{\left(\sum_{i=i_j}^{I} \widehat{C}_{i,J}^{CL}\right)^2}{\left(\sum_{i=i_0}^{i=i_j-1} w_{i,j}\right)}. \tag{43}$$

Remarks:

– Note that

$$\widehat{EE}_{tot} > \sum_{i=0}^{I} \widehat{EE}_i = \sum_{j=i_0}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{\sum_{i=i_j}^{I} \left(\widehat{C}_{i,J}^{CL}\right)^2}{\left(\sum_{i=i_0}^{i=i_j-1} w_{i,j}\right)}.$$

c) **result**
The results in the following Theorem except formula (46) are a summary of (30) , (35) , (40) , (42) , (43) .

Theorem 4.1 *The msep of the ultimate run-off risk can be estimated by*

a) single accident year i

$$\widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\frac{1}{w_{i,j}} + \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}} \right) \right\} \tag{44}$$

where the weights $w_{i,j}$ are defined in (4) ;

b) total over all accident years

$$\widehat{msep}_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right) = \sum_{j=j_0}^{J-1} \left\{ \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\sum_{i=i_j}^{I} \frac{\left(\widehat{C}_{i,J}^{CL}\right)^2}{w_{i,j}} + \frac{\left(\sum_{i=i_j}^{I} \widehat{C}_{i,J}^{CL}\right)^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}} \right) \right\} \tag{45}$$

$$= \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left(\frac{\sum_{i=i_j}^{I} \widehat{C}_{i,J}^{CL}}{\widehat{C}_{tot,J}^{CL}} \right) \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}}, \tag{46}$$

where

$$\widehat{C}_{tot,J}^{CL} = \sum_{i=i_0}^{I} \widehat{C}_{i,J}^{CL}.$$

Remarks:

– The first summand in (44) and (45) represent the process variance and the second one the estimation error.

– (44) is the same formula as the Mack formula (95) in appendix A. However, (45) and (46) look different to the Mack formula (96) in appendix A. The covariance terms have disappeared, they are much simpler and have an intuitively understandable interpretation (see the second last bullet point of these remarks). But they give the same results as (96) .

– Formula (46) was already found by Ancus Röhr in [12]. But his derivation is less stringent, as he did not calculate the msep of the ultimate claims development result CDR_{ult}, but the msep of $\widetilde{CDR}_{ult}$, where $\widetilde{CDR}_{ult}$ is a first order Taylor expansion of CDR_{ult}.

– For the process error the uncertainties due to the $F_{i,j}$ sum up, whereas for the estimation error several accident years are affected simultaneously by the uncertainty of $\widehat{f}_j^{CL}$. This is the reason why

$$\widehat{PV}_{tot} = \sum_{i=i_0}^{I} \widehat{PV}_i, \text{ but } \widehat{EE}_{tot} > \sum_{i=i_0}^{I} \widehat{EE}_i.$$

– **intuitively accessible interpretation**
From (36) and (29) we see that (44) and (45) can be written as

$$\widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \sum_{j=j_i}^{J-1} \left(\widehat{\mathrm{CoV}}\left(F_{i,j}|\,\mathcal{B}_j\right)^2 + \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL}\middle|\,\mathcal{B}_j\right)^2 \right) \right\}, \tag{47}$$

$$\widehat{msep}_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left(\sum\nolimits_{j=j_0}^{J-1} q_j \left(\widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL}\middle|\,\mathcal{B}_j\right)\right)^2 \right), \tag{48}$$

$$\text{where } q_j = \frac{\sum_{i=i_j}^{I} \widehat{C}_{i,J}^{CL}}{\widehat{C}_{tot,J}^{CL}} \text{ is the fraction of } \widehat{C}_{tot,J}^{CL}\text{, which is affected by the uncertainty of } \widehat{f}_j^{CL}. \tag{49}$$

(47) and (48) give a good intuitive understanding of (44) and (46), as the coefficients of variation $\mathrm{CoV}\left(F_{i,j}|\,\mathcal{B}_j\right)$ and $\mathrm{CoV}\left(\widehat{f}_j^{CL}\middle|\,\mathcal{B}_j\right)$ are good intuitive measures for the relative deviation of $F_{i,j}$ and $\widehat{f}_j^{CL}$ from the "true" CL-factors f_j. In particular, (48) is a very intuitive formula.

– Instead of distinguishing between the process variance and the estimation error we could apply the telescope formula (25) directly to the ultimate run-off risk

$$\begin{aligned} Z_{i,ult} &= C_{i,J} - \widehat{C}_{i,J}^{CL} \\ &= w_{i,j_i}\left(\prod_{j=j_i}^{J-1} F_{i,j} - \prod_{j=j_i}^{J-1} \widehat{f}_j^{CL}\right) \\ &= w_{i,j_i}\sum_{j=j_i}^{J-1}\left\{\prod_{k=j_i}^{j-1} F_{i,k}\left(F_{i,j} - \widehat{f}_j^{CL}\right)\prod_{m=j+1}^{J-1}\widehat{f}_m^{CL}\right\} \\ &= \underbrace{\sum_{j=j_i}^{J-1} C_{i,j}\left(F_{i,j} - f_j\right)\prod_{m=j+1}^{J-1}\widehat{f}_m^{CL}}_{A_i} + \underbrace{\sum_{j=j_i}^{J-1} C_{i,j}\left(f_j - \widehat{f}_j^{CL}\right)\prod_{m=j+1}^{J-1}\widehat{f}_m^{CL}}_{B_i}. \end{aligned} \tag{50}$$

However the summands in (50) are correlated, such that the calculation of $E\left[Z_{i,ult}^2\middle|\,\mathcal{D}_I\right]$ is rather complicated. But note that the $C_{i,j}$ in B_i play the role of "stochastic weights", which are not yet known. A natural procedure is to replace them by the forecasted weights $w_{i,j} = \widehat{C}_{i,j}^{CL}$ and to consider

$$\widetilde{Z}_{i,ult} = \sum_{j=j_i}^{J-1} C_{i,j}\left(F_{i,j} - f_j\right) \prod_{m=j+1}^{J-1} \widehat{f}_m^{CL} + \sum_{j=j_i}^{J-1} w_{i,j}\left(f_j - \widehat{f}_j^{CL}\right) \prod_{m=j+1}^{J-1} \widehat{f}_m^{CL}. \tag{51}$$

Then we immediately get,

$$\widehat{E}\left[\widetilde{Z}_{i,ult}^2 \middle| \mathcal{D}_I\right] = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\left(\frac{1}{w_{i,j}} + \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}}\right)\right\},$$

which is the same result as the one found in Theorem 4.1. Hence by replacing the not yet known stochastic weights in B_i by the forecasted weights $w_{i,j}$ and then calculating the msep of the modified r.v. $\widetilde{Z}_{i,ult}$ we end up with the "correct" estimator (44).

Proof of Theorem 4.1:
It only remains to prove that (45) can be expressed by (46).

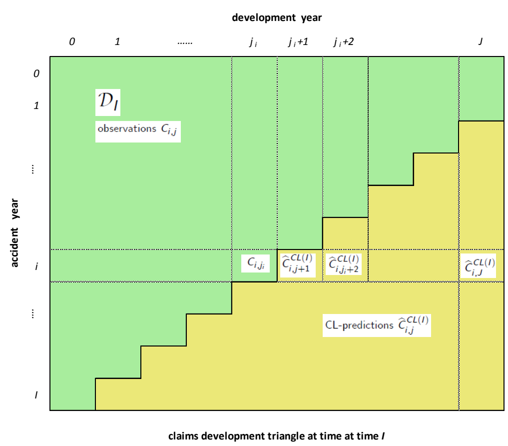
$$\begin{aligned}
&\sum_{j=j_0}^{J-1}\left\{\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\left(\sum_{i=i_j}^{I}\frac{\left(\widehat{C}_{i,J}^{CL}\right)^2}{w_{i,j}} + \frac{\left(\sum_{i=i_j}^{I}\widehat{C}_{i,J}^{CL}\right)^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}\right)\right\} = \\
&= \sum_{j=j_0}^{J-1}\left\{\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\left(\prod_{k=j}^{J-1}\widehat{f}_k^{CL}\right)\left(\left(\sum_{i=i_j}^{I}\widehat{C}_{i,J}^{CL}\right)\left(1+\frac{\sum_{i=i_j}^{I}\widehat{C}_{i,J}^{CL}}{\sum_{i=i_0}^{i_j-1}\widehat{C}_{i,J}^{CL}}\right)\right)\right\} \\
&= \sum_{j=j_0}^{J-1}\left\{\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\left(\prod_{k=j}^{J-1}\widehat{f}_k^{CL}\right)\left(\sum_{i=i_j}^{I}\widehat{C}_{i,J}^{CL}\right)\left(\frac{\widehat{C}_{tot,J}^{CL}}{\sum_{i=i_0}^{i_j-1}\widehat{C}_{i,J}^{CL}}\right)\right\} \\
&= \left(\widehat{C}_{tot,J}^{CL}\right)^2\left(\sum_{j=j_0}^{J-1}\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}}\frac{\sum_{i=i_j}^{I}\widehat{C}_{i,J}^{CL}}{\widehat{C}_{tot,J}^{CL}}\right).
\end{aligned}$$

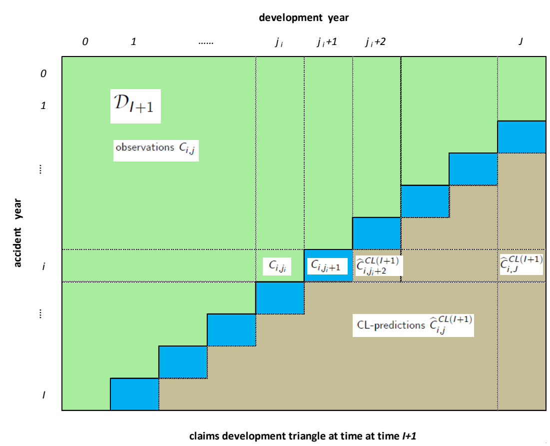
□

5 The one-year run-off prediction uncertainty of the next accounting year

In the new solvency regulation (SST, Solvency II), a time horizon of one year is considered. Therefore the reserve risk relevant for solvency purposes is the one-year run-off risk of the next accounting year.

In the previous section, claims reserving was considered from a static perspective. For solvency purposes, the claims reserving has to be seen as a dynamic process, where the predictions are updated based on new information which become available during the run-off process. At the end of the next accounting year $I+1$ there will be available the data $\mathcal{D}_{I+1}$. The CL-factors and the prediction of the ultimate claim will then be made on the basis of $\mathcal{D}_{I+1}$.





As in the previous sections we denote by $\widehat{f}_j^{CL}$ and $\widehat{C}_{i,j}^{CL}$ the CL factors and CL forecasts at time I. But for the future accounting year $I+1$ we will indicate by a superscript the time-point of the corresponding estimates, e.g. $\widehat{f}_j^{CL(I+1)}$ and $\widehat{C}_{i,J}^{CL(I+1)}$. Note that, conditional on $\mathcal{D}_I$, $\widehat{f}_j^{CL(I+1)}$ and $\widehat{C}_{i,J}^{CL(I+1)}$ are r.v., whereas $\widehat{f}_j^{CL}$ and $\widehat{C}_{i,j}^{CL}$ are given constants.

In the following we derive the msep of this one-year run-off risk strictly within the classical model of Mack.

a) single accounting year

The claims development result of accident year i in the P&L statement of the next accounting year $I+1$ is given by

$$CDR_i^{(I+1)} = \widehat{C}_{i,J}^{CL} - \widehat{C}_{i,J}^{CL(I+1)}$$

and

$$\begin{aligned} Z_i^{(I+1)} &= -CDR_i^{(I+1)} \\ &= w_{i,j_i} \left\{ F_{i,j_i} \prod_{j=j_i+1}^{J-1} \widehat{f}_j^{CL(I+1)} - \widehat{f}_{j_i}^{CL} \prod_{j=j_i+1}^{J-1} f_j^{CL} \right\}. \end{aligned} \tag{52}$$

With the telescope formula (25) we can write (52) as

$$Z_i^{(I+1)} = \underbrace{w_{i,j_i} \left\{ \left(F_{i,j_i} - \widehat{f}_{j_i}^{CL} \right) \prod_{j=j_i+1}^{J-1} \widehat{f}_j^{CL} \right\}}_{A_i} + \underbrace{w_{i,j_i} F_{i,j_i} \left(\prod_{j=j_i+1}^{J-1} \widehat{f}_j^{CL(I+1)} - \prod_{j=j_i+1}^{J-1} \widehat{f}_j^{CL)} \right)}_{B_i}.$$

By definition of best estimate reserves we forecast a CDR of 0. Hence the msep of the one-year run-off risk of the next accounting year is given by

$$\begin{aligned} msep_{CDR_i^{(I+1)} \mid \mathcal{D}_I}(0) &= E\left[\left(Z_i^{(I+1)} \right)^2 \middle| \mathcal{D}_I \right] \\ &= E\left[A_i^2 \middle| \mathcal{D}_I \right] + E\left[B_i^2 \middle| \mathcal{D}_I \right] + 2E\left[A_i B_i \middle| \mathcal{D}_I \right]. \end{aligned} \tag{53}$$

For the first summand in (53) we get

$$\begin{aligned} E\left[A_i^2\middle|\mathcal{D}_I\right] &= w_{i,j_i}^2\left\{E\left[(F_{i,j_i}-f_{j_i})^2\middle|\mathcal{D}_I\right]+\left(f_{j_i}-\widehat{f}_{j_i}^{CL}\right)^2\right\}\prod_{j=j_i+1}^{J-1}\left(\widehat{f}_j^{CL}\right)^2 \\ &= w_{i,j_i}\sigma_{j_i}^2\prod_{j=j_i+1}^{J-1}\left(\widehat{f}_j^{CL}\right)^2+w_{i,j_i}^2\left(f_{j_i}-\widehat{f}_{j_i}^{CL}\right)^2\prod_{j=j_i+1}^{J-1}\left(\widehat{f}_j^{CL}\right)^2, \end{aligned}$$

which, by use of the the estimation principle 3.6, is estimated by

$$\widehat{E}\left[A_i^2\middle|\mathcal{D}_I\right]=\left(\widehat{C}_{i,J}^{CL}\right)^2\frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\left\{\frac{1}{w_{i,j_i}}+\frac{1}{\sum_{k=i_0}^{i-1}w_{k,j_i}}\right\}. \tag{54}$$

For estimating the second and the third summand in (53) we first note that

$$\begin{aligned} \widehat{f}_j^{CL(I+1)}-\widehat{f}_j^{CL} &= \frac{\sum_{i=i_0}^{i_j}w_{i,j}F_{i,j}}{\sum_{i=i_0}^{i_j}w_{i,j}}-\frac{\sum_{i=i_0}^{i_j-1}w_{i,j}F_{i,j}}{\sum_{i=i_0}^{i_j-1}w_{i,j}} \\ &= a_j\left(F_{i_j,j}-\widehat{f}_j^{CL}\right), \end{aligned} \tag{55}$$

where

$$a_j=\frac{w_{i_j,j}}{\sum_{i=i_0}^{i_j}w_{i,j}}. \tag{56}$$

Since the observations of different accident years are independent it follows that

$$\left\{F_{i,j_i},\ \widehat{f}_{j_i+1}^{CL(I+1)},\ldots,\widehat{f}_{J-1}^{CL(I+1)}\right\}\quad\text{are independent,} \tag{57}$$

$$\left\{\widehat{f}_{j_0}^{CL(I+1)},\ \widehat{f}_{j_0+1}^{CL(I+1)},\ldots,\widehat{f}_{J-1}^{CL(I+1)}\right\}\quad\text{are independent.} \tag{58}$$

From the model assumptions 3.1 and from (55) and (57) follows that

$$\begin{aligned} E\left[F_{i,j_i}^2\middle|\mathcal{D}_I\right] &= f_{j_i}^2+\frac{\sigma_{j_i}^2}{w_{i,j_i}}, \\ E\left[\widehat{f}_j^{CL(I+1)}\middle|\mathcal{D}_I\right] &= \widehat{f}_j^{CL}+a_j\left(f_j-\widehat{f}_j^{CL}\right), \\ \operatorname{Var}\left(\widehat{f}_j^{CL(I+1)}\middle|\mathcal{D}_I\right) &= a_j^2\frac{\sigma_{j_i}^2}{w_{i,j_i}}, \\ E\left[\left(\widehat{f}_j^{CL(I+1)}\right)^2\middle|\mathcal{D}_I\right] &= \left(\widehat{f}_j^{CL}+a_j\left(f_j-\widehat{f}_j^{CL}\right)\right)^2+a_j^2\frac{\sigma_j^2}{w_{i,j_i}}, \end{aligned}$$

and, by applying the estimation principle 3.6,

$$\widehat{E}\left[F_{i,j_i}^2\middle|\mathcal{D}_I\right] = \left(\widehat{f}_{j_i}^{CL}\right)^2+\frac{\widehat{\sigma}_{j_i}^2}{w_{i,j_i}} \tag{59}$$

$$\widehat{E}\left[\widehat{f}_j^{CL(I+1)}\middle|\mathcal{D}_I\right] = \widehat{f}_j^{CL}, \tag{60}$$

$$\begin{aligned} \widehat{E}\left[\left(\widehat{f}_j^{CL(I+1)}\right)^2\middle|\mathcal{D}_I\right] &= \left(\widehat{f}_j^{CL}\right)^2+a_j^2\widehat{\sigma}_j^2\left(\frac{1}{\sum_{i=i_0}^{i_j-1}w_{i,j}}+\frac{1}{w_{i,j_i}}\right) \\ &= \left(\widehat{f}_j^{CL}\right)^2+b_j\widehat{\sigma}_j^2, \end{aligned} \tag{61}$$

where

$$b_j = \frac{w_{i_j,j}}{\left(\sum_{i=i_0}^{i_j-1} w_{i,j}\right)\left(\sum_{i=i_0}^{i_j} w_{i,j}\right)}. \tag{62}$$

With the estimation principle 3.6 we immediately obtain that

$$\widehat{E}\left[A_i B_i \middle| \mathcal{D}_I\right] = 0, \tag{63}$$

since

$$\widehat{E}\left[\left(\prod_{j=j_i+1}^{J-1} \widehat{f}_j^{CL(I+1)} - \prod_{j=j_i+1}^{J-1} f_j^{CL)}\right)\middle| \mathcal{D}_I\right] = 0.$$

From (57) and (59) follows that the second summand in (53) can be estimated by

$$\begin{aligned}\widehat{E}\left[B_i^2 \middle| \mathcal{D}_I\right] &= w_{i,j_i}^2 \left\{\left(\left(\widehat{f}_{j_i}^{CL}\right)^2 + \frac{\widehat{\sigma}_{j_i}^2}{w_{i,j_i}}\right)\left(\prod_{j=j_i+1}^{J-1}\left(\left(\widehat{f}_j^{CL}\right)^2 + b_j \widehat{\sigma}_j^2\right) - \prod_{j=j_i+1}^{J-1}\left(\widehat{f}_j^{CL}\right)^2\right)\right\} \\ &= w_{i,j_i}^2 \prod_{j=j_i}^{J-1}\left(\widehat{f}_j^{CL}\right)^2 \left\{\left(1 + \frac{1}{w_{i,j_i}}\frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\right)\left(\prod_{j=j_i+1}^{J-1}\left(1 + b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1\right)\right\}. \end{aligned} \tag{64}$$

By plugging (54) , (63) , (64) into (53) we get

$$\begin{aligned}\widehat{E}\left[\left(Z_i^{(I+1)}\right)^2 \middle| \mathcal{D}_I\right] = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\left\{\frac{1}{w_{i,j_i}} + \frac{1}{\sum_{k=i_0}^{i-1} w_{k,j_i}}\right\} + \\ + \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\left(1 + \frac{1}{w_{i,j_i}}\frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\right)\left(\prod_{j=j_i+1}^{J-1}\left(1 + b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1\right)\right\}.\end{aligned} \tag{65}$$

Remarks:

- From (60) and (61) we see that

$$b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} = \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1)} \middle| \mathcal{D}_I\right)^2, \tag{66}$$

which is an intuitively accessible interpretation.

- Note the difference between

$\widehat{\mathrm{CoV}}\left(\widehat{f}_{j_i}^{CL} \middle| \mathcal{B}_{j_i}\right)$ = measure, how much $\widehat{f}_{j_i}^{CL}$ will deviate from f_j, if only the data $\mathcal{B}_{j_i}$ are considered as known and non random

and

$\widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1)} \middle| \mathcal{D}_I\right)$ = measure how much $\widehat{f}_j^{CL(I+1)}$ will deviate from $\widehat{f}_j^{CL}$ if all data $\mathcal{D}_I$ are considered as known and non random.

b) **total over all accident years**

We complement the observed triangle by filling up the lower right not yet observed part with the CL-forecasts $\widehat{C}_{i,J}^{CL}$ and take the total over each column, that is we define

$$\widehat{C}_{tot,j}^{CL} := w_{tot,j=} \sum\nolimits_{i=i_0}^{I} w_{i,j},$$

where the weights $w_{i,j}$ are defined in (4) . Analogously we define

$$\widehat{C}_{tot,j}^{CL(I+1)} := \sum\nolimits_{i=i_0}^{I} w_{i,j}^{(I+1)},$$

where

$$w_{i,j}^{(I+1)} = \begin{cases} C_{i,J} & \text{if } C_{i,J} \in \mathcal{D}_I \\ \widehat{C}_{i,j}^{CL(I+1)} & \text{otherwise} \end{cases} .$$

By definition of $\widehat{f}_j^{CL}$ and $\widehat{f}_j^{CL(I+1)}$ and of the CL-forecasts it holds that

$$\begin{aligned} \widehat{C}_{tot,j}^{CL} &= w_{tot,j_0} \prod_{j=j_0}^{j-1} \widehat{f}_j^{CL}, \\ \widehat{C}_{tot,j}^{CL(I+1)} &= w_{tot,j_0} \prod_{j=j_0}^{j-1} \widehat{f}_j^{CL(I+1)}. \end{aligned}$$

The CDR of the next accounting year for the total over all accident year is given by

$$CDR_{tot}^{(I+1)} = \widehat{C}_{tot,J}^{CL} - \widehat{C}_{tot,J}^{CL(I+1)}$$

and

$$\begin{aligned} Z_{tot}^{(I+1)} &= -CDR_{tot}^{(I+1)} \\ &= w_{tot,j_0} \left(\prod_{j=j_0}^{J-1} \widehat{f}_j^{CL(I+1)} - \prod_{j=j_0}^{J-1} \widehat{f}_j^{CL} \right). \end{aligned}$$

From the independence property (58) and from (61) we immediately get

$$\widehat{E}\left[\left(Z_{tot}^{(I+1)} \right)^2 \middle| \mathcal{D}_I \right] = \left(\widehat{C}_{tot,J}^{CL} \right)^2 \left\{ \prod_{j=j_0}^{J-1} \left(1 + b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL} \right)^2} \right) - 1 \right\} \tag{67}$$

c) **result**

The following Theorem is a summary of (65) and (67) .

Theorem 5.1 *The msep of the one year run-off risk in the next accounting year I+1 can be estimated by*

a) single accident year

$$\widehat{msep}_{CDR_i^{(I+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2} \left\{\frac{1}{w_{i,j_i}} + \frac{1}{\sum_{k=i_0}^{i-1} w_{k,j_i}}\right\} +$$
$$+ \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\left(1 + \frac{1}{w_{i,j_i}} \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\right)\left(\prod_{j=j_i+1}^{J-1}\left(1 + b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1\right)\right\}, \tag{68}$$

where

$$b_j = \frac{w_{i_j,j}}{\left(\sum_{i=i_0}^{i_j-1} w_{i,j}\right)\left(\sum_{i=i_0}^{i_j} w_{i,j}\right)}, \tag{69}$$

$$w_{i,j} \text{ as defined in } (4)\,.$$

b) total over all accident years

$$\widehat{msep}_{CDR_{tot}^{(I+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{\prod_{j=j_0}^{J-1}\left(1 + b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1\right\}, \tag{70}$$

where b_j is given by (69) *.*

Remarks:

- The first summand in (68) reflects the risk that the observation on the next diagonal will deviate from the forecast at time I.
- The second summand in (68) reflects the risk of updating the forecasts of later development years due to an update of the estimated CL-factors from time I to time $I+1$.
- The estimator (70) is surprisingly simple and even simpler than the estimator (71) for a single accident year .
- **intuitively accessible interpretation**
  From (29), (36) and (66) we see that

$$\begin{aligned}
\frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\frac{1}{w_{i,j_i}} &= \widehat{\text{CoV}}\left(F_{i,j}\middle|\,\mathcal{B}_j\right)^2, \\
\frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2}\frac{1}{\sum_{k=i_0}^{i-1} w_{k,j_i}} &= \widehat{\text{CoV}}\left(\widehat{f}_j^{CL}\middle|\,\mathcal{B}_j\right)^2, \\
b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} &= \widehat{\text{CoV}}\left(\widehat{f}_j^{CL}\middle|\,\mathcal{D}_I\right)^2,
\end{aligned}$$

where the right hand side are easily understandable and intuitively accessible interpretations of the items on the left-hand side.

The following Theorem is obtained by taking a first order Taylor expansion of (68) and (70). It is a good approximation, if $b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} << 1$,which is the case in most practical situations.

Theorem 5.2 *The msep of the one year run-off risk in the next accounting year I+1 can alternatively (Taylor approximation) be estimated by*

a) single accident year

$$\widehat{msep}_{CDR_i^{(I+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL}\right)^2} \left\{ \left(\frac{1}{w_{i,j_i}} + \frac{1}{\sum_{k=i_0}^{i-1} w_{k,j_j}} \right) \right\} + \\ + \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \sum_{j=j_i+1}^{J-1} b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \right\} \tag{71}$$

where

$$b_j = \frac{w_{i_j,j}}{\left(\sum_{i=i_0}^{i_j-1} w_{i,j}\right)\left(\sum_{i=i_0}^{i_j} w_{i,j}\right)}, \tag{72}$$

$w_{i,j}$ *as defined in* (4).

b) total over all accident years

$$\widehat{msep}_{CDR_{tot}^{(I+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{ \sum_{j=j_0}^{J-1} b_j \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \right\}. \tag{73}$$

Remarks

- (71) is the Merz-Wüthrich formula (see (97) in appendix A). Hence the Merz-Wüthrich formula is obtained by a first order Taylor approximation from (68). In practical applications the numerical results of the two estimators are very often so close to each other that the difference is negligible for practical purposes.
- (73) can directly be compared with the corresponding Merz-Wüthrich formula (99) in appendix A. It looks much nicer, since all the covariance terms occurring in (99) have disappeared.

– **intuitively accessible interpretation**
Analogously as in Theorem 5.1 we can look behind the formulas and we can write (71) and (73) as

$$\widehat{msep}_{CDR_i^{(I+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\widehat{\mathrm{CoV}}\left(F_{i,j_i}\middle|\mathcal{B}_{j_i}\right)^2 + \widehat{\mathrm{CoV}}\left(\widehat{f}_{j_i}^{CL}\middle|\mathcal{B}_{j_i}\right)^2\right\} + \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\sum_{j=j_i+1}^{J-1} \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1)}\middle|\mathcal{D}_I\right)^2\right\}, \quad (74)$$

$$\widehat{msep}_{CDR_{tot}^{(I+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{\sum_{j=j_0}^{J-1} \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1)}\middle|\mathcal{D}_I\right)^2\right\}, \quad (75)$$

which are intuitively accessible interpretations of (71) and (73).

6 The one-year run-off prediction uncertainty in future accounting years

In the SST and in solvency II, the market value margin corresponding to the run-off risk equals the discounted cost of capital, which is needed at the end of each accounting year for the one year run-off risk in the next accounting year. For this purpose we need estimates of the one year run-off risk for the accounting years $I+k$, $k=1,\ldots,J-1$, evaluated at time I.

6.1 The Compatibility Condition

Note that the one-year CDR in the accounting years $I+k$, $k=1,\ldots,J$ are given by

$$CDR_i^{(I+k)} = \widehat{C}_{i,J}^{CL(I+k-1)} - \widehat{C}_{i,J}^{CL(I+k)},$$
$$CDR_{tot}^{(I+k)} = \widehat{C}_{tot,J}^{CL(I+k-1)} - \widehat{C}_{tot,J}^{CL(I+1+k)}.$$

Note also that accident year i is already fully developed in the accounting years $\{I+k : k \geq J-j_i\}$ and that therefore

$$\widehat{C}_{i,J}^{CL(I+k)} = C_{i,J} \quad \text{for } k \geq J-j_i.$$

The sum of the one year claims development results over all future development years is equal to the ultimate claims development result. Hence the one-year run-off risks $Z_i^{(I+k)} = -CDR_i^{(I+k)}$ satisfy

$$\sum_{k=1}^{J-j_i} Z_i^{(I+k)} = C_{i,J} - \widehat{C}_{i,J}^{CL}, \quad (76)$$
$$\sum_{k=1}^{J} Z_{tot}^{(I+k)} = C_{tot,J} - \widehat{C}_{tot,J}^{CL}, \quad (77)$$

and hence

$$E\left[\left(\sum_{k=1}^{J} Z_i^{(I+k)}\right)^2 \middle| \mathcal{D}_I\right] = E\left[\left(\widehat{C}_{i,J}^{CL} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] = msep_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right), \quad (78)$$
$$E\left[\left(\sum_{k=1}^{J} Z_{tot}^{(I+k)}\right)^2 \middle| \mathcal{D}_I\right] = E\left[\left(\widehat{C}_{tot,J}^{CL} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] = msep_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right). \quad (79)$$

By definition of best estimate reserves the forecast of the claims development result in future periods is always zero. For this reason it is often argued that the process of best estimate forecasts $\left\{\widehat{C}_{i,J}^{BE(I+k)} : k=0,\dots,J\right\}$ is a martingale and that therefore the one-year CDR, which are the increments of this process, are independent. Based on this martingale argument it is then required that the estimators of the one-year run-off risk should satisfy the following "splitting" property:

Definition 6.1 (splitting property) *Estimators of the msep of the one year run-off risk fulfil the "splitting property" if*

$$\sum_{k=1}^{J} \widehat{msep}_{CDR_i^{(I+k)}|\mathcal{D}_I}(0) = \widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) \quad \text{for } i=i_0,\dots,I \quad , \tag{80}$$

$$\sum_{k=1}^{J} \widehat{msep}_{CDR_{tot}^{(I+k)}|\mathcal{D}_I}(0) = \widehat{msep}_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right), \tag{81}$$

where the right hand side of (80) *and* (81) *are given by Theorem 4.1.*

Remarks:

- The "splitting property" means that the ultimate run-off prediction uncertainty is split over all future accounting years until final development.

However best estimate forecasts are usually not a martingale, which is also the case for the CL-forecasts. The CL forecasts fulfil

$$\widehat{E}\left[\widehat{C}_{i,J}^{CL(I+k+1)} \middle| \widehat{C}_{i,J}^{CL(I+k)}\right] = \widehat{C}_{i,J}^{CL(I+k)},$$

but they do not satisfy the martingale condition

$$E\left[\widehat{C}_{i,J}^{CL(I+k+1)} \middle| \widehat{C}_{i,J}^{CL(I+k)}\right] = \widehat{C}_{i,J}^{CL(I+k)},$$

because the unknown CL factors f_j are replaced in the CL-forecasts by its estimates $\widehat{f}_j^{CL}$ and $\widehat{f}_j^{CL(I+1+k)}$respectively.

Assume, that the required capital (risk margin) for the run-off risk is calculated by a fixed percentage of the msep (variance risk measure) and that we consider only nominal cash-flows (no interest income). The requirement, to have enough capital for the ultimate run-off risk is a weaker condition than the requirement, that this capital is also sufficient to meet, from a today's perspective, the capital requirement for the one-year run-off risk at the beginning of each future accounting year until final claims development. Therefore the msep for the ultimate run-off risk is a lower bound for the sum of the msep of the one-year run-off risk. For this reason the estimators for the one-year run-off risk should satisfy the following "compatibility" condition:

Condition 6.2 (compatibility) *Estimators of the msep of the one year run-off risk fulfil the "compatibility condition" if for $i=i_0,\dots,I$ and for the total over all accident years it holds that*

$$\sum_{k=0}^{J} \widehat{msep}_{CDR_i^{(I+k)}|\mathcal{D}_I}(0) > \widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) \quad \text{for } i=i_0,\dots,I, \tag{82}$$

$$\sum_{k=0}^{J} \widehat{msep}_{CDR_{tot}^{(I+k)}|\mathcal{D}_I}(0) > \widehat{msep}_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right), \tag{83}$$

where the right hand side of (82) *and* (83) *are given by Theorem 4.1.*

6.2 Estimators of the msep of the one-year run-off risk in future accounting years

It holds that

$$\begin{aligned} msep_{CDR_i^{(I+k+1)}|\mathcal{D}_I}(0) &= E\left[\left(Z_i^{(I+k+1)}\right)^2\middle|\mathcal{D}_I\right] \\ &= E\left[E\left[\left(Z_i^{(I+k+1)}\right)^2\middle|\mathcal{D}_{I+k}\right]\middle|\mathcal{D}_I\right], \qquad (84) \\ msep_{CDR_{tot}^{(I+k+1)}|\mathcal{D}_I}(0) &= E\left[\left(Z_{tot}^{(I+k+1)}\right)^2\middle|\mathcal{D}_I\right] \\ &= E\left[E\left[\left(Z_{tot}^{(I+k+1)}\right)^2\middle|\mathcal{D}_{I+k}\right]\middle|\mathcal{D}_I\right]. \qquad (85) \end{aligned}$$

The following estimator of the inner expected value of (84) is immediately obtained with Theorem 5.1.

$$\begin{aligned} \widehat{msep}_{CDR_i^{(I+k+1)}|\mathcal{D}_{I+k}}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL}\right)^2}\left\{\frac{1}{C_{i,j_i+k}} + \frac{1}{\sum_{l=i_0}^{i-1} C_{l,j_i+k}}\right\} + \qquad (86) \\ + \left(\widehat{C}_{i,J}^{CL}\right)^2\left\{\left(1 + B_{j_i+k}^{(I+k)}\frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL}\right)^2}\right)\left(\prod_{j=j_i+k+1}^{J-1}\left(1 + B_j^{(I+k)}\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1\right)\right\}, \end{aligned}$$

where

$$\begin{aligned} B_j^{(I+k)} &= \frac{C_{i_j+k,j}}{\left(\sum_{i=i_0}^{i_j+k-1} C_{i,j}\right)\left(\sum_{i=i_0}^{i_j+k} C_{i,j}\right)}, \qquad (87) \\ \widehat{f}_j^{CL(I+k)} &= \sum_{i=i_0}^{i_j+k-1}\frac{C_{i,j}}{C_{\bullet,j}}F_{i,j}, \quad \text{where } C_{\bullet,j} = \sum_{i=i_0}^{i_j+k-1} w_{i,j}, \qquad (88) \\ \widehat{C}_{i,J}^{CL(I+k)} &= C_{i,j_i+k}\prod_{j=j_i+k}^{J-1}\widehat{f}_j^{CL(I+k)}. \qquad (89) \end{aligned}$$

Note that the $C_{i,j}$ appearing in (86) play the role of a "weight". The only difference to (68) is that some of these "weights" are "stochastic weights" and not exactly known at time I. A natural procedure is to replace them with the weights $w_{i,j}$ defined in (4), which are either the already known weights $C_{i,j} \in \mathcal{D}_I$ or the forecasts $\widehat{C}_{i,j}^{CL}$ at time I. As mentioned in the remarks to Theorem 4.1 on page 13 in the last bullet point, this procedure has given there the correct Mack-result.

We do the same here and check then whether the resulting estimators fulfil the compatibility condition 6.2.

Theorem 6.3 *The msep of the one-year run-off risk in future accounting years can be estimated by*

i) single accident year i, accounting years $I+k+1$, $k=0,\ldots,J-j_i-1$

$$\widehat{msep}_{CDR_i^{(I+k+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL)}\right)^2} \left\{ \frac{1}{w_{i,j_i+k}} + \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j_i+k}} \right\} + \tag{90}$$

$$+ \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \left(1 + \frac{1}{w_{i,j_i+k}} \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL}\right)^2}\right) \left(\prod_{j=j_i+k+1}^{J-1} \left(1 + b_j^{(I+k)} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1 \right) \right\}$$

where

$$b_j^{(I+k)} = \frac{w_{i_j+k,j}}{\left(\sum_{i=i_0}^{i_j+k-1} w_{i,j}\right)\left(\sum_{i=i_0}^{i_j+k} w_{i,j}\right)}, \tag{91}$$

weights $w_{i,j}$ as defined in (4)*.*

ii) total over all accident years, accounting years $I+k+1$, $k=0,\ldots,J-1$

$$\widehat{msep}_{CDR_{tot}^{(I+k+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{ \prod_{j=j_i+k}^{J-1} \left(1 + b_j^{(I+k)} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2}\right) - 1 \right\}. \tag{92}$$

iii) The estimators (90) *and* (92) *fulfil the compatibility condition.*

Remarks:

- For $k=0$ (90) and (92) are of course the same as (68) and (70) in Theorem 5.1.
- The msep for accounting years with $k>0$ can also be calculated by filling up the next k diagonals with the chain ladder forecasts to get an arificial new data set $\widetilde{\mathcal{D}}_{I+k}$ and by applying Theorem (70) on the remaining triangle in $\widetilde{\mathcal{D}}_{I+k}$, but by keeping the estimators $\widehat{\sigma}_j^2$ from the original triangle $\mathcal{D}_I$.
- (92) is again a surprisingly simple formula.
- **intuitively accessible interpretation**
  Analogously as in Theorem 5.1 the expressions appearing in Theorem 6.3 can be interpreted in the following easily understandable and intuitively accessible way:

$$\frac{1}{w_{i,j_i+k}} \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL)}\right)^2} = \widehat{\mathrm{CoV}}\left(F_{i,j_i+k} \middle| \mathcal{D}_{I+k}\right)^2,$$

$$\frac{1}{\sum_{l=i_0}^{i-1} w_{l,j_i+k}} \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL)}\right)^2} = \widehat{\mathrm{CoV}}\left(\widehat{f}_{j_i+k}^{CL} \middle| \mathcal{B}_{j+k}\right)^2,$$

$$b_j^{(I+k)} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} = \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1)} \middle| \mathcal{D}_{I+k}\right)^2.$$

Proof of Theoerm 6.3:
The derivation of the estimators in i) was already explained. ii) is obtained analogously. iii) follows from the next Theorem 6.4 and the fact, that the estimators (90) and (92) are greater than the estimators (93) and (94). □

As in section 5 the estimators in Theorem 6.3 can again be approximated by a first Taylor approximation.

Theorem 6.4 *The msep of the one-year run-off risk in future accounting years can alternatively (Taylor approximation) be estimated by*

i) single accident year i, accounting years $I+k+1, \quad k=0,\dots,J-j_i-1$

$$\widehat{msep}_{CDR_i^{(I+k+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}^2_{j_{i+k}}}{\left(\widehat{f}^{CL}_{j_{i+k}}\right)^2} \left\{\left(\frac{1}{w_{i,j_i+k}} + \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j_i+k}}\right)\right\} + \\ + \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{\sum_{j=j_i+k+1}^{J-1} b_j^{(I+k)} \frac{\widehat{\sigma}^2_j}{\left(\widehat{f}^{CL}_j\right)^2}\right\} \qquad (93)$$

where $b_j^{(I+k)}$ *are given by* (91)*.*

ii) total over all accident years, accounting years $I+k+1, \quad k=0,\dots,J-1$

$$\widehat{msep}_{CDR_{tot}^{(I+k+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{\sum_{j=j_0+k}^{J-1} b_j^{(I+k)} \frac{\widehat{\sigma}^2_j}{\left(\widehat{f}^{CL}_j\right)^2}\right\} \qquad (94)$$

where $b_j^{(I+k)}$ *is given by* (91)*.*

iii) the estimators (93) *and* (94) *fulfil the splitting property.*

Remarks:

- The first bullet point in the remarks after Theorem 6.3 is analogously valid here.
- (93) and (94) are different to the Merz-Wüthrich formulas (100) and (101). However they give the same numerical results. This cannot be a pure coincidence. Thus they must be same just written differently.
- The first summand in (93) reflects again the risk that the observation on the next diagonal will deviate from the forecast at time I, whereas the second summand entails the risk of updating the forecasts of the later development years. The first summand looks similar to the one in Merz-Wüthrich, but note, that there is a difference even in this first summand.
- (94) is again a surprisingly simple formula compared to the rather complicated expression with all the covariance terms in (101).

- **intuitively accessible interpretation**
  Analogously as in Theorem 5.1 (93) and (94) can be written as

$$\widehat{msep}_{CDR_i^{(I+k+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \left(\widehat{\mathrm{CoV}}\left(F_{i,j_i+k}\middle|\, \mathcal{B}_{j_i+k}\right)\right)^2 + \widehat{\mathrm{CoV}}\left(\widehat{f}_{j_i+k}^{CL}\middle|\, \mathcal{B}_{j_i+k}\right)^2 \right\} + $$
$$+ \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \sum_{j=j_i+k+1}^{J-1} \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+1+k)}\middle|\, \mathcal{D}_{I+k}\right)^2 \right\},$$

$$\widehat{msep}_{CDR_{tot}^{(I+k+1)}\big|\mathcal{D}_I}(0) = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{ \sum_{j=j_i+k}^{J-1} \widehat{\mathrm{CoV}}\left(\widehat{f}_j^{CL(I+k+1)}\middle|\, \mathcal{D}_{I+k}\right)^2 \right\},$$

  which are intuitively accessible and understandable formulas.

Proof of Theorem 6.4:
(93) and (94) are obtained by a straightforward first order Taylor approximation from (90) and (92) .The proof of the splitting property iii) is given in appendix D □

7 Numerical example

Table 1 shows in the upper left part the data of a claims development triangle of medical costs in accident insurance and in the lower right part the chain-ladder forecasts. Column 20 contains the CL-forecasts of the ultimate claims (in red) and the corresponding chain-ladder reserves. The total reserves at the end of 2010 for this line of business amounts to CHF 66.697 mio.

Table 1: triangel of cumulative payments in CHF 1'000 and CL-forecasts

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Reserves
1984	1'324	2'230	2'373	2'465	2'505	2'559	2'593	2'683	2'755	2'849	2'968	2'999	3'096	3'333	3'420	3'532	3'678	3'724	3'797	3'875	3'966	
1985	1'430	2'796	3'042	3'136	3'231	3'400	3'526	3'582	3'872	3'905	3'966	4'064	4'388	4'586	4'914	5'394	5'495	5'597	5'719	5'824	6'003	
1986	1'612	2'970	3'244	3'352	3'412	3'469	3'568	3'618	3'659	3'680	3'698	3'729	3'751	3'787	3'793	3'799	3'815	3'824	3'837	3'851	3'858	
1987	2'075	3'458	3'724	3'820	3'864	3'938	3'974	3'997	4'017	4'048	4'059	4'100	4'120	4'127	4'144	4'157	4'201	4'222	4'233	4'246	4'253	
1988	2'673	4'457	4'866	4'960	5'055	5'087	5'135	5'171	5'258	5'407	5'431	5'451	5'465	5'485	5'497	5'507	5'519	5'541	5'568	5'582	5'604	
1989	2'918	4'730	5'138	5'302	5'359	5'437	5'501	5'550	5'569	5'600	5'605	5'608	5'610	5'653	5'672	5'703	5'742	5'771	5'776	5'794	5'804	
1990	3'052	4'900	5'371	5'600	5'734	5'812	5'865	5'932	5'980	6'058	6'118	6'141	6'217	6'264	6'309	6'354	6'391	6'408	6'429	6'447	6'501	
1991	2'649	4'340	4'805	4'995	5'097	5'163	5'315	5'375	5'500	5'586	5'635	5'697	5'731	5'758	5'804	5'834	5'859	5'880	5'893	5'900	**5'961**	61
1992	2'779	4'834	5'283	5'498	5'629	5'670	5'738	5'808	5'835	5'935	6'028	6'059	6'092	6'111	6'143	6'169	6'186	6'200	6'216	6'256	**6'321**	105
1993	2'492	4'344	4'782	4'968	5'033	5'107	5'191	5'230	5'266	5'349	5'388	5'399	5'431	5'459	5'466	5'471	5'475	5'477	5'512	5'548	**5'605**	128
1994	3'026	5'403	5'820	6'063	6'137	6'197	6'254	6'512	6'610	6'750	6'851	6'910	6'977	7'059	7'156	7'217	7'263	7'302	7'348	7'396	**7'473**	210
1995	4'154	7'574	8'418	8'836	9'099	9'276	9'440	9'591	9'825	10'038	10'187	10'361	10'479	10'564	10'647	10'776	10'865	10'923	10'993	11'065	**11'179**	403
1996	3'490	6'266	7'052	7'375	7'505	7'646	7'746	7'809	7'860	7'894	7'922	7'947	7'976	8'046	8'110	8'221	8'289	8'334	8'387	8'442	**8'529**	419
1997	3'557	7'089	7'812	8'065	8'303	8'445	8'542	8'616	8'697	8'744	8'816	8'874	8'946	9'008	9'108	9'233	9'309	9'359	9'419	9'481	**9'579**	570
1998	4'742	8'819	9'821	10'355	10'662	11'000	11'288	11'509	11'650	11'771	11'890	11'983	12'087	12'225	12'361	12'530	12'633	12'701	12'783	12'866	**12'999**	912
1999	6'508	11'826	13'199	13'889	14'277	14'574	14'855	15'140	15'332	15'469	15'628	15'758	15'930	16'112	16'291	16'514	16'650	16'740	16'847	16'956	**17'132**	1'374
2000	6'708	13'382	15'155	15'797	16'130	16'389	16'658	16'865	17'040	17'205	17'321	17'461	17'652	17'853	18'051	18'299	18'450	18'549	18'668	18'789	**18'983**	1'662
2001	6'283	11'983	13'552	14'280	14'625	14'956	15'180	15'281	15'364	15'474	15'624	15'750	15'922	16'104	16'283	16'506	16'642	16'732	16'839	16'948	**17'123**	1'649
2002	6'297	12'810	14'166	14'883	15'326	15'568	15'731	15'973	16'062	16'253	16'411	16'543	16'724	16'915	17'103	17'337	17'480	17'574	17'687	17'802	**17'986**	1'924
2003	6'369	12'594	14'450	15'191	15'561	15'902	16'050	16'186	16'386	16'581	16'742	16'878	17'062	17'257	17'448	17'687	17'833	17'929	18'044	18'161	**18'349**	2'163
2004	7'735	15'339	17'274	18'177	18'783	19'247	19'584	19'850	20'096	20'335	20'532	20'698	20'925	21'163	21'398	21'691	21'870	21'988	22'129	22'272	**22'503**	2'918
2005	9'022	17'415	19'896	21'148	22'027	22'522	22'874	23'184	23'471	23'751	23'981	24'175	24'439	24'718	24'992	25'335	25'544	25'681	25'846	26'014	**26'283**	3'761
2006	10'311	21'215	24'530	26'123	27'200	27'736	28'170	28'551	28'905	29'250	29'533	29'772	30'098	30'441	30'778	31'201	31'458	31'627	31'829	32'036	**32'368**	5'168
2007	10'945	21'346	23'646	24'651	25'341	25'841	26'245	26'600	26'930	27'251	27'515	27'738	28'041	28'361	28'675	29'069	29'308	29'466	29'654	29'847	**30'156**	5'505
2008	12'073	22'274	25'170	26'390	27'129	27'664	28'096	28'477	28'830	29'173	29'456	29'694	30'019	30'361	30'698	31'119	31'376	31'544	31'746	31'953	**32'283**	7'113
2009	10'667	21'295	23'857	25'013	25'714	26'221	26'631	26'991	27'326	27'652	27'920	28'145	28'453	28'778	29'097	29'496	29'739	29'899	30'090	30'286	**30'599**	9'304
2010	12'385	23'476	26'299	27'574	28'346	28'905	29'357	29'755	30'123	30'483	30'778	31'027	31'366	31'724	32'076	32'516	32'784	32'960	33'171	33'387	**33'732**	21'347
Total																						**66'697**
f_j^{CL}	1.896	1.120	1.048	1.028	1.020	1.016	1.014	1.012	1.012	1.010	1.008	1.011	1.011	1.011	1.014	1.008	1.005	1.006	1.007	1.010		

Table 2 shows the square root of the estimated msep for the ultimate run-off and for the one year run-off in the next accounting year where the latter is calculated with the "exact" estimator according to Theorem 5.1 as well as with the Taylor approximation according to Theorem 5.2. We see, that in this example the numerical results of the two estimators differ not more than in the second digit after the decimal point and that the results with the "exact" estimator are greater, but only very minor in the second digit after the decimal point.

Table 2

accident year	**Reserves**	**ultimate run-off** msep$^{1/2}$	in % reserves	**one year run-off** next accounting year msep$^{1/2}$ exact	Taylor appr	in % reserves
1991	61	71	116%	70.74	70.74	116%
1992	105	87	82%	47.58	47.58	45%
1993	128	92	72%	45.87	45.87	36%
1994	210	115	55%	40.51	40.51	19%
1995	403	169	42%	88.48	88.48	22%
1996	419	238	57%	190.98	190.98	46%
1997	570	289	51%	139.94	139.94	25%
1998	912	378	41%	163.51	163.51	18%
1999	1'374	482	35%	198.78	198.78	14%
2000	1'662	517	31%	106.76	106.76	6%
2001	1'649	493	30%	110.51	110.51	7%
2002	1'924	516	27%	120.35	120.35	6%
2003	2'163	549	25%	187.36	187.36	9%
2004	2'918	632	22%	155.02	155.02	5%
2005	3'761	703	19%	160.31	160.31	4%
2006	5'168	814	16%	201.54	201.54	4%
2007	5'505	798	14%	224.48	224.48	4%
2008	7'113	862	12%	265.29	265.29	4%
2009	9'304	930	10%	437.82	437.81	5%
2010	21'347	1'795	8%	1'507.37	1'507.36	7%
Total	**66'697**	5'033	8%	2'435.88	2'435.86	4%

Table 3 shows the results obtained for the square-root of the one-year msep in future accounting years until final development together with the ingoing reserves. The numerical results obtained with the "exact" formula and with the Taylor approximation are again practically the same with differences only in the second digit after the decimal point. From the table we can also the splitting property of the Taylor-approximation.

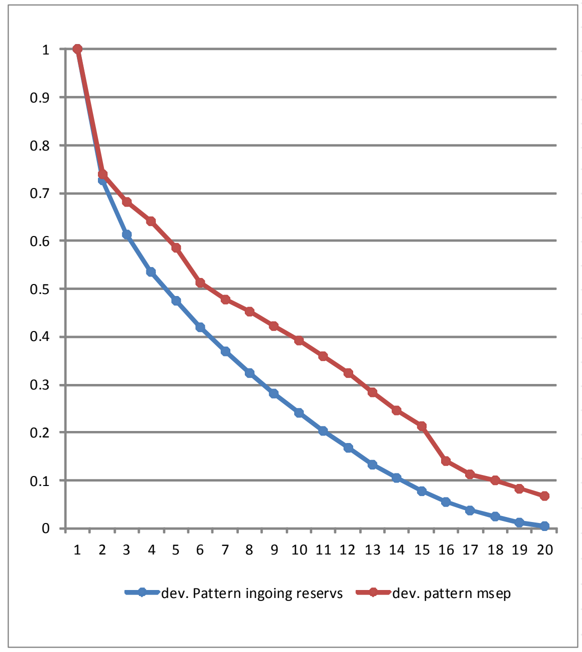
In the current formula for calculating the risk margin in solvency II it assumed that the required capital for the remaining one-year run-off risk in future accounting years decreases proportionally to the remaining reserves. This was due to the lack of formulas to calculate the prediction uncertainty for future accounting years. But now the formulas have been developed and are here. Comparing the development pattern of the reserves with the development pattern of the msep we see, that the latter decreases much slower. This is not a surprise. Complex and complicated claims such as severe bodily injury claims stay open for a long time, whereas "normal" claims can be settled much quicker. Hence the proportion of the reserves stemming from complex claims is bigger in later development years. But the prediction uncertainty of this kind of claims is bigger as for the "normal" claims. But it also means that one will need more capital in solvency II with this new formulas, since the risk margin will become bigger.

Table 3

total over all accident years

accounting year	**ingoing Reserves** in 1'000	dev. pattern	**one year run-off** msep$^{1/2}$ "exact"	Taylor appr.	dev.pattern
2011	66'697	100%	2'435.88	2'435.86	100%
2'012	48'513	73%	1'801.67	1'801.67	74%
2'013	40'919	61%	1'661.06	1'661.05	68%
2'014	35'786	54%	1'564.28	1'564.27	64%
2'015	31'614	47%	1'426.15	1'426.14	59%
2'016	27'960	42%	1'250.72	1'250.71	51%
2'017	24'694	37%	1'163.14	1'163.14	48%
2'018	21'662	32%	1'099.81	1'099.81	45%
2'019	18'791	28%	1'027.23	1'027.23	42%
2'020	16'067	24%	953.60	953.60	39%
2'021	13'531	20%	874.67	874.67	36%
2'022	11'164	17%	788.65	788.65	32%
2'023	8'933	13%	692.48	692.48	28%
2'024	6'930	10%	602.20	602.20	25%
2'025	5'164	7.7%	518.85	518.85	21%
2'026	3'648	5.5%	341.16	341.16	14%
2'027	2'494	3.7%	274.70	274.70	11%
2'028	1'611	2.4%	244.81	244.81	10%
2'029	874	1.3%	198.87	198.87	8%
2'030	345	0.5%	162.87	162.87	7%
			sum 1-year msep	**25'326'904**	
			ultimate msep	**25'326'904**	

development pattern reserves and msep



Acknowledgement

I am grateful to Ancus Röhr. Discussions with him on earlier drafts of his paper [12] have motivated me to work further on this topic. The result is the present paper. I am also grateful to Hans Bühlmann. He told me that Lemma 3.5 is called the telescope-formula and how it can be proved best. Finally I would like to thank Patrick Helbling, who had done most of the calculations of the numerical example.

References

[1] Buchwalder M., Bühlmann, H., Merz M., Wüthrich M. (2006). The Mean Square Error of Prediction in the Chain Ladder Reserving Method-Final Remark 36/2, 521-542.

[2] Buchwalder M., Bühlmann, H., Merz M., Wüthrich M. (2006). The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited).ASTIN Bulletin 36/2, 553.

[3] Bühlmann, H., De Felice M., Gisler, A., Moriconi F., Wüthrich, M.V. (2009). Recursive Credibility Formula for Chain Ladder Factors and the Claims Development Result. ASTIN Bulletin 39/1,275-306.

[4] Gisler, A. (2006). The estimation error in the chain-ladder reserving method: a Bayesian approach. ASTIN Bulletin 36/2, 554-565.

[5] Gisler, A. , Wüthrich M.V.

[6] (2008). Credibility for the chain-ladder reserving method.38/2, 565-600.

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

[8] Mack, T., Quarg B., Braun C. (2006).The Mean Square Error of Prediction in the Chain Ladder Reserving Method- A Comment. ASTIN Bulletin 36/2, 523-552.

[9] Merz M., Wüthrich M.V. (2008). Stochastic Claims Reserving Methods in Insurance, Wiley.

[10] Merz M., Wüthrich M.V. (2008). Modelling the claims development result for solvency purposes. CAS Forum, Fall 2008, 542-568.

[11] Merz M., Wüthrich M.V. (2014). Claims run-off uncertainty: the full picture, SSRN manuscript 2524352.

[12] Roehr A. (2016). Chain ladder and error propagation, to appear in the next ASTIN Bulletin.

Appendices

A Formulae of Mack and Merz-Wüthrich

A.1 Formula of Mack

Written in the notation of this paper Mack [7] has found the following formulas for estimating the total run-off uncertainties for the CL reserves.

Theorem A.1 (Mack) *The msep can be estimated by*

i) single accident year i

$$\widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\frac{1}{w_{i,j}} + \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}}\right) \tag{95}$$

where $\widehat{C}_{i,j}^{CL}$ *are the CL-forecasts and where* $w_{i,j}$ *are as defined in* (4) *.*

ii) total over all accident years

$$\widehat{msep}_{C_{tot,J}|\mathcal{D}_I}\left(\widehat{C}_{tot,J}^{CL}\right) = \sum_{i=i_J+1}^{I} \widehat{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) \tag{96}$$

$$+ 2 \sum_{i=i_J+1}^{I} \widehat{C}_{i,J}^{CL} \left(\sum_{k=i+1}^{I} \widehat{C}_{k,J}^{CL}\right) \sum_{j=j_i}^{J-1} \left(\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{m=i_0}^{i_j-1} w_{m,j}}\right).$$

A.2 Formulae of Merz-Wüthrich

In [10] Merz and Wüthrich found the following formula for the msep of the one year run-off risk in the next accounting year (see formulas (1.2) and (2.3) in [11]).

Theorem A.2 *The msep of the one-year run-off risk in the next accounting year can be estimated by*

i) single accident year i

$$\begin{aligned}\widehat{msep}_{CDR_{i,I+1}|\mathcal{D}_I}(0) = & \left(\widehat{C}_{i,J}^{CL(I)}\right)^2 \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL(I)}\right)^2} \left(\frac{1}{w_{i,j_i}} + \frac{1}{\sum_{k=i_0}^{i_j-1} w_{k,j_i}}\right) \\ & + \left(\widehat{C}_{i,J}^{CL(I)}\right)^2 \sum_{j=j_i+1}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL(I)}\right)^2} \left(\alpha_j^{(I)} \frac{1}{\sum_{k=i_0}^{i_j-1} w_{k,j}}\right),\end{aligned} \tag{97}$$

where

$$\alpha_j^{(I)} = \frac{w_{i_j,j}}{\sum_{k=i_0}^{i_j} w_{k,j}}. \tag{98}$$

ii) total over all accident years

$$\widehat{msep}_{CDR_{tot,I+1}|\mathcal{D}_I}(0) = \sum_{i=i_J+1}^{I} \widehat{msep}_{CDR_{i,I+1}|\mathcal{D}_I}(0) \tag{99}$$

$$+2\sum_{i=i_J+1}^{I}\sum_{m=i+1}^{I} \widehat{C}_{i,J}^{CL(I)}\widehat{C}_{m,J}^{CL(I)} \left\{ \frac{\widehat{\sigma}_{j_i}^2}{\left(\widehat{f}_{j_i}^{CL(I)}\right)^2} \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j_i}} + \sum_{j=j_i+1}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL(I)}\right)^2} \alpha_j^{(I)} \frac{1}{\sum_{k=i_0}^{i_j-1} w_{k,j}} \right\}.$$

In [11] Merz and Wüthrich also derived the following formulas for estimating the msep of the one year run-off risk in future accounting years (see formulas (1.4) and (2.4) in [11]).

Theorem A.3 *The msep of the one-year run-off risk of future accounting years $I+1+k$ can be estimated by*

i) single accident year i (for $k = 1, \ldots, J - j_i - 1$)

$$\widehat{msep}_{CDR_{i,I+k+1}|\mathcal{D}_I}(0) =$$

$$= \left(\widehat{C}_{i,J}^{CL}\right)^2 \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL}\right)^2} \left(\frac{1}{\widehat{C}_{i,j_i+k}^{CL}} + \prod_{m=1}^{k} \left(1-\alpha_{j_i+m}^{(I)}\right) \frac{1}{\sum_{m=i_0}^{i-1-k} w_{m,j_i+k}} \right) \tag{100}$$

$$+ \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=j_i+k+1}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\alpha_{j-k}^{(I)} \prod_{m=0}^{k-1} \left(1-\alpha_{j-m}^{(I)}\right) \frac{1}{\sum_{m=i_0}^{i_j-1} w_{m,j}} \right).$$

ii) total over all accident years (for $k = 1, \ldots, J - 1$)

$$\widehat{msep}_{CDR_{tot,I+k+1}|\mathcal{D}_I}(0) = \sum_{i=i_J+k+1}^{I} \widehat{msep}_{CDR_{i,I+k+1}|\mathcal{D}_I}(0)$$

$$+2\sum_{i=i_J+k+1}^{I}\sum_{m=i+1}^{I} \widehat{C}_{i,J}^{CL}\widehat{C}_{m,J}^{CL} \frac{\widehat{\sigma}_{j_i+k}^2}{\left(\widehat{f}_{j_i+k}^{CL}\right)^2} \prod_{m=1}^{k}\left(1-\alpha_{j_i+m}^{(I)}\right) \frac{1}{\sum_{m=i_0}^{i-k-1} w_{m,j_i+k}} \tag{101}$$

$$+2\sum_{i=i_J+k+1}^{I}\sum_{m=i+1}^{I} \widehat{C}_{i,J}^{CL}\widehat{C}_{m,J}^{CL} \sum_{j=j_i+k+1}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\alpha_{j-k}^{(I)} \prod_{m=0}^{k-1}\left(1-\alpha_{j-m}^{(I)}\right) \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} \right).$$

B Discussion on the Estimation Error in 2006: comparison between BBMW estimator and Mack estimator

As mentioned in the introduction, there was was quite some discussion how to estimate the estimation error in the classical Mack model (see [2], [8], [4]). In [4] Gisler introduced the Bayesian CL-model. He then considered two specific Bayesian models, where one of them was essentially the same as the model, which is considered in [11]. By taking a non-informative prior he obtained the BBMW estimator for the estimation error. Thus the Bayesian approach seemed to speak in favour of the BBMW estimator.

The introduction of the Bayesian CL-model turned out to be very fruitful (see [5],[3]). It is a useful model for many situations in practice.

However, the Bayesian CL-model is a different model. It is different to the Mack model, and in the mean-time I have come to the conclusion that the BBMW estimator is the appropriate estimator in the Bayesian CL model, but not in the classical model of Mack.

In (37) we have seen that the Mack estimator is given by

$$\widehat{EE}_i^{Mack} = \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{i=i_0}^{i_j-1} w_{i,j}}. \tag{102}$$

In [2] the authors suggested an alternative estimator

$$\widehat{EE}_i^{BBMW} = w_{i,j_i}^2 \left(\prod_{j=j_i}^{J-1} \left(\left(\widehat{f}_j^{CL}\right)^2 + \frac{\widehat{\sigma}_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}} \right) - \prod_{j=j_i}^{J-1} \left(\widehat{f}_j^{CL}\right)^2 \right), \tag{103}$$

what they called "conditional resampling". More discussions on the two estimators at that time are found in [2], [8], [4].

We now reconsider the two estimators.

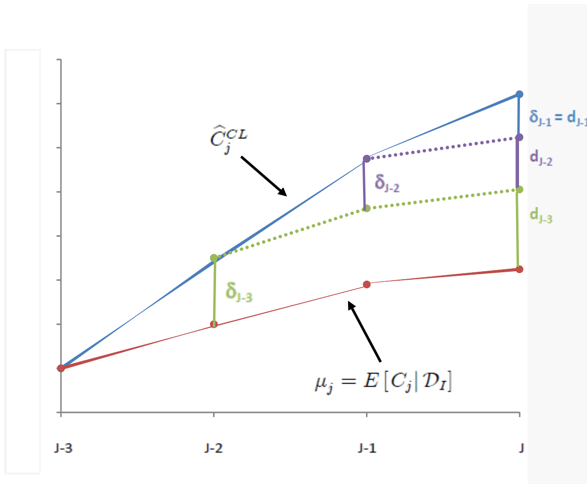
From the telescope formula we have obtained that the estimation error is given by

$$EE_i = B_i^2,$$

where

$$B_i = \sum_{j=j_i}^{J-1} \underbrace{\underbrace{\widehat{C}_{i,j}^{CL} \left(\widehat{f}_j^{CL} - f_j\right)}_{\delta_{ij}} \prod_{k=j+1}^{J-1} f_k}_{d_{ij}}. \tag{104}$$

The graphics below visualises (104) for accident year $i = i_{J-3}$. For simplicity we have dropped there the index i. Note that both, the trajectory $\left\{\widehat{C}_j^{CL} : j = J-3, \ldots, J\right\}$ as well as the trajectory $\left\{\mu_j := E\left[C_{i,j}\middle|\mathcal{D}_I\right] : j = J-3,\ldots,J\right\}$, are fixed and non stochastic.



To estimate B_i^2 we should have in mind that the δ_{ij} are realisations of r.v.

$$\begin{aligned} \Delta_{ij} &= \widehat{C}_{i,j}^{CL}\left(F_j - f_j\right) && (105) \\ \text{where } F_j &= \sum_{i=i_0}^{i_j-1} \frac{w_{i,j}}{w_{\bullet,j}} F_{i,j}, \quad w_{\bullet,j} = \sum_{i=i_0}^{i_j-1} w_{i,j} \ . \end{aligned}$$

and that

$$\delta_{ij}^2 = E\left[\Delta_{ij}^2\middle|\mathcal{D}_I\right].$$

Since $\widehat{C}_{i,j}^{CL}$ is a known and given multiplicative factor it is obvious that one should consider the conditional distribution of Δ_{ij} given $\mathcal{B}_j$ for estimating δ_{ij}^2. This is exactly what Mack does, and he obtains

$$\begin{aligned} E\left[\Delta_{ij}^2\middle|\mathcal{B}_j\right] &= \left(\widehat{C}_{i,j}^{CL}\right)^2 \frac{\sigma_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}, \\ \widehat{\delta}_{ij}^{2(Mack)} &= \widehat{E}\left[\Delta_{ij}^2\middle|\mathcal{B}_j\right] = \left(\widehat{C}_{i,j}^{CL}\right)^2 \frac{\widehat{\sigma}_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}. && (106) \end{aligned}$$

Since $\{\Delta_{i,j_0},\ldots\Delta_{i,j-1}\}$ are conditional on $\mathcal{B}_j$ known constants, it follows that the r.v. $\{\Delta_{ij}|\mathcal{B}_j : j = j_i,\ldots,J-1\}$ are uncorrelated. Hence $\{d_{ij} : j = j_i,\ldots,J-1\}$ can be considered as realisations of uncorrelated r.v. which then leads immediately to the Mack estimator (102).

In terms of the telescope formula, the BBMW estimator is obtained if one considers the r.v.

$$\widetilde{B}_i = \sum_{j=j_i}^{J-1} \widetilde{C}_{i,j}\left(\widetilde{F}_j - f_j\right) \prod_{k=j+1}^{J-1} f_k. \qquad (107)$$

where

$$\widetilde{C}_{i,j} = w_{i,j_j} \prod_{k=j_i}^{j-1} \widetilde{F}_k \quad \text{and where}$$

$$\text{the } \widetilde{F}_j \text{ are independent r.v. with } E\left[\widetilde{F}_j\right] = f_j \text{ and } \operatorname{Var}\left(F_j \middle| \mathcal{B}_j\right) = \frac{\sigma_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}.$$

Hence the known multiplicative constants $\widehat{C}_{i,j}^{CL}$ in (104) are replaced by r.v. $\widetilde{C}_{i,j}$ in (107) ,which hardly makes sense in the classical Mack model.

However, in the Bayesian CL-model the CL factors f_j are realisations of r.v. F_j^B (the superscript B indicates that we are in the Bayesian model). The estimation error is then given by

$$EE_i = B_i^{*^2},$$

where

$$\begin{aligned} B_i^* &= \left(E\left[C_{i,J}\middle|\mathcal{D}_I\right] - \widehat{C_{i,J}}^{CL}\right) \\ &= w_{i,j_i}\left(\prod_{j=j_i}^{J-1} F_j^B - \prod_{j=j_i}^{J-1} \widehat{f}_j^{CL}\right). \end{aligned} \tag{108}$$

With the telescope formula we also obtain that

$$B_i^* = \sum\nolimits_{j=j_i}^{J-1} \widetilde{C}_{i,j}^* \left(F_j^B - \widehat{f}_j^{CL}\right) \prod\nolimits_{k=j+1}^{J-1} \widehat{f}_k^{CL}, \tag{109}$$

where

$$\widetilde{C}_{i,j}^* = w_{i,j_j} \prod_{k=j_i}^{j} F_k^B.$$

which should be compared with (104) . In the following we always refer to the Bayesian CL-model considered in [11]. In this model the r.v. F_j^B are conditional on $\mathcal{D}_I$ independent. By taking a non informative prior and in the cases where the two first moments of the posterior distribution exist, these moments are given by

$$E\left[F_j^B\middle|\mathcal{D}_I\right] = \widehat{f}_j^{CL}, \tag{110}$$

$$\operatorname{Var}\left(F_j^B\middle|\mathcal{D}_I\right) = \frac{\sigma_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}. \tag{111}$$

From (108) , (110) , (111) and the conditional independence of the F_j^B we immediately see that

$$E\left[B_i^{*^2}\middle|\mathcal{D}_I\right] = w_{i,j_i}^2\left(\prod_{j=j_i}^{J-1}\left(\left(\widehat{f}_j^{CL}\right)^2 + \frac{\sigma_j^2}{\sum_{i=i_0}^{i_j-1} w_{i,j}}\right) - \prod_{j=j_i}^{J-1}\left(\widehat{f}_j^{CL}\right)^2\right).$$

Hence, the BBMW estimator is the appropriate estimator in the Bayesian CL-model, but not in the classical model of Mack.

C Proof of Properties 3.3.

We have to show that

a)
$$E\left[\left.F_{i,k}\right|\mathcal{B}_j\right] = f_k \quad \text{for} \quad j \leq k \leq J-1,$$

b)
$$\left\{\left.F_{i,k} : k = j, \ldots, J-1\right|\mathcal{B}_j\right\} \text{ are uncorrelated,}$$

c)
$$\left(\text{CoV}\left(\left.F_{i,k}\right|\mathcal{B}_j\right)\right)^2 = \frac{\sigma_k^2}{f_k} E\left[\left.\frac{1}{C_{i,k}}\right|\mathcal{B}_j\right] \text{ for } j \leq k \leq J-1,$$

Proof:

a)
$$E\left[\left.F_{i,k}\right|\mathcal{B}_j\right] = E\left[\left.E\left[\left.F_{i,k}\right|\mathcal{B}_k\right]\right|\mathcal{B}_j\right] = f_k.$$

b) Consider $F_{i,k}$ and $F_{i,l}$ with $j \leq k < l$. Then

$$\begin{aligned}
E\left[\left.F_{i,k}F_{i,l}\right|\mathcal{B}_j\right] &= E\left[\left.E\left[\left.F_{i,k}F_{i,l}\right|\mathcal{B}_l\right]\right|\mathcal{B}_j\right] \\
&= E[F_{i,k}\underbrace{E\left[\left.F_{i,l}\right|\mathcal{B}_l\right]}_{f_l}|\mathcal{B}_j] \\
&= f_l E\left[\left.F_{i,k}\right|\mathcal{B}_j\right] \\
&= f_l f_k.
\end{aligned}$$

c)

$$\begin{aligned}
\left(\text{CoV}\left(\left.F_{i,k}\right|\mathcal{B}_j\right)\right)^2 &= \frac{\text{Var}\left(\left.F_{i,k}\right|\mathcal{B}_j\right)}{f_k^2} \\
&= \frac{E\left[\left.\text{Var}\left(\left.F_{i,k}\right|\mathcal{B}_k\right)\right|\mathcal{B}_j\right]}{f_k^2} + \frac{\text{Var}\left(\left.E\left[\left.F_{i,k}\right|\mathcal{B}_k\right]\right|\mathcal{B}_j\right)}{f_k^2} \\
&= \frac{\sigma_k^2}{f_k^2} E\left[\left.\frac{1}{C_{i,k}}\right|\mathcal{B}_j\right] \quad \text{for} \quad j < k \leq J-1,
\end{aligned}$$

because the second summand in the second equation is equal to zero, since $E\left[\left.F_{i,k}\right|\mathcal{B}_k\right] = f_k$. □

D Theorem 6.4: Proof of the splitting property iii)

Proof:

i) single accident year i

$$\begin{aligned}
&\sum_{k=0}^{J-j_i-1} \widehat{msep}_{CDR_i^{(I+k+1)}|\mathcal{D}_I}(0) \\
&= \left(\widehat{C}_{i,J}^{CL}\right)^2 \left\{ \sum_{k=0}^{J-j_i-1} \frac{\widehat{\sigma}_{j_{i+k}}^2}{\left(\widehat{f}_{j_{i+k}}^{CL}\right)^2} \left(\frac{1}{w_{i,j_i+k}} + \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j_i+k}} \right) + \sum_{j=j_i+k+1}^{J-1} b_j^{(I+k)} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \right\} \\
&= \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\frac{1}{w_{i,j}} + \underbrace{\frac{1}{\sum_{l=i_0}^{i-1} w_{l,j}} + \sum_{k=0}^{j-j_i-1} b_j^{(I+k)}}_{B_{i,j}} \right). \qquad (112)
\end{aligned}$$

By plugging (91) into (112) we obtain

$$\begin{aligned}
B_{i,j_i} &= \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j}}, \\
B_{i,j_i+1} &= B_{i,j_i} + \frac{w_{i-1,j}}{\sum_{l=i_0}^{i-1} w_{l,j} \sum_{l=i_0}^{i-2} w_{l,j}} \\
&= \frac{1}{\sum_{l=i_0}^{i-1} w_{l,j}} + \frac{w_{i-1,j}}{\sum_{l=i_0}^{i-1} w_{l,j} \sum_{l=i_0}^{i-2} w_{l,j}} \\
&= \frac{1}{\sum_{l=i_0}^{i-2} w_{l,j}} = \frac{1}{\sum_{l=i_0}^{i_{j_i+1}-1} w_{l,j}} \\
&\vdots \\
B_{i,j_i+k} &= \frac{1}{\sum_{l=i_0}^{i-k-1} w_{l,j}} = \frac{1}{\sum_{l=i_0}^{i_{j_i+k}-1} w_{l,j}} \qquad (113)
\end{aligned}$$

Hence

$$\sum_{k=0}^{J-j_i-1} \widehat{msep}_{CDR_i^{(I+k+1)}|\mathcal{D}_I}(0) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=j_i}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \left(\frac{1}{w_{i,j}} + \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} \right),$$

which is identical to (44) .

ii) Total over all accident years

$$\sum_{k=0}^{J_i-1} \widehat{msep}_{CDR_{tot}^{(I+k+1)}|\mathcal{D}_I}(0 = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left\{ \sum_{k=0}^{J_i-1} \sum_{j=j_0}^{J-1} b_j^{(I+k)} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \right\}$$
$$= \left\{ \left(\widehat{C}_{tot,J}^{CL}\right)^2 \sum_{j=j_0}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \underbrace{\sum_{k=0}^{j-j_0} b_j^{(I+k)}}_{B_j} \right\}.$$

Since $j_I = j_0$ we see from (113) and the definition of $B_{i,j}$ in (112) that

$$\begin{aligned}
B_j &= \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} - \left(\frac{1}{w_{I,j}} - b_j^{(I)}\right) \\
&= \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} - \left(\frac{1}{\sum_{l=i_0}^{I-1} w_{l,j}} - \frac{w_{I,j}}{\left(\sum_{l=i_0}^{I-1} w_{l,j}\right)\left(\sum_{l=i_0}^{I-1} w_{l,j}\right)}\right) \\
&= \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} - \frac{1}{\sum_{l=i_0}^{I} w_{l,j}} \\
&= \frac{\sum_{l=i_j}^{I} w_{l,j}}{\sum_{l=i_0}^{I} w_{l,j}} \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} \\
&= \frac{\sum_{l=i_j}^{I} \widehat{C}_{l,,j}^{CL}}{\widehat{C}_{tot,,j}^{CL}} \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}}.
\end{aligned}$$

Hence

$$\sum_{k=0}^{J_i-1} \widehat{msep}_{CDR_{tot}^{(I+k+1)}|\mathcal{D}_I}(0 = \left(\widehat{C}_{tot,J}^{CL}\right)^2 \left(\sum_{j=j_0}^{J-1} \frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^{CL}\right)^2} \frac{1}{\sum_{l=i_0}^{i_j-1} w_{l,j}} q_j \right), \tag{114}$$

where

$$q_j = \frac{\sum_{l=i_j}^{I} \widehat{C}_{l,,j}^{CL}}{\widehat{C}_{tot,,j}^{CL}}.$$

(114) is identical to(46). □