# Prediction uncertainties in the Cape Cod reserving method

**Journal Article**

**Author(s):**
Saluz, Annina

# Prediction uncertainties in the Cape Cod reserving method

## Annina Saluz*

*Department of Mathematics, RiskLab, ETH Zurich, 8092 Zurich, Switzerland*

## Abstract

The Cape Cod (CC) method was designed by Bühlmann and Straub in order to overcome some shortcomings of the chain ladder (CL) method. Owing to its simplicity and because of the advantages over the CL method, the CC method has become a well-established method in practice. In this paper we consider a distribution-free stochastic model for the CC method. Within this model we give the parameter estimates and we derive estimates for the conditional mean square error of prediction for the CC method. In addition, we derive an estimate for the uncertainty in the claims development result.

## Keywords

## 1. Introduction

The Cape Cod (CC) method was developed by Bühlmann & Straub (1983). A derivation of the CC method is published in Straub (1988). In the CC method the reserve of an accident year is the product of an estimate of the expected ultimate claim and the estimated "still-to-come" factor of the corresponding accident year. The estimate of the expected ultimate claim is the product of the premium and an estimate for the loss ratio. For the estimation of the "still-to-come" factor, it is further assumed that there is a development pattern, which is the same for all accident years. Hereby, the development pattern describes the proportion of claims that evolve up to a certain development year relative to the ultimate claim amount.

The CC method was designed in order to overcome some shortcomings of the chain ladder (CL) method, which is one of the most popular claims reserving methods. The CL predictor of the ultimate claim is obtained by multiplying the current claims amount of an accident year (i.e. the claims amount known so far) by a product of development factors. Hence, the CL predictors of the ultimate claims are directly proportional to the current claims amount and therefore, if the current claims amount is zero or an outlier, the CL prediction gives unsatisfactory results. This is in particular the case for long-tailed lines of business in recent accident years. Moreover, the CL method is very sensitive to changes in individual claim numbers. Finally, the CL method is purely based on the claims data and it disregards the information in the premiums. These shortcomings are addressed by the CC method. The CC method is more robust than the CL method and, in contrast to the CL method, the CC method incorporates the information contained in the premiums.

The CC method is also closely related to the Bornhuetter–Ferguson (BF) method (Bornhuetter & Ferguson, 1972). As in the CC method, the BF reserve of an accident year is the product of an

*Correspondence to: Annina Saluz, Department of Mathematics, RiskLab, ETH Zurich, 8092 Zurich, Switzerland. Tel: + 41 44 632 68 30; Fax: + 41 44 632 15 23; E-mail: annina.saluz@math.ethz.ch

estimate of the expected ultimate claim and the estimated "still-to-come" factor of the corresponding accident year. For the estimation of the latter one also assumes that there is a development pattern, which is the same for all accident years. The estimate of the expected ultimate claim is given by an unbiased a priori estimate. The a priori estimate is assumed to be independent from the data of the corresponding accident year and it is not adjusted during the claims development. Often the a priori estimate is the product of the earned premium and an a priori estimate for the loss ratio. On the other hand, the loss ratio in the CC method is estimated from the data and hence it depends on the claims development. In the past, the development pattern for the BF method was often estimated by the development pattern resulting from the CL method. However, Mack (2006) criticises the use of the CL development pattern in the BF method, because the CL method assumes a multiplicative connection between past and future loss amounts, whereas the BF method establishes an additive connection (independence). Moreover, the incorporation of the a priori estimates in the BF method is a fundamental difference to the CL method. Hence, Saluz *et al.* (2011) argue that this additional information should also be incorporated in the estimates of the development pattern and that the use of the CL development pattern is inconsistent with the BF philosophy.

In a numerical example Bühlmann & Straub (1983) use the CL development pattern for the CC method. However, they remark that the estimation of the development pattern is still unsolved. Similar to the BF method, we think that an estimate of the development pattern for the CC method should incorporate the additional information available from the premiums, too.

Owing to its simplicity, the CC method is a well-established method in practice. However, in current literature there is no formula for the estimation of the prediction uncertainty in the CC method (see Dal Moro & Lo, 2014). In this paper we consider a stochastic model for the CC method and we derive estimates for the development pattern within this model. The incorporation of the information from the premiums in the estimation of the development pattern is consistent with our model assumptions. In order to quantify prediction uncertainty we derive an estimate for the conditional mean square error of prediction (MSEP) of the CC reserve and we give an estimate for the uncertainty in the one-year claims development result (CDR).

*Organisation of the paper*: in the remainder of this section we introduce the notation and we review the CC, CL and BF method. In section 2, we consider a distribution-free model for the CC method. In this model we calculate a development pattern that incorporates the information from the premiums and we derive estimates for the prediction uncertainty of the ultimate claim and for the uncertainty in the CDR. In section 2.2, we consider a distributional model and we show that in this special case the parameter estimates given in the distribution-free model can be derived with maximum likelihood (ML) estimation. A numerical example is given in section 3.

## 1.1. Notation and data structure

We denote the cumulative claims (cumulative payments or incurred losses) in accident year $i \in \{0, \dots, I\}$ at the end of development year $j \in \{0, \dots, J\}$ by $C_{i,j} > 0$ and we assume $J \leq I$. Let $X_{i,j} = C_{i,j} - C_{i,j-1}$ denote the incremental claims, where we set $C_{i,-1} = 0$. The summation over an index starting from 0 is denoted with a square bracket, for example:

$$C_{[k],j} = \sum_{i=0}^{k} C_{i,j}, \quad 0 \leq k \leq I, \quad 0 \leq j \leq J.$$

240

**Table 1.** Claims development triangle.

| accident | development year $j$ | | | | | |
|---|---|---|---|---|---|---|
| year $i$ | 0 | 1 | ... | $j$ | ... | $J$ |
| 0 | | | | | | |
| 1 | | | observations $\mathcal{D}_I$ | | | |
| $\vdots$ | | | | | | |
| $i$ | | | | | | |
| $\vdots$ | | | $\mathcal{D}_I^c$ | | | |
| $I$ | | | | | | |

The observations are given in the upper left triangle/trapezoid $\mathcal{D}_I$. The lower right triangle $\mathcal{D}_I^c$ needs to be predicted.

We assume that all claims are settled after development year $J$ and therefore the total ultimate claim of accident year $i$ is given by $C_{i,J}$. At time $I$ we have information in the upper left trapezoid/triangle:

$$\mathcal{D}_I = \big\{ C_{i,j} : i+j \leq I, \ j \leq J \big\},$$

and our goal is to predict the lower right triangle $\mathcal{D}_I^c = \{ C_{i,j} : i+j > I, i \leq I, \ j \leq J \}$. An illustration of the data is given in Table 1.

We define the outstanding loss liabilities for accident year $i$ at time $I$ by

$$R_i = C_{i,J} - C_{i,I-i} = \sum_{j=I-i+1}^{J} X_{i,j}, \quad I-J+1 \leq i \leq I, \tag{1.1}$$

and the total outstanding loss liabilities are defined by $R = \sum_{i=I-J+1}^{I} R_i$. For accident years $i = 0, \dots, I-J$ the ultimate claim $C_{i,J}$ is known at time $I$ and hence $R_i = 0$, $i = 0, \dots, I-J$.

**Remark 1.1** If $C_{i,j}$ denote cumulative payments then formula (1.1) gives the "true" outstanding loss liabilities. For incurred losses $C_{i,j}$ the "true" outstanding loss liabilities are given by

$$\overline{R}_i = C_{i,J} - C_{i,I-i} + C_{i,I-i} - C_{i,I-i}^{\text{paid}}, \quad I-J+1 \leq i \leq I, \tag{1.2}$$

with $C_{i,j}^{\text{paid}}$ denoting the cumulative payments of accident year $i$ up to development year $j$. Note that the additional term $C_{i,I-i} - C_{i,I-i}^{\text{paid}}$ is observable at time $I$ and has no impact on the claims prediction problem. Therefore, we only consider the outstanding loss liabilities as defined in (1.1).

The budgeting and planning process in an insurance company gives information on the premium level of each accident year. Often premium levels are adjusted to business cycles. For instance, in soft markets, the premium level is often increased by a certain percentage. We denote the earned on-level premium for accident year $i$ by $\nu_i$, $0 \leq i \leq I$. This means that $\nu_i$ is the earned premium, adjusted for business cycles, such that the expected loss ratio is the same for all accident years. These adjustments are based on the information from the budgeting and planning process, which should be available in a company. In periods of a stable premium level, the earned premiums can be used for the on-level premiums $\nu_i$. Otherwise, if it is for instance known that the premiums of an accident year are

increased by a certain percentage owing to soft market conditions, then the on-level premiums are corrected for this increase. Mack (2006) also derives a procedure for the calculation of on-level premiums from the data. However, with the procedure of Mack (2006) the resulting on-level premiums depend on the data $\mathcal{D}_I$. Here, we assume that the calculation of the $\nu_i$'s does not incorporate the data in the triangle and that the $\nu_i$'s can be considered as known constants.

In order to simplify notation we denote the index of the last observed development year in accident year $i$ by $\iota(i) = \min(I-i, J)$.

## 1.2. CL method

The CL prediction of the ultimate claim $C_{i,J}$ of accident year $i > I - J$ is given by

$$\hat{C}_{i,J}^{\mathrm{CL}} = C_{i,\iota(i)} \prod_{j=\iota(i)}^{J-1} \hat{f}_j, \quad \text{where} \quad \hat{f}_j = \frac{C_{[I-j-1],j+1}}{C_{[I-j-1],j}}. \tag{1.3}$$

The CL reserve is given by

$$\begin{aligned}
\hat{R}_i^{\mathrm{CL}} &= C_{i,\iota(i)} \left( \prod_{j=\iota(i)}^{J-1} \hat{f}_j - 1 \right) = C_{i,\iota(i)} \prod_{j=\iota(i)}^{J-1} \hat{f}_j \left( 1 - \prod_{j=\iota(i)}^{J-1} \hat{f}_j^{-1} \right) \\
&= \hat{C}_{i,J}^{\mathrm{CL}} \left( 1 - \hat{\beta}_{\iota(i)}^{\mathrm{CL}} \right).
\end{aligned}$$

The CL development pattern

$$\hat{\beta}_j^{\mathrm{CL}} = \prod_{k=j}^{J-1} \hat{f}_k^{-1}, \quad 0 \le j \le J-1, \quad \hat{\beta}_J^{\mathrm{CL}} = 1, \tag{1.4}$$

is an estimate of the development pattern $\beta_j$, $0 \le j \le J$, which describes the percentage of claims that evolve up to development year $j$ relative to the ultimate claim.

There are different stochastic models that justify the use of the CL algorithm. A distribution-free model was suggested by Mack (1993).

As mentioned above, the proportionality of the CL predictor to the current claims amount $C_{i,\iota(i)}$ is sometimes problematic in long-tailed lines of business for recent accident years. Note that the $\hat{f}_j$ for late development years are based on few observations. A change of $C_{0,J}$ has a multiplicative impact on the prediction of the ultimate claims of all accident years. Therefore, the CL method is very sensitive to changes of individual numbers.

## 1.3. CC method

The CC predictor (see Bühlmann & Straub, 1983) for the ultimate claim $C_{i,J}$ is given by

$$\hat{C}_{i,J}^{\mathrm{CC}} = C_{i,\iota(i)} + \nu_i \hat{q} \left( 1 - \hat{\beta}_{\iota(i)} \right), \tag{1.5}$$

where

$$\hat{q} = \frac{\sum_{i=0}^{I} C_{i,\iota(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{\iota(i)}}. \tag{1.6}$$

242

Moreover, $\hat{\beta}_{\iota(i)}$ is an estimate of $\beta_{\iota(i)}$. As above, the development pattern $\beta_j$, $0 \leq j \leq J$, describes the percentage of claims that evolve up to development year $j$ relative to the ultimate claim. Note that for each $i$

$$\hat{q}_i = \frac{C_{i,\iota(i)}}{\nu_i \hat{\beta}_{\iota(i)}}$$

is an estimate for the loss ratio $q$ and hence $\hat{q}$ is a weighted average loss ratio with the weights $\nu_i \hat{\beta}_{\iota(i)}$.

In the original article of Bühlmann & Straub (1983) it is mentioned that the estimation of the development pattern $\beta_j$ is an unsolved problem. In practice the development pattern is often estimated by the CL development pattern given in (1.4). In this case the CC predictor of the ultimate claim can be written as

$$\hat{C}_{i,J}^{\text{CC}} = C_{i,\iota(i)} + \nu_i \hat{q} \hat{\beta}_{\iota(i)}^{\text{CL}} \left( \prod_{j=\iota(i)}^{J-1} \hat{f}_j - 1 \right),$$

that is, we replace the diagonal element $C_{i,\iota(i)}$ in the CL prediction

$$\hat{C}_{i,J}^{\text{CL}} = C_{i,\iota(i)} \prod_{j=\iota(i)}^{J-1} \hat{f}_j = C_{i,\iota(i)} + C_{i,\iota(i)} \left( \prod_{j=\iota(i)}^{J-1} \hat{f}_j - 1 \right),$$

by the more robust value

$$\nu_i \hat{q} \hat{\beta}_{\iota(i)}^{\text{CL}} = \nu_i \hat{\beta}_{\iota(i)}^{\text{CL}} \frac{\sum_{i=0}^{I} C_{i,\iota(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{\iota(i)}^{\text{CL}}}.$$

Therefore, one can interpret the CC method as an application of the CL algorithm to these more robust diagonal values (see Wüthrich & Merz, 2008). On the other hand, if we take formula (1.5), the CC method is close to the BF method in which the predictor of the ultimate claim is given by

$$\hat{C}_{i,J}^{\text{BF}} = C_{i,\iota(i)} + \hat{\mu}_i \left( 1 - \hat{\beta}_{\iota(i)} \right), \tag{1.7}$$

where $\hat{\mu}_i$ is an unbiased a priori estimate of the expected ultimate claim $E[C_{i,J}]$. In the BF method $\hat{\mu}_i$ is usually an estimate that incorporates information from the earned premiums $p_i$. In this case it is assumed that for each accident year an a priori estimate for the loss ratio $\hat{q}_i$ is available and $\hat{\mu}_i = p_i \hat{q}_i$. The loss ratio estimates $\hat{q}_i$ are typically assumed to be independent from the data of the corresponding accident year $i$ (see, for instance Mack, 2008; Alai *et al.*, 2010; Saluz *et al.*, 2011). These a priori estimates $\hat{q}_i$ are known at the beginning of each accident year and they are not adjusted during the claims development. In contrast to the BF method, the loss ratio $q$ in the CC method is estimated from the data $\mathcal{D}_I$. Moreover, the CC method assumes a constant loss ratio, which means that the earned premiums $p_i$ need to be adjusted to on-level premiums $\nu_i$. Instead of assuming the availability of unbiased a priori estimates of $E[C_{i,J}]$, the CC method assumes that the expected ultimate claims are known up to a constant factor. In the BF predictor (1.7) we further note the additive relation with the current claims amount $C_{i,\iota(i)}$. This additive relation leads to a reserve estimate:

$$\hat{R}_i^{\text{BF}} = \hat{\mu}_i (1 - \hat{\beta}_{I-i}),$$

which does not directly depend on $C_{i,\iota(i)}$. The BF method typically assumes independence between past and future claims (see Mack, 2008). In contrast, the CL predictor given in (1.3) and the corresponding reserve are directly proportional to $C_{i,\iota(i)}$. Owing to this different relation (additive or multiplicative), Mack (2006) criticises the use of the CL development pattern in the BF method and suggests different estimates. Saluz *et al.* (2011) argue that the information from the a priori estimates should be

incorporated in the estimation of the development pattern in the BF method and derive corresponding estimators based on ML considerations. For the CC method we suggest similar estimators that incorporate the premiums $\nu_i$. The suggested estimators are derived from the best linear unbiased estimators. In section 2.2 it will be seen that under the assumption that incremental claims are overdispersed Poisson (ODP) distributed, these estimators coincide with the ML estimators.

In order to quantify the uncertainties in the CC predictions we will consider second moments. The conditional MSEP for accident year $i > I - J$, given $\mathcal{D}_I$, is given by

$$\text{MSEP}_{C_{i,J}\,|\,\mathcal{D}_I}(\hat{C}_{i,J}^{\text{CC}}) = E\left[\left(C_{i,J} - \hat{C}_{i,J}^{\text{CC}}\right)^2 \middle| \mathcal{D}_I\right]$$

$$\overset{(1.5)}{=} E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \nu_i \hat{q}\left(1 - \hat{\beta}_{I-i}\right)\right)^2 \middle| \mathcal{D}_I\right].$$

Similarly, for aggregated accident years we want to estimate

$$\text{MSEP}_{\sum_{i=I-J+1}^{I} C_{i,J}\,|\,\mathcal{D}_I}\left(\sum_{i=I-J+1}^{I} \hat{C}_{i,J}^{\text{CC}}\right) = E\left[\left(\sum_{i=I-J+1}^{I} \left(C_{i,J} - \hat{C}_{i,J}^{\text{CC}}\right)\right)^2 \middle| \mathcal{D}_I\right].$$

In order to estimate the conditional MSEP a stochastic model will be introduced in the following section.

## 2. Stochastic model for the CC method

### 2.1. CC method in a distribution-free model

For the estimation of the prediction uncertainty in the CC method we assume the following underlying distribution-free model.

**Model Assumptions 2.1 (CC model)** Incremental claims $X_{i,j}$ are independent and there exist positive parameters $q$, $\sigma_j^2$, $0 \leq j \leq J$, and a development pattern $\gamma_0, \dots, \gamma_J$, with $\sum_{j=0}^{J} \gamma_j = 1$ such that

$$E[X_{i,j}] = \nu_i q \gamma_j \quad \text{and} \quad \text{Var}(X_{i,j}) = \nu_i q \sigma_j^2$$

for $0 \leq i \leq I$ and $0 \leq j \leq J$.

Note that $\gamma_j$ denotes the incremental development pattern $\gamma_j = \beta_j - \beta_{j-1}$, $1 \leq j \leq J$, and $\gamma_0 = \beta_0$. Further, we remark that the expected loss ratio $q$ is the same for all accident years. This means that the premiums $\nu_i$ are "on-level" premiums as explained below Remark 1.1.

### 2.1.1. Parameter estimation
Under Model Assumptions 2.1 the best linear unbiased estimate for $\gamma_j$ is given by

$$\tilde{\gamma}_j^{\text{raw}} = \frac{X_{[I-j],j}}{q\nu_{[I-j]}}.$$

In general, the $\tilde{\gamma}_j^{\text{raw}}$'s do not sum up to 1 and hence we consider the normalised development pattern

$$\hat{\gamma}_j = \frac{\tilde{\gamma}_j^{\text{raw}}}{\sum_{k=0}^{J} \tilde{\gamma}_k^{\text{raw}}}.$$

Note that the normalised development pattern $\hat{\gamma}_j$ does not depend on the unknown loss ratio $q$, that is, we have

$$\hat{\gamma}_j = \frac{\tilde{\gamma}_j^{\text{raw}}}{\sum_{k=0}^{J} \tilde{\gamma}_k^{\text{raw}}} = \frac{\hat{\gamma}_j^{\text{raw}}}{\sum_{k=0}^{J} \hat{\gamma}_k^{\text{raw}}} = \frac{\hat{\gamma}_j^{\text{raw}}}{\hat{\beta}_J^{\text{raw}}}, \tag{2.1}$$

where

$$\hat{\gamma}_j^{\text{raw}} = \frac{X_{[I-j],j}}{\nu_{[I-j]}} \tag{2.2}$$

and $\hat{\beta}_j^{\text{raw}} = \sum_{k=0}^{j} \hat{\gamma}_k^{\text{raw}}$. The cumulative development pattern $\beta_j$ is then estimated by

$$\hat{\beta}_j = \sum_{k=0}^{j} \hat{\gamma}_k = \frac{\hat{\beta}_j^{\text{raw}}}{\hat{\beta}_J^{\text{raw}}}. \tag{2.3}$$

The best linear unbiased estimator for $q$ based on the observations $X_{i,j}$, $i+j \leq I$, $j \leq J$, is given by

$$\tilde{q} = \frac{\sum_{j=0}^{J} \sum_{i=0}^{I-j} X_{i,j} \frac{\gamma_i}{\sigma_j^2}}{\sum_{j=0}^{J} \sum_{i=0}^{I-j} \nu_i \frac{\gamma_j^2}{\sigma_j^2}}. \tag{2.4}$$

If we insert the estimates (2.1) for the $\gamma_j$'s in $\tilde{q}$ we obtain

$$\hat{\tilde{q}} = \frac{\sum_{j=0}^{J} \sum_{i=0}^{I-j} X_{i,j} \frac{X_{[I-j],j}/\nu_{[I-j]}}{\hat{\beta}_J^{\text{raw}} \sigma_j^2}}{\sum_{j=0}^{J} \nu_{[I-j]} \frac{X_{[I-j],j}^2/(\nu_{[I-j]})^2}{(\hat{\beta}_J^{\text{raw}})^2 \sigma_j^2}} = \hat{\beta}_J^{\text{raw}} = \hat{q}, \tag{2.5}$$

where $\hat{q}$ was defined in (1.6). The third equality in (2.5) is derived as follows:

$$\hat{q} \overset{(1.6)}{=} \frac{\sum_{i=0}^{I} C_{i,t(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{t(i)}} \overset{(2.1)}{=} \frac{\sum_{i=0}^{I} C_{i,t(i)}}{\sum_{\substack{i+j \leq I, \\ j \leq J}} \nu_i \frac{X_{[I-j],j}/\nu_{[I-j]}}{\hat{\beta}_J^{\text{raw}}}}$$

$$= \frac{\hat{\beta}_J^{\text{raw}} \sum_{j=0}^{J} X_{[I-j],j}}{\sum_{j=0}^{J} \nu_{[I-j]} \frac{X_{[I-j],j}}{\nu_{[I-j]}}} = \hat{\beta}_J^{\text{raw}} = \hat{\tilde{q}}. \tag{2.6}$$

Hence, in the CC model 2.1 the use of the CC loss ratio estimate (1.6) can be justified by a derivation from the best linear unbiased estimate of $q$. Under Model Assumptions 2.1 we therefore use the following estimates

$$\hat{\gamma}_j = \frac{\frac{X_{[I-j],j}}{\nu_{[I-j]}}}{\sum_{l=0}^{J} \frac{X_{[I-l],l}}{\nu_{[I-l]}}} = \frac{\hat{\gamma}_j^{\text{raw}}}{\hat{\beta}_J^{\text{raw}}},$$

$$\hat{q} = \frac{\sum_{i=0}^{I} C_{i,t(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{t(i)}} = \sum_{l=0}^{J} \frac{X_{[I-l],l}}{\nu_{[I-l]}} = \hat{\beta}_J^{\text{raw}}. \tag{2.7}$$

Note that the loss ratio $\hat{q}$ corresponds to the multiplicative normalisation factor $\hat{\beta}_J^{\text{raw}}$ of the non-normalised development pattern given in (2.2). For the CC reserve and ultimate prediction for accident years $i > I - J$ we then have

$$\hat{R}_i^{\text{CC}} = \nu_i \hat{q}(1 - \hat{\beta}_{I-i}) = \nu_i \hat{\beta}_J^{\text{raw}} \left(1 - \frac{\hat{\beta}_{I-i}^{\text{raw}}}{\hat{\beta}_J^{\text{raw}}}\right) = \nu_i(\hat{\beta}_J^{\text{raw}} - \hat{\beta}_{I-i}^{\text{raw}}), \tag{2.8}$$

$$\hat{C}_{i,J}^{\text{CC}} = C_{i,I-i} + \hat{R}_i^{\text{CC}}. \tag{2.9}$$

245

From (2.8) we note that the loss ratio estimate $\hat{q}$ in $\hat{R}_i^{CC}$ can be omitted if the development pattern is not normalised.

**Remark 2.2**

- In order to incorporate a tail development $\hat{\gamma}_{J+1}$ one should normalise the raw development pattern $\hat{\gamma}_j^{raw}$ such that for the normalised development pattern $\hat{\gamma}_j = \hat{\gamma}_j^{raw}/c$ we obtain

$$\sum_{j=0}^{J} \hat{\gamma}_j = \sum_{j=0}^{J} \frac{\hat{\gamma}_j^{raw}}{c} = \frac{\hat{\beta}_J^{raw}}{c} = \hat{\beta}_J = 1 - \hat{\gamma}_{J+1},$$

where $c$ denotes the normalising constant. Therefore, we obtain $\hat{\gamma}_j = \hat{\gamma}_j^{raw}/c = \hat{\gamma}_j^{raw}(1-\hat{\gamma}_{J+1})/\hat{\beta}_J^{raw}$. Inserting $\hat{\gamma}_j$ in (2.4) one gets the estimate

$$\hat{q} = \frac{\hat{\beta}_J^{raw}}{1 - \hat{\gamma}_{J+1}}.$$

For the reserves one obtains in this case:

$$\hat{R}_i^{CC} = \nu_i \hat{q}(1 - \hat{\beta}_{I-i}) = \nu_i \left( \frac{\hat{\beta}_J^{raw}}{1 - \hat{\gamma}_{J+1}} - \hat{\beta}_{I-i}^{raw} \right).$$

Here, the tail development $\hat{\gamma}_{J+1}$ is assumed to be given by an expert.

- The representation (2.8) of the CC reserve with the non-normalised development pattern $\hat{\beta}_J^{raw}$ substantially simplifies the analysis of prediction uncertainty. For instance, for the non-normalised development pattern one can explicitly calculate the covariances $\text{Cov}(\hat{\beta}_j^{raw}, \hat{\beta}_k^{raw})$, $0 \leq j$, $k \leq J$. For the normalised development pattern this is not possible because of the random variables in the denominators.

- In the BF method the estimation of the development pattern suggested by Mack (2008) is also based on estimators of the form

$$\tilde{\gamma}_j^{raw} = \frac{X_{[I-j],j}}{\mu_{[I-j]}}, \quad \text{where} \quad \mu_i = E[C_{i,J}], \quad 0 \leq i \leq I.$$

However, the unknown $\mu_i$ are replaced by the a priori estimates $\hat{\mu}_i$, which possibly incorporate non-constant a priori loss ratios $\hat{q}_i$. In contrast to CC, the BF reserve cannot be rewritten as simply as the CC reserve (2.8). For BF, a normalised development pattern has to be used. Moreover, in the prediction error of the BF reserve the uncertainties in the a priori estimates need to be incorporated.

For the estimation of the conditional MSEP we further need estimates for $q\sigma_j^2$. Denote by $Y_{i,j}$ the "incremental loss ratios" $Y_{i,j} = X_{i,j}/\nu_i$. Observe that

$$\text{Var}(Y_{i,j}) = \frac{q\sigma_j^2}{\nu_i},$$

and $q\sigma_j^2$ is the variance of $Y_{i,j}$ per unit of $\nu_i$. An unbiased estimator for $q\sigma_j^2$, $j \neq I$, is given by

$$\widehat{q\sigma^2}_j = \frac{1}{I-j} \sum_{i=0}^{I-j} \nu_i \left( \frac{X_{i,j}}{\nu_i} - \hat{\gamma}_j^{raw} \right)^2, \quad 0 \leq j \leq J, \, j \neq I. \tag{2.10}$$

246

In the case where $I = J$ an estimate for $q\sigma_J^2$ is obtained with the extrapolation from Mack (1993):

$$\widehat{q\sigma_J^2} = \min\left(\widehat{q\sigma^2}_{J-1}, \widehat{q\sigma^2}_{J-2}, \frac{\widehat{q\sigma^2}^2_{J-1}}{\widehat{q\sigma^2}_{J-2}}\right) = \min\left(\widehat{q\sigma^2}_{J-2}, \frac{\widehat{q\sigma^2}^2_{J-1}}{\widehat{q\sigma^2}_{J-2}}\right).$$

### 2.1.1. Prediction uncertainty

Under Model Assumptions 2.1 we have for the conditional MSEP of accident year $i > I - J$, given $\mathcal{D}_I$:

$$\mathrm{MSEP}_{C_{i,J} \mid \mathcal{D}_I}(\hat{C}_{i,J}^{\mathrm{CC}}) = E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \nu_i \hat{q}(1-\hat{\beta}_{I-i})\right)^2 \middle| \mathcal{D}_I\right]$$

$$= \underbrace{\sum_{j=I-i+1}^{J} \mathrm{Var}(X_{i,j})}_{\mathrm{PV}_i} + \underbrace{\nu_i^2\left(q(1-\beta_{I-i}) - \hat{q}(1-\hat{\beta}_{I-i})\right)^2}_{=\mathrm{PEE}_i}, \qquad (2.11)$$

where $\mathrm{PV}_i$ is the process variance of accident year $i$, which describes the randomness of the $X_{i,j}$, and $\mathrm{PEE}_i$ the parameter estimation error, which describes the uncertainties in the parameter estimates. The conditional MSEP can be estimated as follows:

**Estimate 2.3 (MSEP, CC model)** Under Model Assumptions 2.1 the conditional MSEP of the CC predictor $\hat{C}_{i,J}^{\mathrm{CC}}$ given in (2.9) is estimated by

$$\widehat{\mathrm{MSEP}}_{C_{i,J} \mid \mathcal{D}_I}\left(\hat{C}_{i,J}^{\mathrm{CC}}\right) = \widehat{\mathrm{PV}}_i + \widehat{\mathrm{PEE}}_i,$$

where $\widehat{\mathrm{PV}}_i$ and $\widehat{\mathrm{PEE}}_i$ are given by

$$\widehat{\mathrm{PV}}_i = \sum_{j=I-i+1}^{J} \nu_i \widehat{q\sigma^2}_j, \quad \widehat{\mathrm{PEE}}_i = \nu_i^2 \sum_{j=I-i+1}^{J} \frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}}, \qquad (2.12)$$

and where $\widehat{q\sigma^2}_j$ is given in (2.10). For aggregated accident years the corresponding conditional MSEP is estimated by

$$\widehat{\mathrm{MSEP}}_{\sum_{i=I-J+1}^{I} C_{i,J} \mid \mathcal{D}_I}\left(\sum_{i=I-J+1}^{I} \hat{C}_{i,J}^{\mathrm{CC}}\right) = \sum_{i=I-J+1}^{I} \widehat{\mathrm{MSEP}}_{C_{i,J} \mid \mathcal{D}_I}\left(\hat{C}_{i,J}^{\mathrm{CC}}\right) + 2 \sum_{I-J+1 \le i < k \le I} \widehat{\mathrm{PEE}}_{i,k},$$

where $\widehat{\mathrm{PEE}}_{i,k}$ is given by

$$\widehat{\mathrm{PEE}}_{i,k} = \nu_i \nu_k \sum_{j=I-i+1}^{J} \frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}}, \quad i < k.$$

The derivation of Estimate 2.3 is given in Appendix A.

### 2.1.2. Conditional MSEP of the one-year CDR

So far we have considered the prediction of the outstanding liabilities at time $I$. At the end of accounting year $I$ we have an additional diagonal of observations in the claims triangle/trapezoid and the predictions of the ultimate claims are updated according to this new information. The one-year

CDR is the difference between two successive predictors of the ultimate claim (see, for instance Merz & Wüthrich, 2008). More precisely, we define the CDR for accounting year $I+1$ and for accident year $i$ by

$$\mathrm{CDR}_i^{(I+1)} = \hat{C}_{i,J}^{(I)} - \hat{C}_{i,J}^{(I+1)}, \tag{2.13}$$

where $\hat{C}_{i,J}^{(k)}$, $k = I, I+1$, denotes the prediction of $C_{i,J}$ at time $k$. Under the CC model 2.1 we have for $I-J+1 \leq i \leq I$

$$\hat{C}_{i,J}^{(I)} = C_{i,I-i} + \nu_i \hat{q}^{(I)} \left(1 - \hat{\beta}_{I-i}^{(I)}\right) = C_{i,I-i} + \nu_i \left(\hat{\beta}_J^{(I)\mathrm{raw}} - \hat{\beta}_{I-i}^{(I)\mathrm{raw}}\right), \tag{2.14}$$

$$\hat{C}_{i,J}^{(I+1)} = C_{i,I-i+1} + \nu_i \hat{q}^{(I+1)} \left(1 - \hat{\beta}_{I-i+1}^{(I+1)}\right) = C_{i,I-i+1} + \nu_i \left(\hat{\beta}_J^{(I+1)\mathrm{raw}} - \hat{\beta}_{I-i+1}^{(I+1)\mathrm{raw}}\right), \tag{2.15}$$

where $\hat{q}^{(k)} = \hat{\beta}_J^{(k)\mathrm{raw}}$, $k = I, I+1$, and

$$\hat{\beta}_j^{(k)\mathrm{raw}} = \sum_{l=0}^j \hat{\gamma}_l^{(k)\mathrm{raw}}, \quad \text{with} \quad \hat{\gamma}_j^{(k)\mathrm{raw}} = \frac{X_{[k-j],j}}{\nu_{[k-j]}}. \tag{2.16}$$

Similarly, we define the CDR for aggregated accident years at time $I+1$ by

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

At time $I$ we predict $\mathrm{CDR}_i^{(I+1)}$ by 0 as we consider $\hat{C}_{i,J}^{(I)}$ as best estimate based on the available information at time $I$. New solvency regulations such as the Swiss Solvency Test (FINMA, 2006) require additional risk capital for protection against possible shortfalls in this one-year CDR. In order to quantify uncertainties in the CDRs we consider second moments given by

$$\mathrm{MSEP}_{\mathrm{CDR}_i^{(I+1)}|\mathcal{D}_I}(0) = E\left[\left(\mathrm{CDR}_i^{(I+1)}\right)^2 \middle| \mathcal{D}_I\right],$$

$$\mathrm{MSEP}_{\mathrm{CDR}^{(I+1)}|\mathcal{D}_I}(0) = E\left[\left(\mathrm{CDR}^{(I+1)}\right)^2 \middle| \mathcal{D}_I\right].$$

In Appendix A the following estimators for these conditional MSEPs are derived:

**Estimate 2.4 (MSEP CDR, CC model)** Under Model Assumptions 2.1 the conditional MSEP of the CDR of accident year $i$ given in (2.13) is estimated by

$$\widehat{\mathrm{MSEP}}_{\mathrm{CDR}_i^{(I+1)}|\mathcal{D}_I}(0) = \nu_i \widehat{q\sigma^2}_{I+1-i} \frac{\nu_{[i]}}{\nu_{[i-1]}} + \nu_i^2 \sum_{j=I+2-i}^J \widehat{q\sigma^2}_j \frac{\nu_{I+1-j}}{\nu_{[I-j]}\nu_{[I+1-j]}},$$

where the $\widehat{q\sigma^2}_j$'s are given in (2.10). For aggregated accident years the corresponding conditional MSEP is estimated by

$$\widehat{\mathrm{MSEP}}_{\mathrm{CDR}^{(I+1)}|\mathcal{D}_I}(0) = \sum_{i=I-J+1}^I \widehat{\mathrm{MSEP}}_{\mathrm{CDR}_i^{(I+1)}|\mathcal{D}_I}(0) + 2 \sum_{I-J+1 \leq i < k \leq I} \nu_i \nu_k \left(\frac{\widehat{q\sigma^2}_{I+1-i}}{\nu_{[i-1]}} + \sum_{j=I+2-i}^J \frac{\widehat{q\sigma^2}_j \nu_{I+1-j}}{\nu_{[I-j]}\nu_{[I+1-j]}}\right).$$

In order to compare the formula given in Estimate 2.4 with the formula for the conditional MSEP of the predictor for the ultimate claim given in Estimate 2.3, we rewrite Estimate 2.4 as follows

$$\widehat{\mathrm{MSEP}}_{\mathrm{CDR}_i^{(I+1)}|\mathcal{D}_I}(0) = \nu_i \widehat{q\sigma^2}_{I+1-i} + \nu_i^2 \frac{\widehat{q\sigma^2}_{I+1-i}}{\nu_{[i-1]}} + \nu_i^2 \sum_{j=I+2-i}^J \frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}} \frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}.$$

248

Note that the conditional MSEP of the CDR considers only the first term of $\widehat{\text{PV}}_i$, which corresponds to the variance in the next diagonal. For the parameter estimation error the first term is fully incorporated and all other terms in $\widehat{\text{PEE}}_i$ are scaled down by the factors $\nu_{I+1-j}/\nu_{[I+1-j]} < 1$. Similar relations are observed in the case of the CL method (see Merz & Wüthrich, 2008; Bühlmann *et al.*, 2009).

## 2.2. ODP model

A special case of Model Assumptions 2.1 is obtained by assuming that the incremental claims are ODP distributed with constant dispersion parameter $\phi$. Note that incremental claims are in this case assumed to be positive.

**Model Assumptions 2.5 (ODP model)** Incremental claims $X_{i,j}$ are independent and ODP distributed with

$$E[X_{i,j}] = \nu_i q \gamma_j \quad \text{and} \quad \text{Var}(X_{i,j}) = \phi \nu_i q \gamma_j = \phi E[X_{i,j}]$$

for $0 \leq i \leq I$ and $0 \leq j \leq J$, where $\gamma_0, \ldots, \gamma_J$, $\phi$ and $q$ are positive parameters and $\sum_{j=0}^{J} \gamma_j = 1$.

The ODP model is well known in claims reserving and was for instance used in England & Verrall (2002).

**Theorem 2.6** Under Model Assumptions 2.5 the ML estimates for $q$ and $\gamma_j$ coincide with the estimates given in (2.1) and (2.7).

*Proof:* The log-likelihood function is given by

$$l_{\mathcal{D}_I} = \sum_{i+j \leq I, j < J} \frac{1}{\phi} \left( X_{i,j} \left( \log \nu_i + \log(q\gamma_j) \right) - \nu_i q \gamma_j \right) + r$$

$$+ \sum_{i=0}^{I-J} \frac{1}{\phi} \left( X_{i,J} \left( \log \nu_i + \log \left( q - \sum_{j=0}^{J-1} q\gamma_j \right) \right) - \nu_i \left( q - \sum_{j=0}^{J-1} q\gamma_j \right) \right),$$

where $r$ is a term that does not depend on the parameters $q$ and $\gamma_0, \gamma_1, \ldots, \gamma_J$. Note that the ML estimate for $q$ yields

$$\hat{q} = \frac{\sum_{i=0}^{I} C_{i,\iota(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{\iota(i)}}, \tag{2.17}$$

which is the formula for the CC loss ratio estimate given in (1.6). The ML estimate for $\gamma_j$, $0 \leq j < J$, is given by the equation

$$\hat{\gamma}_j = \frac{X_{[I-j],j}}{\nu_{[I-j]}\hat{q} + \frac{X_{[I-J],J}}{\hat{\gamma}_J} - \nu_{[I-J]}\hat{q}},$$

where $\hat{\gamma}_J = 1 - \sum_{j=0}^{J-1} \hat{\gamma}_j$. The equation for $\hat{\gamma}_j$ can be rewritten as

$$\hat{\gamma}_j = \frac{X_{[I-j],j}}{\nu_{[I-j]}\hat{q} + \kappa}, \tag{2.18}$$

where $\kappa = \frac{X_{[I-J],J}}{\hat{\gamma}_J} - \nu_{[I-J]}\hat{q} > -\nu_{[I-J]}\hat{q}$ ensures that $\sum_{j=0}^{J} \hat{\gamma}_j = 1$. We insert the estimates $\hat{\gamma}_j$ in the ML equation for $q$ and obtain

$$\hat{q} = \frac{\sum_{i=0}^{I} C_{i,\iota(i)}}{\sum_{i=0}^{I} \nu_i \sum_{j=0}^{\iota(i)} \frac{X_{[I-j],j}}{\nu_{[I-j]}\hat{q} + \kappa}}. \tag{2.19}$$

It follows that

$$\sum_{\substack{i+j \le I, \\ j \le J}} X_{i,j} = \sum_{i=0}^{I} C_{i,t(i)} \overset{(2.19)}{=} \hat{q} \sum_{\substack{i+j \le I, \\ j \le J}} \nu_i \frac{X_{[I-j],j}}{\nu_{[I-j]}\hat{q}+\kappa}$$

$$= \hat{q} \sum_{j=0}^{J} \nu_{[I-j]} \frac{X_{[I-j],j}}{\nu_{[I-j]}\hat{q}+\kappa},$$

and hence

$$0 = \sum_{j=0}^{J} X_{[I-j],j} \left(1 - \frac{\hat{q}\nu_{[I-j]}}{\nu_{[I-j]}\hat{q}+\kappa}\right) = \sum_{j=0}^{J} X_{[I-j],j} \frac{\kappa}{\nu_{[I-j]}\hat{q}+\kappa} = \kappa \sum_{j=0}^{J} \hat{\gamma}_j = \kappa.$$

From $\kappa = 0$ and the constraint $\sum_{j=0}^{J} \hat{\gamma}_j = 1$ it follows with (2.18) that

$$\hat{q} = \sum_{j=0}^{J} \frac{X_{[I-j],j}}{\nu_{[I-j]}},$$

and from (2.6) we know that $\hat{q}$ also satisfies the ML equation (2.17). We conclude that the ML estimates for $q$ and for the $\gamma_j$'s under Model Assumptions 2.5 are given by

$$\hat{\gamma}_j = \frac{\frac{X_{[I-j],j}}{\nu_{[I-j]}}}{\sum_{l=0}^{J} \frac{X_{[I-l],l}}{\nu_{[I-l]}}} = \frac{\hat{\gamma}_j^{\mathrm{raw}}}{\hat{\beta}_J^{\mathrm{raw}}}, \qquad \hat{q} = \frac{\sum_{i=0}^{I} C_{i,t(i)}}{\sum_{i=0}^{I} \nu_i \hat{\beta}_{t(i)}} = \sum_{l=0}^{J} \frac{X_{[I-l],l}}{\nu_{[I-l]}} = \hat{\beta}_J^{\mathrm{raw}},$$

as in (2.1) and (2.7), which proves the claim. $\qquad \square$

As the parameter estimates are as in (2.1) and (2.7), we can use Estimate 2.3 for the estimation of the conditional MSEP, we just need to replace the estimates for $q\sigma_j^2$, $0 \le j \le J$, by some estimates for $\phi q \gamma_j$. As in Wüthrich & Merz (2008) we suggest to estimate $\phi$ with the help of Pearson residuals:

$$\hat{\phi} = \frac{1}{|\mathcal{D}_I| - (J+1)} \sum_{\substack{i+j \le I \\ j \le J}} \frac{(X_{i,j} - \nu_i \hat{\gamma}_j^{\mathrm{raw}})^2}{\nu_i \hat{\gamma}_j^{\mathrm{raw}}}, \qquad (2.20)$$

where $|\mathcal{D}_I|$ denotes the number of observations in $\mathcal{D}_I$. Replacing the estimates $\widehat{q\sigma}_j^2$ by $\hat{\phi}\hat{\gamma}_j^{\mathrm{raw}}$ in Estimate 2.3 we obtain the following estimators:

**Estimate 2.7 (MSEP, ODP model)** Under Model Assumptions 2.5 the conditional MSEP of the CC predictor $\hat{C}_{i,J}^{\mathrm{CC}}$ given in (2.9) is estimated by

$$\widehat{\mathrm{MSEP}}_{C_{i,J}|\mathcal{D}_I}\left(\hat{C}_{i,J}^{\mathrm{CC}}\right) = \widehat{\mathrm{PV}}_i + \widehat{\mathrm{PEE}}_i = \hat{\phi}\hat{R}_i^{\mathrm{CC}} + \nu_i^2 \hat{\phi} \sum_{j=I-i+1}^{J} \frac{X_{[I-j],j}}{(\nu_{[I-j]})^2},$$

where $\hat{\phi}$ and $\hat{R}_i^{\mathrm{CC}}$ are given in (2.20) and (2.8). For aggregated accident years the corresponding conditional MSEP is estimated by

$$\widehat{\mathrm{MSEP}}_{\sum_{i=I-J+1}^{I} C_{i,J}|\mathcal{D}_I}\left(\sum_{i=I-J+1}^{I} \hat{C}_{i,J}^{\mathrm{CC}}\right) = \sum_{i=I-J+1}^{I} \widehat{\mathrm{MSEP}}_{C_{i,J}|\mathcal{D}_I}\left(\hat{C}_{i,J}^{\mathrm{CC}}\right) + 2 \sum_{I-J+1 \le i < k \le I} \widehat{\mathrm{PEE}}_{i,k}$$

$$= \sum_{i=I-J+1}^{I} \widehat{\mathrm{MSEP}}_{C_{i,J}|\mathcal{D}_I}\left(\hat{C}_{i,J}^{\mathrm{CC}}\right) + 2 \sum_{I-J+1 \le i < k \le I} \nu_i \nu_k \hat{\phi} \sum_{j=I-i+1}^{J} \frac{X_{[I-j],j}}{(\nu_{[I-j]})^2}.$$

Analogously an estimator for the uncertainty in the one-year CDR is obtained from Estimate 2.4 by replacing $\widehat{q\sigma}_j^2$ by $\hat{\phi}\hat{\gamma}_j^{\mathrm{raw}}$, with $\hat{\phi}$ given in (2.20).

**Remark 2.8**

- Under Model Assumptions 2.5, if $\mu_i = \nu_i q$, $0 \leq i \leq I$, are considered as unknown parameters and if the $\mu_i$'s and $\gamma_j$'s are both estimated by ML estimation then it is well-known that the resulting development pattern is the CL development pattern (see e.g. Mack, 1991; Renshaw & Verrall, 1998). The result goes back to Hachemeister & Stanard (1975). For the estimation of the correlations in the CL development pattern one can therefore use the Fisher information matrix. This approach is also used by Alai *et al.* (2010) to derive an estimate for the conditional MSEP in the BF method in the case where the CL development pattern is used. However, in the CC method the premiums are assumed to be known and hence ignoring this information for the estimation of the development pattern is not consistent with the CC assumptions and thus we do not consider this case in more detail.

- Taylor (2002) remarks that the assumption of a constant dispersion parameter is often not realistic in practice. In Model Assumptions 2.5 the assumption of a constant dispersion parameter $\phi$ can be replaced by using a different dispersion parameter $\phi_j$ for each development year. The derivation of the ML estimates in this case is completely analogous and leads to the same estimates for the development pattern and the loss ratio. Further, the estimator (2.10) for $q\sigma_j^2$ gives in this case an estimator for $q\phi_j\gamma_j$. Therefore, if we use (2.10) for the estimation of the variance parameters $q\phi_j\gamma_j$, we obtain the estimators given in Estimates 2.3 and 2.4 for the conditional MSEPs. Hence, in this case we have the same formulas as in the distribution-free CC model 2.1.

## 3. Example

To illustrate the results, we apply the method to a data set from Wüthrich & Merz (2008), which is given in Appendix B. As explained below Remark 1.1, in a company premiums should be adjusted based on the information from the budgeting and planning process. As we do not have information from the budgeting and planning process we assume that the earned premiums given in Appendix B can be used as on-level premiums.

We calculate the reserves, process standard deviation, parameter estimation error and conditional MSEP for the CC model 2.1 and for the ODP model 2.5. Results from the CC model 2.1 are summarised in Table 2. The corresponding parameter estimates are given in Table 3 and the loss ratio estimate is given by $\hat{q} = \hat{\beta}_J^{\text{raw}} = 0.674$. The coefficient of variation of a reserve estimate $\hat{R}_i$ is estimated as

$$\widehat{\text{CoVa}}\left(\hat{R}_i\right) = \frac{\sqrt{\widehat{\text{MSEP}}_{R_i \mid \mathcal{D}_I}\left(\hat{R}_i\right)}}{\hat{R}_i} = \frac{\sqrt{\widehat{\text{MSEP}}_{C_{i,J} \mid \mathcal{D}_I}\left(\hat{C}_{i,J}\right)}}{\hat{R}_i}.$$

For the CC model 2.1 we also calculate the uncertainty in the CDR given by Estimate 2.4 and we compare these numbers with the conditional MSEP of the ultimate claim (see Estimate 2.3). The results are given in Table 4. Note that

$$\left(\frac{\widehat{\text{MSEP}}_{\text{CDR}^{(I+1)} \mid \mathcal{D}_I}(0)}{\widehat{\text{MSEP}}_{\sum_{i=I-J+1}^{I} C_{i,J} \mid \mathcal{D}_I}\left(\sum_{i=I-J+1}^{I} \hat{C}_{i,J}\right)}\right)^{\frac{1}{2}} = 89.4\,\%,$$

which means that most of the uncertainty in the prediction of the ultimate claim amount is contained in the next accounting year. As stated in Merz & Wüthrich (2008), this is to be expected for a triangle with short-tailed development. In the claims triangle of our example the development is

**Table 2.** Results CC model 2.1: reserves, process standard deviation ($PV^{1/2}$), parameter estimation error (PEE), conditional MSEP and coefficient of variation (CoVa) according to Estimate 2.3.

| $i$ | Reserves | $PV^{1/2}$ | $PEE^{1/2}$ | $MSEP^{1/2}$ | CoVa (%) |
|---|---|---|---|---|---|
| 1 | 15,209 | 175 | 172 | 245 | 1.6 |
| 2 | 25,619 | 684 | 486 | 840 | 3.3 |
| 3 | 35,874 | 2,596 | 1,481 | 2,989 | 8.3 |
| 4 | 90,234 | 7,556 | 3,836 | 8,474 | 9.4 |
| 5 | 166,584 | 29,088 | 13,300 | 31,984 | 19.2 |
| 6 | 314,665 | 67,560 | 28,248 | 73,227 | 23.3 |
| 7 | 528,056 | 80,012 | 32,234 | 86,261 | 16.3 |
| 8 | 1,200,821 | 152,676 | 55,740 | 162,533 | 13.5 |
| 9 | 4,240,563 | 393,823 | 135,844 | 416,594 | 9.8 |
| Total | 6,617,625 | 436,215 | 201,730 | 480,602 | 7.3 |

**Table 3.** CC model 2.1: parameter estimates (2.2) and (2.10).

| $j$ | $\hat{\gamma}_j^{\mathrm{raw}}$ (%) | $\widehat{q\sigma^2}_j$ |
|---|---|---|
| 0 | 39.49 | 9,760 |
| 1 | 19.58 | 8,585 |
| 2 | 4.67 | 1,172 |
| 3 | 1.51 | 132 |
| 4 | 1.01 | 251 |
| 5 | 0.49 | 52 |
| 6 | 0.37 | 3.5 |
| 7 | 0.08 | 0.45 |
| 8 | 0.08 | 0.03 |
| 9 | 0.10 | 0.002 |

very short. Most of the claims evolve in the first two development years. This can be observed when considering the normalised development pattern $\hat{\gamma}_j$ given in Table 5.

The results obtained from the ODP model 2.5 are given in Table 6. As shown in section 2.2, the reserves and development pattern are the same as in the CC model 2.1. In the ODP model 2.5 we have a constant dispersion parameter $\phi$ for all development years. The estimates $\hat{\phi}\hat{\gamma}_j^{\mathrm{raw}}$ for the variance parameters $\phi q\gamma_j$ should be compared with the corresponding estimates $\widehat{q\sigma^2}_j$ for $q\sigma_j^2$ in the CC model 2.1 (see Tables 3 and 7). As $\hat{\phi}\hat{\gamma}_j^{\mathrm{raw}} >> \widehat{q\sigma^2}_j$ for late development years (see Tables 3 and 7), the process standard deviation and the parameter estimation error in early accident years are much higher in this model compared with the CC model 2.1 (see Tables 2 and 6). However, for the total over all accident years the process standard deviations in the CC model 2.1 (Table 2) are higher than in the ODP model 2.5 (Table 6). As stated in Remark 2.8, the assumption of a constant dispersion parameter is often not appropriate. With the estimates $\widehat{q\sigma^2}_j$ given in (2.10) we define

$$\hat{\phi}_j = \frac{\widehat{q\sigma^2}_j}{\hat{\gamma}_j^{\mathrm{raw}}}, \quad j = 0, \dots, J. \tag{3.1}$$

Note that $\hat{\phi}_j$ gives an estimate for a dispersion parameter $\phi_j$ that varies between development years. In our example these dispersion parameters are given in Table 8. The estimate $\hat{\phi}$ of the dispersion

252

**Table 4.** CC model 2.1: uncertainty in the CDR according to Estimate 2.4 and $\widehat{\mathrm{MSEP}}^{1/2}_{\mathrm{CDR}_i^{(I+1)} | \mathcal{D}_I}(0) / \widehat{\mathrm{MSEP}}^{1/2}_{C_{i,J} | \mathcal{D}_I}\left(\hat{C}_{i,J}\right)$ for comparison with Estimate 2.3.

| $i$ | $\widehat{\mathrm{MSEP}}^{1/2}_{\mathrm{CDR}_i^{(I+1)} | \mathcal{D}_I}(0)$ | % of total uncertainty |
|---|---|---|
| 1 | 245 | 100 |
| 2 | 813 | 97 |
| 3 | 2,886 | 97 |
| 4 | 7,943 | 94 |
| 5 | 30,845 | 96 |
| 6 | 66,215 | 90 |
| 7 | 48,072 | 56 |
| 8 | 138,473 | 85 |
| 9 | 382,113 | 92 |
| Total | 429,567 | 89.4 |

**Table 5.** Normalised development pattern given in (2.1) and CL development pattern.

| $j$ | $\hat{\gamma}_j$ (%) | $\hat{\gamma}_j^{\mathrm{CL}}$ (%) |
|---|---|---|
| 0 | 58.62 | 58.96 |
| 1 | 29.06 | 29.04 |
| 2 | 6.94 | 6.84 |
| 3 | 2.24 | 2.17 |
| 4 | 1.50 | 1.44 |
| 5 | 0.72 | 0.69 |
| 6 | 0.54 | 0.51 |
| 7 | 0.12 | 0.11 |
| 8 | 0.11 | 0.10 |
| 9 | 0.15 | 0.14 |

**Table 6.** Results ODP model 2.5: reserves, process standard deviation ($\mathrm{PV}^{1/2}$), parameter estimation error (PEE), conditional MSEP and coefficient of variation (CoVa) according to Estimate 2.7.

| $i$ | Reserves | $\mathrm{PV}^{1/2}$ | $\mathrm{PEE}^{1/2}$ | $\mathrm{MSEP}^{1/2}$ | CoVa (%) |
|---|---|---|---|---|---|
| 1 | 15,209 | 18,130 | 17,780 | 25,393 | 167.0 |
| 2 | 25,619 | 23,530 | 20,246 | 31,041 | 121.2 |
| 3 | 35,874 | 27,843 | 21,489 | 35,172 | 98.0 |
| 4 | 90,234 | 44,159 | 27,869 | 52,218 | 57.9 |
| 5 | 166,584 | 60,000 | 33,988 | 68,958 | 41.4 |
| 6 | 314,665 | 82,463 | 40,870 | 92,035 | 29.2 |
| 7 | 528,056 | 106,826 | 47,701 | 116,992 | 22.2 |
| 8 | 1,200,821 | 161,092 | 63,570 | 173,182 | 14.4 |
| 9 | 4,240,563 | 302,725 | 108,952 | 321,734 | 7.6 |
| Total | 6,617,625 | 378,170 | 290,414 | 476,815 | 7.2 |

parameter in the ODP model 2.5 is given by $\hat{\phi} = 21,611$ (see (2.20)). The variability of the estimates $\hat{\phi}_j$ given in Table 8 also suggests that the assumption of a constant dispersion parameter is too restrictive.

**Table 7.** ODP model 2.5: parameter estimates (2.2) and $\hat{\phi}$ as in (2.20).

| $j$ | $\hat{\gamma}_j^{\text{raw}}$ (%) | $\hat{\phi}\hat{\gamma}_j^{\text{raw}}$ |
|---|---|---|
| 0 | 39.49 | 8,535 |
| 1 | 19.58 | 4,231 |
| 2 | 4.67 | 1,010 |
| 3 | 1.51 | 326 |
| 4 | 1.01 | 219 |
| 5 | 0.49 | 105 |
| 6 | 0.37 | 79 |
| 7 | 0.08 | 17 |
| 8 | 0.08 | 16 |
| 9 | 0.10 | 22 |

**Table 8.** Parameter estimates (3.1) for comparison with $\hat{\phi} = 21,611$ calculated from (2.20).

| $j$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\hat{\phi}_j$ | 24,712 | 43,854 | 25,071 | 8,772 | 24,824 | 10,746 | 943 | 573 | 40 | 2 |

**Table 9.** CL results according to Mack (1993).

| $i$ | Reserves | $PV^{1/2}$ | $PEE^{1/2}$ | $MSEP^{1/2}$ | CoVa (%) |
|---|---|---|---|---|---|
| 1 | 15,126 | 191 | 187 | 267 | 1.8 |
| 2 | 26,257 | 742 | 535 | 914 | 3.5 |
| 3 | 34,538 | 2,669 | 1,493 | 3,058 | 8.9 |
| 4 | 85,302 | 6,832 | 3,392 | 7,628 | 8.9 |
| 5 | 156,494 | 30,478 | 13,517 | 33,341 | 21.3 |
| 6 | 286,121 | 68,212 | 27,286 | 73,467 | 25.7 |
| 7 | 449,167 | 80,077 | 29,675 | 85,398 | 19.0 |
| 8 | 1,043,242 | 126,960 | 43,903 | 134,337 | 12.9 |
| 9 | 3,950,815 | 389,783 | 129,769 | 410,817 | 10.4 |
| Total | 6,047,061 | 424,379 | 185,024 | 462,960 | 7.7 |

For comparison we further give the results obtained with Mack's CL method (Mack, 1993) in Table 9. These results can be found in Wüthrich & Merz (2008). In Table 5 we compare the CL development pattern to the normalised development pattern given in (2.1). In this example the development patterns are very similar.

The CL reserves are substantially lower than the CC reserves (see Tables 2 and 9). From the data triangle, we observe that the diagonal values $C_{i,t(i)}$ are rather low for accident years 6–8. The relative differences in the reserves of CC and CL are also the highest for these accident years.

As in Wüthrich & Merz (2008) we plot the individual loss ratios $C_{i,t(i)}/(\nu_i\hat{\beta}_{t(i)})$ and compare it to the estimated loss ratio $\hat{q} = 0.674$ (see Figure 1). The picture also shows that the diagonal claims $C_{i,t(i)}$ in accident years 7 and 8 are rather low. On the other hand, it is possible that the premium level was increased in these accident years. In a company, information from the budgeting and planning
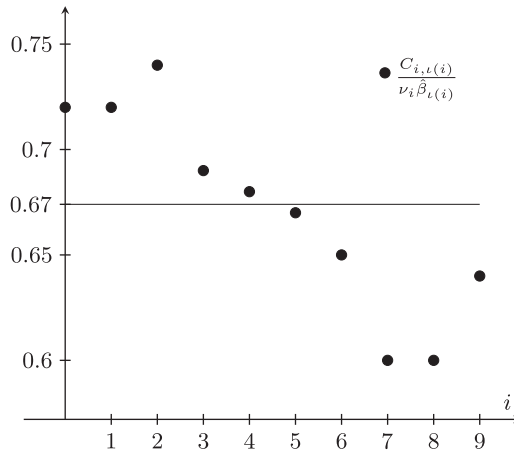
**Figure 1.** Individual loss ratios $\frac{C_{i,t(i)}}{\nu_i \hat{\beta}_{t(i)}}$ and $\hat{q} = 0.674$.

process should be used to adjust the premiums in this case. The decision whether one rather relies on the given diagonal values (by using CL) or whether one trusts the premium information more (by using CC), needs to be based on further information on the data and the quality of the on-level premiums.

## 4. Conclusion

In this paper we studied the CC method in a stochastic framework. We derive parameter estimates for the development pattern and the loss ratio within the considered stochastic model. The resulting estimates allow to rewrite the CC reserve in a simple form as given in (2.8). Moreover, we derive closed formulas for the conditional MSEP of the ultimate claim as well as for the uncertainties in the one-year CDR. Although the CC method is well established in practice, such formulas cannot be found in current literature (see Dal Moro & Lo, 2014). We further remark that most contributions on the CDR in current literature are in the context of the CL method.

In this paper the relations between the ultimate uncertainties and the uncertainties on a one-year horizon are highlighted, too. In the numerical example it is observed that most of the uncertainty lies in the next accounting year. Such observations are typical for short-tailed lines of business (see Merz & Wüthrich, 2008).

### 4.1. Limitations and possible extensions

In our stochastic model we have assumed that on-level premiums are known. In practice, such on-level premiums are not always available, and the calculation of on-level premiums can be afflicted with uncertainties. As an alternative to on-level premiums, Bühlmann & Straub (1983) also allow for varying loss ratio estimates given by

$$\hat{q}_i = \frac{\sum_{i \in S_i} C_{i,t(i)}}{\sum_{i \in S_i} \nu_i \hat{\beta}_{t(i)}},$$

where the set $S_i \subset \{0,1,\ldots,I\}$ needs to be chosen in such a way that the years in $S_i$ are comparable with accident year $i$ in terms of the expected loss ratio. This means that only the premiums $v_k$, $k \in S_i$, need to be adjusted to the same level. However, if non-constant loss ratios are used, then the estimation of the development pattern and the prediction uncertainties needs to be adjusted. The simple representation of the CC reserves given in (2.8) does then no longer hold. Such adjustments and a corresponding stochastic model could be addressed in future research.

Further, we showed in Remark 2.2 how an externally given tail development can be incorporated in the estimation of the reserves. Future research could investigate how such tail development estimators can be obtained and how the incorporation of such tail developments affects prediction uncertainties.

## Acknowledgements

## References

Alai, D.H., Merz, M. & Wüthrich, M.V. (2010). Prediction uncertainty in the Bornhuetter-Ferguson claims reserving method: revisited. *Annals of Actuarial Science*, **5**(1), 7–17.

Bornhuetter, R.L. & Ferguson, R.E. (1972). Proceedings of the Casualty Actuarial Society. The actuary and IBNR. *Proc. Cas*, **LIX**, 181–195.

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.

Bühlmann, H. & Straub, E. (1983). *Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod and Complementary Loss Ratio*. International Summer School of the Swiss Association of Actuaries, Leysin, Switzerland.

Dal Moro, E. & Lo, J. (2014). An industry question: the ultimate and one-year reserving uncertainty for different non-life reserving methodologies. *ASTIN Bulletin*, **44**(3), 495–499.

England, P.D. & Verrall, R.J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8**(3), 443–518.

FINMA (2006). Technisches Dokument zum Swiss Solvency Test, Version 02/10/2006, technical report, Bundesamt für Privatversicherungen, Bern, Switzerland.

Hachemeister, C. & Stanard, J. (1975). IBNR claims count estimation with static lag functions. ASTIN Colloquium, Portimao, Portugal. September 30 – October 3, 1975.

Mack, T. (1991). A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin*, **21**(1), 93–109.

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

Mack, T. (2006). Parameter estimation for Bornhuetter/Ferguson. *CAS Forum*, (Fall 2006), 141–157.

Mack, T. (2008). The prediction error of Bornhuetter/Ferguson. *ASTIN Bulletin*, **38**(1), 87–103.

Merz, M. & Wüthrich, M.V. (2008). Modelling the claims development result for solvency purposes. *CAS Forum*, (Fall 2008), 542–568.

Renshaw, A.E. & Verrall, R.J. (1998). A stochastic model for the chain ladder method. *British Actuarial Journal*, **14**, 903–923.

Saluz, A., Gisler, A. & Wüthrich, M.V. (2011). Development pattern and prediction error for the stochastic Bornhuetter-Ferguson claims reserving method. *ASTIN Bulletin*, **41**(2), 279–313.

Straub, E. (1988). *Non-Life Insurance Mathematics*. Springer-Verlag, Association of Swiss Actuaries, Zürich.

Taylor, G.C. (2002). Written discussion on the paper "Stochastic claims reserving in general insurance" by P. D. England and R. J. Verrall. *British Actuarial Journal*, **8**(3), 540–542.

Wüthrich, M.V. & Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. Wiley, Chichester.

## Appendix A

## Derivations of the estimators

### Derivation of Estimate 2.3

For the conditional MSEP of accident year $i \geq I - J$ we have

$$\text{MSEP}_{C_{i,J} \mid \mathcal{D}_I}\left(\hat{C}_{i,J}^{\text{CC}}\right) = E\left[\left(C_{i,J} - \hat{C}_{i,J}^{\text{CC}}\right)^2 \middle| \mathcal{D}_I\right]$$

$$\overset{(2.9)}{=} E\left[\left(C_{i,I-i} + \sum_{j=I-i+1}^{J} X_{i,j} - C_{i,I-i} - \hat{R}_{i,J}^{\text{CC}}\right)^2 \middle| \mathcal{D}_I\right]$$

$$\overset{(2.8)}{=} E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \nu_i q(1 - \beta_{I-i}) + \nu_i q(1 - \beta_{I-i}) - \nu_i\left(\hat{\beta}_J^{\text{raw}} - \hat{\beta}_{I-i}^{\text{raw}}\right)\right)^2 \middle| \mathcal{D}_I\right]$$

$$= \sum_{j=I-i+1}^{J} \text{Var}(X_{i,j}) + \nu_i^2\left(\hat{\beta}_J^{\text{raw}} - \hat{\beta}_{I-i}^{\text{raw}} - q(1 - \beta_{I-i})\right)^2 \tag{A.1}$$

$$+ 2\nu_i\left(q(1 - \beta_{I-i}) - \hat{\beta}_J^{\text{raw}} - \hat{\beta}_{I-i}^{\text{raw}}\right) E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \nu_i q(1 - \beta_{I-i})\right) \middle| \mathcal{D}_I\right]$$

$$= \underbrace{\sum_{j=I-i+1}^{J} \nu_i q \sigma_j^2}_{\text{PV}_i} + \underbrace{\nu_i^2\left(\sum_{j=I-i+1}^{J}\left(\hat{\gamma}_j^{\text{raw}} - q\gamma_j\right)\right)^2}_{\text{PEE}_i},$$

where the mixed term in the second line of equation (A.1) vanishes owing to the independence between $(X_{i,j})_{j = I-i+1, \dots, J}$ and $\mathcal{D}_I$. We estimate the process variance $\text{PV}_i$ by inserting the parameter estimates given in (2.10):

$$\widehat{\text{PV}}_i = \sum_{j=I-i+1}^{J} \nu_i \widehat{q\sigma^2}_j.$$

For the estimation of the parameter estimation error of accident year $i$ we observe that $\hat{\gamma}_j^{\text{raw}}$ is an unbiased estimator for $q\gamma_j$ (this follows from (2.2)). We approximate parameter estimation error by taking the average overall possible observations $\mathcal{D}_I$, that is, we consider the expected value

$$E[\text{PEE}_i] = \nu_i^2 \text{Var}\left( \sum_{j=I-i+1}^{J} \hat{\gamma}_j^{\text{raw}} \right).$$

Further, $\hat{\gamma}_j^{\text{raw}}$ depends only on the data from development year $j$ and therefore the $(\hat{\gamma}_j^{\text{raw}})_{j=0,\ldots,J}$ are independent. Moreover, we have

$$\text{Var}\left(\hat{\gamma}_j^{\text{raw}}\right) = \text{Var}\left(\frac{\sum_{l=0}^{I-j} X_{l,j}}{\nu_{[I-j]}}\right) = \frac{1}{\left(\nu_{[I-j]}\right)^2} \sum_{l=0}^{I-j} \text{Var}(X_{l,j})$$

$$= \frac{1}{\left(\nu_{[I-j]}\right)^2} \sum_{l=0}^{I-j} \nu_l q \sigma_j^2 = \frac{q\sigma_j^2}{\nu_{[I-j]}}. \qquad (A.2)$$

We therefore obtain

$$E[\text{PEE}_i] = \nu_i^2 \text{Var}\left( \sum_{j=I-i+1}^{J} \hat{\gamma}_j^{\text{raw}} \right) = \nu_i^2 \sum_{j=I-i+1}^{J} \text{Var}\left(\hat{\gamma}_j^{\text{raw}}\right) = \nu_i^2 \sum_{j=I-i+1}^{J} \frac{q\sigma_j^2}{\nu_{[I-j]}}.$$

Hence we estimate the parameter estimation error by

$$\widehat{\text{PEE}}_i = \nu_i^2 \sum_{j=I-i+1}^{J} \frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}},$$

which yields the estimator $\widehat{\text{MSEP}}_{C_{i,J}\,|\,\mathcal{D}_I}\left(\hat{C}_{i,J}^{\text{CC}}\right)$ given in Estimate 2.3.

Next we consider the conditional MSEP of aggregated accident years given by

$$\text{MSEP}_{\sum_{i=I-J+1}^{I} C_{i,J}\,|\,\mathcal{D}_I}\left( \sum_{i=I-J+1}^{I} \hat{C}_{i,J}^{\text{CC}} \right) = E\left[ \left( \sum_{i=I-J+1}^{I} \left(C_{i,J} - \hat{C}_{i,J}^{\text{CC}}\right) \right)^2 \Bigg| \mathcal{D}_I \right]$$

$$= \sum_{i=I-J+1}^{I} \text{MSEP}_{C_{i,J}\,|\,\mathcal{D}_I}(\hat{C}_{i,J}^{\text{CC}}) + 2 \sum_{i<k} E\left[ \left(C_{i,J} - \hat{C}_{i,J}^{\text{CC}}\right)\left(C_{k,J} - \hat{C}_{k,J}^{\text{CC}}\right) \Big| \mathcal{D}_I \right].$$

As in the case of a single accident year, we use (2.8) and (2.9) to rewrite

$$C_{i,J} - \hat{C}_{i,J}^{\text{CC}} = \sum_{j=I-i+1}^{J} X_{i,j} - \nu_i\left(\hat{\beta}_j^{\text{raw}} - \hat{\beta}_{I-i}^{\text{raw}}\right).$$

258

For the mixed terms $E\left[\left(C_{i,J}-\hat{C}_{i,J}^{\text{CC}}\right)\left(C_{k,J}-\hat{C}_{k,J}^{\text{CC}}\right)\Big|\mathcal{D}_I\right]$, $i<k$, we then obtain

$$E\left[\left(C_{i,J}-\hat{C}_{i,J}^{\text{CC}}\right)\left(C_{k,J}-\hat{C}_{k,J}^{\text{CC}}\right)\Big|\mathcal{D}_I\right]$$

$$= E\left[\left(\sum_{j=I-i+1}^{J}X_{i,j}-\nu_i\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\right)\left(\sum_{j=I-k+1}^{J}X_{k,j}-\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)\right)\Bigg|\mathcal{D}_I\right]$$

$$= E\left[\left(\sum_{j=I-i+1}^{J}X_{i,j}\right)\left(\sum_{j=I-k+1}^{J}X_{k,j}\right)\Bigg|\mathcal{D}_I\right] - E\left[\sum_{j=I-i+1}^{J}X_{i,j}\Bigg|\mathcal{D}_I\right]\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)$$

$$- \nu_i\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)E\left[\sum_{j=I-k+1}^{J}X_{k,j}\Bigg|\mathcal{D}_I\right] + \nu_i\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)$$

$$= \nu_i\nu_k q^2(1-\beta_{I-i})(1-\beta_{I-k}) - \nu_i q(1-\beta_{I-i})\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)$$

$$- \nu_i\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\nu_k q(1-\beta_{I-k}) + \nu_i\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right),$$

where we used the independence of the incremental claims $X_{i,j}$. This expression can be rewritten as follows

$$E\left[\left(C_{i,J}-\hat{C}_{i,J}^{\text{CC}}\right)\left(C_{k,J}-\hat{C}_{k,J}^{\text{CC}}\right)\Big|\mathcal{D}_I\right]$$

$$= \nu_i\nu_k q(1-\beta_{I-i})\left(q(1-\beta_{I-k})-\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)\right)$$

$$- \nu_i\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\left(q(1-\beta_{I-k})-\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)\right)$$

$$= \nu_i\nu_k\left(q(1-\beta_{I-i})-\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}\right)\right)\left(q(1-\beta_{I-k})-\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}\right)\right).$$

Therefore, we need to estimate for $i<k$

$$\text{PEE}_{i,k} = \nu_i\nu_k\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-i}^{\text{raw}}-q(1-\beta_{I-i})\right)\left(\hat{\beta}_J^{\text{raw}}-\hat{\beta}_{I-k}^{\text{raw}}-q(1-\beta_{I-k})\right)$$

$$= \nu_i\nu_k\left(\sum_{j=I-i+1}^{J}\left(\hat{\gamma}_j^{\text{raw}}-q\gamma_j\right)\right)\left(\sum_{l=I-k+1}^{J}\left(\hat{\gamma}_l^{\text{raw}}-q\gamma_l\right)\right).$$

We approximate $\text{PEE}_{i,k}$ by its expected value

$$E[\text{PEE}_{i,k}] = \nu_i\nu_k\text{Cov}\left(\sum_{j=I-i+1}^{J}\hat{\gamma}_j^{\text{raw}}, \sum_{l=I-k+1}^{J}\hat{\gamma}_l^{\text{raw}}\right)$$

$$= \nu_i\nu_k\sum_{j=I-i+1}^{J}\text{Var}\left(\hat{\gamma}_j^{\text{raw}}\right)$$

$$\overset{\text{(A.2)}}{=} \nu_i\nu_k\sum_{j=I-i+1}^{J}\frac{q\sigma_j^2}{\nu_{[I-j]}}, \quad i<k,$$

and estimate $\mathrm{PEE}_{i,k}$ by

$$\widehat{\mathrm{PEE}}_{i,k} = \nu_i \nu_k \sum_{j=I-i+1}^{J} \frac{\widehat{q\sigma^2_{\,j}}}{\nu_{[I-j]}}, \quad i < k.$$

Therefore we derived the formulas given in Estimate 2.3. $\qquad\square$

**Derivation of Estimate 2.4:**

From (2.14) and (2.15) we obtain for $I+1-J \leq i \leq I$

$$
\begin{aligned}
\mathrm{CDR}_i^{(I+1)} &= \hat{C}_{i,J}^{(I)} - \hat{C}_{i,J}^{(I+1)} \\
&= -X_{i,I+1-i} + \nu_i\left(\hat{\beta}_J^{(I)\mathrm{raw}} - \hat{\beta}_{I-i}^{(I)\mathrm{raw}} - \hat{\beta}_J^{(I+1)\mathrm{raw}} + \hat{\beta}_{I+1-i}^{(I+1)\mathrm{raw}}\right) \\
&= -X_{i,I+1-i} + \nu_i\left(\sum_{j=I+1-i}^{J} \hat{\gamma}_j^{(I)\mathrm{raw}} - \sum_{j=I+2-i}^{J} \hat{\gamma}_j^{(I+1)\mathrm{raw}}\right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\hat{\gamma}_j^{(I+1)\mathrm{raw}} &\overset{(2.16)}{=} \frac{X_{[I+1-j],j}}{\nu_{[I+1-j]}} = \nu_{[I-j]}\frac{X_{[I-j],j} + X_{I+1-j,j}}{\nu_{[I-j]}\nu_{[I+1-j]}} \\
&= \hat{\gamma}_j^{(I)\mathrm{raw}}\frac{\nu_{[I-j]}}{\nu_{[I+1-j]}} + \frac{X_{I+1-j}}{\nu_{[I+1-j]}}.
\end{aligned}
$$

Hence we obtain for $I+1-J \leq i \leq I$

$$
\begin{aligned}
\mathrm{CDR}_i^{(I+1)} &= -X_{i,I+1-i} + \nu_i\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} + \sum_{j=I+2-i}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}}\left(1 - \frac{\nu_{[I-j]}}{\nu_{[I+1-j]}}\right) - \frac{X_{I+1-j,j}}{\nu_{[I+1-j]}}\right)\right) \\
&= \nu_i q\gamma_{I+1-i} - X_{i,I+1-i} + \nu_i\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right) \\
&\quad + \nu_i \sum_{j=I+2-i}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}}\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}} - \frac{X_{I+1-j,j}}{\nu_{[I+1-j]}}\right) \\
&= \nu_i q\gamma_{I+1-i} - X_{i,I+1-i} + \nu_i\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right) \\
&\quad + \nu_i \sum_{j=I+2-i}^{J}\left(\left(\hat{\gamma}_j^{(I)\mathrm{raw}} - q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}} - \frac{X_{I+1-j,j} - \nu_{I+1-j}q\gamma_j}{\nu_{[I+1-j]}}\right).
\end{aligned}
\tag{A.3}
$$

Therefore we have for the conditional MSEP

$$
\begin{aligned}
\mathrm{MSEP}_{\mathrm{CDR}_i^{(I+1)}|\mathcal{D}_I}(0) &= E\left[\left(\mathrm{CDR}_i^{(I+1)}\right)^2\middle|\mathcal{D}_I\right] \\
&= \nu_i q\sigma^2_{I+1-i} + \nu_i^2\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right)^2 \\
&\quad + 2\nu_i^2\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right)\left(\sum_{j=I+2-i}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}} - q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}\right)
\end{aligned}
\tag{A.4}
$$

260

$$+\nu_i^2 E\left[\left(\sum_{j=I+2-i}^{J}\left(\left(\hat{\gamma}_j^{(I)\text{raw}}-q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}-\frac{X_{I+1-j,j}-\nu_{I+1-j}q\gamma_j}{\nu_{[I+1-j]}}\right)\right)^2\Bigg|\mathcal{D}_I\right], \qquad (A.5)$$

where we used that

$$0 = \nu_i E\left[\left(\nu_i q\gamma_{I+1-i}-X_{i,I+1-i}\right)\left(\hat{\gamma}_{I+1-i}^{(I)\text{raw}}-q\gamma_{I+1-i}\right)\Big|\mathcal{D}_I\right]$$

$$+\nu_i E\left[\left(\nu_i q\gamma_{I+1-i}-X_{i,I+1-i}\right)\left(\sum_{j=I+2-i}^{J}\left(\left(\hat{\gamma}_j^{(I)\text{raw}}-q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}\right)\right)\Bigg|\mathcal{D}_I\right]$$

$$-\nu_i E\left[\left(\nu_i q\gamma_{I+1-i}-X_{i,I+1-i}\right)\left(\sum_{j=I+2-i}^{J}\left(\frac{X_{I+1-j,j}-\nu_{I+1-j}q\gamma_j}{\nu_{[I+1-j]}}\right)\right)\Bigg|\mathcal{D}_I\right],$$

due to the independence of $X_{i,I+1-i}$ from $(X_{I+1-j,j})_{j=I+2-i,\dots J}$ and $\mathcal{D}_I$. For the conditional expected value given in (A.5) we have

$$E\left[\left(\sum_{j=I+2-i}^{J}\left(\left(\hat{\gamma}_j^{(I)\text{raw}}-q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}-\frac{X_{I+1-j,j}-\nu_{I+1-j}q\gamma_j}{\nu_{[I+1-j]}}\right)\right)^2\Bigg|\mathcal{D}_I\right]$$

$$= \sum_{j=I+2-i}^{J}\left(\hat{\gamma}_j^{(I)\text{raw}}-q\gamma_j\right)^2\frac{\nu_{I+1-j}^2}{\left(\nu_{[I+1-j]}\right)^2}+\sum_{j=I+2-i}^{J}E\left[\left(\frac{X_{I+1-j,j}-\nu_{I+1-j}q\gamma_j}{\nu_{[I+1-j]}}\right)^2\right]$$

$$+2\sum_{j<k}\left(\hat{\gamma}_j^{(I)\text{raw}}-q\gamma_j\right)\left(\hat{\gamma}_k^{(I)\text{raw}}-q\gamma_k\right)\frac{\nu_{I+1-j}\nu_{I+1-k}}{\nu_{[I+1-j]}\nu_{[I+1-k]}}. \qquad (A.6)$$

As in the case of the conditional MSEP of the ultimate claim, we approximate the unknown terms by their expected values. Because of the independence of the $(\hat{\gamma}_j^{\text{raw}})_{j=0,\dots J}$, the expected values of the terms in line (A.4) and (A.6) are 0. After replacing the unknown parameters by their estimates we get the following estimator

$$\widehat{\text{MSEP}}_{\text{CDR}_i^{(I+1)}\mid\mathcal{D}_I}(0)$$

$$= \nu_i\widehat{q\sigma^2}_{I+1-i}+\nu_i^2\widehat{\text{Var}}\left(\hat{\gamma}_{I+1-i}^{(I)\text{raw}}\right)+\nu_i^2\sum_{j=I+2-i}^{J}\left(\widehat{\text{Var}}\left(\hat{\gamma}_j^{(I)\text{raw}}\right)\frac{\nu_{I+1-j}^2}{\left(\nu_{[I+1-j]}\right)^2}+\frac{\widehat{\text{Var}}(X_{I+1-j,j})}{\left(\nu_{[I+1-j]}\right)^2}\right),$$

where the variances $\text{Var}\left(\hat{\gamma}_j^{(I)\text{raw}}\right)$ and $\text{Var}(X_{I+1-j,j})$ are estimated by inserting the parameter estimates for $q\sigma_j^2$ given in (2.10), that is:

$$\widehat{\text{Var}}\left(\hat{\gamma}_j^{(I)\text{raw}}\right)=\frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}} \quad \text{and} \quad \widehat{\text{Var}}(X_{I+1-j,j})=\nu_{I+1-j}\widehat{q\sigma^2}_j.$$

Therefore we arrive at the estimator

$$\widehat{\text{MSEP}}_{\text{CDR}_i^{(I+1)}\mid\mathcal{D}_I}(0) = \nu_i\widehat{q\sigma^2}_{I+1-i}+\nu_i^2\frac{\widehat{q\sigma^2}_{I+1-i}}{\nu_{[i-1]}}+\nu_i^2\sum_{j=I+2-i}^{J}\left(\frac{\widehat{q\sigma^2}_j}{\nu_{[I-j]}}\frac{\nu_{I+1-j}^2}{\left(\nu_{[I+1-j]}\right)^2}+\frac{\nu_{I+1-j}\widehat{q\sigma^2}_j}{\left(\nu_{[I+1-j]}\right)^2}\right)$$

$$= \nu_i\widehat{q\sigma^2}_{I+1-i}\frac{\nu_{[i]}}{\nu_{[i-1]}}+\nu_i^2\sum_{j=I+2-i}^{J}\widehat{q\sigma^2}_j\frac{\nu_{I+1-j}}{\nu_{[I-j]}\nu_{[I+1-j]}},$$

as given in Estimate 2.4.

261

Similarly, we have for aggregated accident years

$$\mathrm{MSEP}_{\mathrm{CDR}^{(I+1)}\,|\,\mathcal{D}_I}(0) = \sum_{i=I-J+1}^{I} \mathrm{MSEP}_{\mathrm{CDR}_i^{(I+1)}\,|\,\mathcal{D}_I}(0) + 2\sum_{i<k} E\left[\mathrm{CDR}_i^{(I+1)}\mathrm{CDR}_k^{(I+1)}\Big|\mathcal{D}_I\right].$$

With the representation of $\mathrm{CDR}_i^{(I+1)}$ given in equation (A.3), it follows that for $i<k$

$$E\left[\mathrm{CDR}_i^{(I+1)}\mathrm{CDR}_k^{(I+1)}\Big|\mathcal{D}_I\right]$$

$$= E\left[(\nu_i q\gamma_{I+1-i} - X_{i,I+1-i})\nu_k\left(-\frac{X_{i,I+1-i}-\nu_i q\gamma_{I+1-i}}{\nu_{[i]}}\right)\Big|\mathcal{D}_I\right]$$

$$+ \nu_i\nu_k\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right)\left(\hat{\gamma}_{I+1-k}^{(I)\mathrm{raw}} - q\gamma_{I+1-k}\right)$$

$$+ \nu_i\nu_k\left(\hat{\gamma}_{I+1-i}^{(I)\mathrm{raw}} - q\gamma_{I+1-i}\right)\sum_{j=I+2-k}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}} - q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}$$

$$+ \nu_i\nu_k\sum_{j=I+2-i}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}} - q\gamma_j\right)\frac{\nu_{I+1-j}}{\nu_{[I+1-j]}}\left(\hat{\gamma}_{I+1-k}^{(I)\mathrm{raw}} - q\gamma_{I+1-k}\right)$$

$$+ \nu_i\nu_k\sum_{j=I+2-i}^{J}\sum_{l=I+2-k}^{J}\left(\hat{\gamma}_j^{(I)\mathrm{raw}} - q\gamma_j\right)\left(\hat{\gamma}_l^{(I)\mathrm{raw}} - q\gamma_l\right)\frac{\nu_{I+1-j}\nu_{I+1-l}}{\nu_{[I+1-j]}\nu_{[I+1-l]}}$$

$$+ \nu_i\nu_k\sum_{j=I+2-i}^{J}\frac{\nu_{I+1-j}}{\left(\nu_{[I+1-j]}\right)^2}q\sigma_j^2.$$

Further we calculate

$$E\left[(\nu_i q\gamma_{I+1-i} - X_{i,I+1-i})\nu_k\left(-\frac{X_{i,I+1-i}-\nu_i q\gamma_{I+1-i}}{\nu_{[i]}}\right)\Big|\mathcal{D}_I\right] = \nu_k\frac{\nu_i}{\nu_{[i]}}q\sigma_{I+1-i}^2.$$

The mixed term $E\left[\mathrm{CDR}_i^{(I+1)}\mathrm{CDR}_k^{(I+1)}\Big|\mathcal{D}_I\right]$, $i<k$, is then approximated by its unconditional expected value which yields

$$\nu_k\frac{\nu_i}{\nu_{[i]}}q\sigma_{I+1-i}^2 + \nu_i\nu_k\frac{\nu_i q\sigma_{I+1-i}^2}{\nu_{[i]}\nu_{[i-1]}} + \nu_i\nu_k\sum_{j=I+2-i}^{J}\frac{q\sigma_j^2\nu_{I+1-j}^2}{\nu_{[I-j]}\left(\nu_{[I+1-j]}\right)^2} + \nu_i\nu_k\sum_{j=I+2-i}^{J}\frac{\nu_{I+1-j}}{\left(\nu_{[I+1-j]}\right)^2}q\sigma_j^2$$

$$= \nu_i\nu_k\left(\frac{q\sigma_{I+1-i}^2}{\nu_{[i-1]}} + \sum_{j=I+2-i}^{J}\frac{q\sigma_j^2\nu_{I+1-j}}{\nu_{[I-j]}\nu_{[I+1-j]}}\right),$$

where we used the independence of the $(\hat{\gamma}_j^{\mathrm{raw}})_{j=0,\dots J}$. After inserting the parameter estimates we arrive at the formula given in Estimate 2.4. $\qquad\square$

# Appendix B

## Data

**Table B.1.** Cumulative payments $C_{i,j}$ and premiums $\nu_i$.

| i/j | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | $\nu_i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 5,946,975 | 9,668,212 | 10,563,929 | 10,771,690 | 10,978,394 | 11,040,518 | 11,106,331 | 11,121,181 | 11,132,310 | 11,148,124 | 15,473,558 |
| 1 | 6,346,756 | 9,593,162 | 10,316,383 | 10,468,180 | 10,536,004 | 10,572,608 | 10,625,360 | 10,636,546 | 10,648,192 | | 14,882,436 |
| 2 | 6,269,090 | 9,245,313 | 10,092,366 | 10,355,134 | 10,507,837 | 10,573,282 | 10,626,827 | 10,635,751 | | | 14,456,039 |
| 3 | 5,863,015 | 8,546,239 | 9,268,771 | 9,459,424 | 9,592,399 | 9,680,740 | 9,724,068 | | | | 14,054,917 |
| 4 | 5,778,885 | 8,524,114 | 9,178,009 | 9,451,404 | 9,681,692 | 9,786,916 | | | | | 14,525,373 |
| 5 | 6,184,793 | 9,013,132 | 9,585,897 | 9,830,796 | 9,935,753 | | | | | | 15,025,923 |
| 6 | 5,600,184 | 8,493,391 | 9,056,505 | 9,282,022 | | | | | | | 14,832,965 |
| 7 | 5,288,066 | 7,728,169 | 8,256,211 | | | | | | | | 14,550,359 |
| 8 | 5,290,793 | 7,648,729 | | | | | | | | | 14,461,781 |
| 9 | 5,675,568 | | | | | | | | | | 15,210,363 |