CREDIBILITY CLAIMS RESERVING WITH STOCHASTIC DIAGONAL EFFECTS

BY

HANS BÜHLMANN AND FRANCO MORICONI

ABSTRACT

An interesting class of stochastic claims reserving methods is given by the models with conditionally independent loss increments (CILI), where the incremental losses are conditionally independent given a risk parameter $\Theta_{i,j}$ depending on both the accident year i and the development year j. The Bühlmann–Straub credibility reserving (BSCR) model is a particular case of a CILI model where the risk parameter is only depending on i. We consider CILI models with additive diagonal risk (ADR), where the risk parameter is given by the sum of two components, one depending on the accident year i and the other depending on the calendar year $t = i + j$. The model can be viewed as an extension of the BSCR model including random diagonal effects, which are often declared to be important in loss reserving but rarely are specifically modeled. We show that the ADR model is tractable in closed form, providing credibility formulae for the reserve and the mean square error of prediction (MSEP). We also derive unbiased estimators for the variance parameters which extend the classical Bühlmann–Straub estimators. The results are illustrated by a numerical example and the estimators are tested by simulation. We find that the inclusion of random diagonal effects can be significant for the reserve estimates and, especially, for the evaluation of the MSEP. The paper is written with the purpose of illustrating the role of stochastic diagonal effects. To isolate these effects, we assume that the development pattern is given. In particular, our MSEP values do not include the uncertainty due to the estimation of the development pattern.

KEYWORDS

Stochastic claims reserving, dependence structure, diagonal effects, credibility theory, variance estimators.

1. INTRODUCTION

1.1. The claims reserving problem

In the standard setup for claims reserving, we have the rectangle of incremental claims $X_{i,j}$ where $i = 0, 1, \ldots, I$ is the index of the accident year (AY) and

Astin Bulletin 45(2), 309-353. doi: 10.1017/asb.2015.3

$j = 0, 1, \ldots, J$ is the index of the development year (DY). The index $t = i + j = 0, 1, \ldots, I + J$ denotes the calendar year (CY). We assume that all claims are settled after DY J where $J \leq I$.

Let us consider the sets

$$\mathcal{D}_I = \{X_{i,j};\ i + j \leq I,\ j \leq J\}, \qquad \mathcal{D}_I^c = \{X_{i,j};\ i + j > I,\ i \leq I,\ j \leq J\}.$$

At time I the random variables $X_{ij} \in \mathcal{D}_I$ are *observed* and the random variables $X_{ij} \in \mathcal{D}_I^c$ — the outstanding claims — must be *predicted*. Denoting by $\widehat{X}_{i,j}$, the prediction of $X_{i,j} \in \mathcal{D}_I^c$ we then have the *reserve estimates*

$$\widehat{R}_i = \sum_{j=I-i+1}^{J} \widehat{X}_{i,j} : \quad \text{reserve estimate for accident year } i = I - J + 1, \ldots, I,$$

$$\widehat{R} = \sum_{i=I-J+1}^{I} \widehat{R}_i : \quad \text{total reserve estimate.}$$

Such predictions — including a measure of precision — can only be made based on a model for the random variables $\{X_{i,j};\ 0 \leq i \leq I,\ 0 \leq j \leq J\}$.

1.2. Conditionally independent loss increments (CILI) models

1.2.1. *A class of CILI models.* We consider a class of claims reserving models characterized by the following essential hypothesis:

- The incremental losses $X_{i,j}$ $(0 \leq i \leq I,\ 0 \leq j \leq J)$ are generated by a two level stochastic mechanism

$$X_{i,j} = w_{i,j}\, \Theta_{i,j} + \sqrt{w_{i,j}}\, \sigma\, \varepsilon_{i,j}, \tag{1}$$

  where
  - the random variable $\Theta_{i,j}$ is a *risk parameter* intrinsic to $X_{i,j}$; it describes the risk characteristics of the "cell" (i, j), i.e. of development year j belonging to accident year i.
  - the random variable $\varepsilon_{i,j}$ is an *observation error* not characteristic of the cell (i, j), i.e. it is pure noise.
  - the observation errors $\varepsilon_{i,j}$ are independent. Moreover ε-variables are independent of Θ-variables.
  - the *deterministic weights* $w_{i,j} > 0$ are chosen in such a way that:

$$\begin{aligned} &\mathbf{E}(\Theta_{i,j}) = \mu_0, && \mathbf{Var}(\Theta_{i,j}) = \sigma_\Theta^2, \\ &\mathbf{E}(\varepsilon_{i,j}) = 0, && \mathbf{Var}(\varepsilon_{i,j}) = 1. \end{aligned} \tag{2}$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

By the independence of the observation errors, the loss increments $X_{i,j}$ of accident year i are conditionally independent, given the Θ-variables; hence we say that (1) provides a class of "CILI models".

The existence of weights satisfying (2) is not automatic. In particular, (1) and (2) imply

$$\mathbf{E}\big[\mathbf{Var}(X_{i,j}|\Theta_{i,j})\big] = \frac{\sigma^2}{\mu_0}\,\mathbf{E}\big[\mathbf{E}(X_{i,j}|\Theta_{i,j})\big],$$

which can be interpreted as a "conditionally overdispersed Poisson" assumption (provided $\sigma^2 \geq \mu_0$).

The deterministic weights $w_{i,j} = \mathbf{E}(X_{i,j})/\mu_0$ are used to model *deterministic effects*.

i. The most common form is

$$w_{i,j} = a_i\,\gamma_j, \qquad 0 \leq i \leq I,\ 0 \leq j \leq J,$$

with the interpretation
- $a_i > 0$: *a priori estimate* for expected ultimate loss $\mathbf{E}\big(\sum_{j=0}^{J} X_{i,j}\big)$,
- $\gamma_j > 0$: *development quota* $\big(\sum_{j=0}^{J}\gamma_j = 1\big)$,
- μ_0 is called the *correction factor*.

For $\mu_0 = 1$, this factorization is the basic idea of the Bornhuetter–Ferguson reserving method (Bornhuetter and Ferguson, 1972). For μ_0 estimated from the data it is the basis of the Cape Cod reserving method (Bühlmann, 1983).

ii. Another form is

$$w_{i,j} = a_i\,\delta_{i+j}\,\gamma_j, \qquad 0 \leq i \leq I,\ 0 \leq j \leq J,$$

with a_i and γ_j as under (i) but with an additional factor $\delta_t = \delta_{i+j}$ for a *deterministic* diagonal effect in calendar year $t = i + j$. This approach is found in Jessen and Rietdorf (2011).

In this paper, we focus on the probabilistic structure of $\Theta_{i,j}$ $(0 \leq i \leq I,\ 0 \leq j \leq J)$ variables which represent the *stochastic effects* of the reserving model. We therefore work with the most simple form of the $w_{i,j}$-weights, $w_{i,j} = a_i\,\gamma_j$, and assume both the *a priori* values $(a_i)_{i=0,\ldots,I}$ and the development quotas $(\gamma_j)_{j=0,\ldots,J}$ as given quantities.

Remark. The CILI-class characterized here is not the most general one. We consider models where the conditioning random variables $\Theta_{i,j}$ are risk parameters having the same distribution (or at least the same first two moments). In the credibility-based additive loss reserving model introduced in Wüthrich and Merz (2012), for example, we have, similarly as in (1),

$$X_{i,j} = a_i\,\Theta_j + \sqrt{a_i}\,\sigma(\Theta_j)\,\varepsilon_{i,j}\,.$$

In this model, the random variables $(\Theta_j)_{j=0,\dots,J}$ are also independent but do no longer have the same distribution. This is also evident by their interpretation as normalized random cashflows.

1.2.2. *The* Θ*-structure.*

The non-informative case. The case of a CILI model where $\Theta_{i,j}$ $(0 \le i \le I,\ 0 \le j \le J)$ are *independent* (in addition to having identical first and second moments) is called non-informative. Looking at (1), we see that not only the ε-variables but also the Θ-variables can be understood as pure noise in this case. This means that the parameters σ^2 and σ_Θ^2 become *stochastically indistinguishable*.

For $X_{i,j} \in \mathcal{D}_I^c$, we may still predict

$$\widehat{X}_{i,j} = \widehat{\mu}_0\, w_{i,j} \quad \text{where} \quad \widehat{\mu}_0 = \frac{\sum_{(i,j)\in\mathcal{D}_I} d_{i,j} \frac{X_{i,j}}{w_{i,j}}}{\sum_{(i,j)\in\mathcal{D}_I} d_{i,j}},$$

for some weights $d_{i,j} > 0$. But for

$$\mathrm{msep}_{X_{i,j}}(\widehat{\mu}_0\, w_{i,j}) = \mathbf{Var}(X_{i,j}) + w_{i,j}^2\, \mathbf{Var}(\widehat{\mu}_0),$$

we are stuck, since for the evaluation of the above right-hand side we need the variance components. We see that it is the *dependence structure* of the Θ-variables that allows us to predict $X_{i,j} \in \mathcal{D}_I^c$ and also to make statements about the prediction error.

Two important cases of dependence in the Θ-structure. Two basic informative cases are specified as follows.

Case A. Dependence on *accident year only*, i.e.

$$\Theta_{i,j} = \eta_i,\ 0 \le i \le I,\ 0 \le j \le J, \quad \text{and} \quad \{\eta_i;\ 0 \le i \le I\}\ \textit{independent}.$$

Case B. Dependence on *accident year* as well as *calendar year*, i.e.

$$\begin{aligned} &\Theta_{i,j} = \eta_i + \zeta_{i+j},\ 0 \le i \le I,\ 0 \le j \le J, \quad \text{and} \\ &\qquad \{\eta_i, \zeta_t;\ 0 \le i \le I,\ 0 \le t \le I+J\}\ \textit{independent}. \end{aligned} \tag{3}$$

Since in (3) the risk parameter is additively separated into an accident year parameter and a calendar year (diagonal) parameter, we call Case B “CILI model with additive diagonal risk (ADR)”. Of course Case A, which is a CILI model without diagonal risk, is contained in the ADR model.

2. The CILI model with additive diagonal risk

Summarizing all the previous considerations, the ADR CILI model can be characterized as follows.

Model assumptions

A1. Let $\boldsymbol{\Theta} := \{\eta_i, \zeta_{i+j};\ \ 0 \le i \le I,\ 0 \le j \le J\}$. There exist positive parameters $a_0, \ldots, a_I, \gamma_0, \ldots, \gamma_J$, and σ^2, with $\sum_{j=0}^{J} \gamma_j = 1$, such that for $0 \le i \le I$ and $0 \le j \le J$

$$\mathbf{E}(X_{ij}|\boldsymbol{\Theta}) = a_i\, \gamma_j\, (\eta_i + \zeta_{i+j}), \tag{4}$$

and

$$\mathbf{Var}(X_{ij}|\boldsymbol{\Theta}) = a_i\, \gamma_j\, \sigma^2\,. \tag{5}$$

A2. All η, ζ variables are independent, with

$$\begin{array}{lll} \mathbf{E}(\eta_i) = \mu_0, & \mathbf{Var}(\eta_i) = \tau^2, & 0 \le i \le I, \\ \mathbf{E}(\zeta_{i+j}) = 0, & \mathbf{Var}(\zeta_{i+j}) = \chi^2, & 0 \le i \le I,\ 0 \le j \le J\,. \end{array}$$

Assumption A1 can be written in the time series format:

A1^{bis}. For $0 \le i \le I$ and $0 \le j \le J$

$$X_{ij} = a_i\, \gamma_j\, (\eta_i + \zeta_{i+j}) + \sqrt{a_i\, \gamma_j}\, \sigma\, \varepsilon_{i,j}, \tag{6}$$

where all η, ζ, ε variables are independent, with $\mathbf{E}(\varepsilon_{ij}) = 0$, $\mathbf{Var}(\varepsilon_{ij}) = 1$.

As usual, we interpret a_i as the *a priori* estimate of ultimate loss of accident year i and γ_j as the development quota of development year j. Moreover we call:

η_i : *random effect of accident year i*,

ζ_{i+j} : *random effect of calendar year $t = i + j$* (sometimes also *random diagonal effect*).

Here the priors $(a_i)_{i=0,\ldots,I}$ and the development quotas $(\gamma_j)_{j=0,\ldots,J}$ are assumed to be given. The parameters $\mu_0, \sigma^2, \tau^2, \chi^2$ must be estimated from the data.

Remarks

- If in the previous assumptions we set $\chi^2 = 0$, the ADR model reduces to the BSCR model, which is an application to claims reserving of the classical Bühlmann–Straub credibility model, see Bühlmann and Gisler (2005). The BSCR model is treated in Section 4.5 of Wüthrich and Merz (2008).
- If we set $\chi^2 = \tau^2 = 0$, the ADR model reduces to the standard model for the additive loss reserving method (see Schmidt (2006), Wüthrich and Merz (2008) Section 8.3). Obviously in this case, the loss increments $X_{i,j}$ are *unconditionally* independent.
- Using a full Bayesian approach, ADR models are also treated in Shi *et al.* (2012), Wüthrich (2012) and Wüthrich (2013). Contrary to their approach, our way of proceeding is based on empirical linear Bayesian (credibility) methodology.

- We are interested in analyzing the effects of the introduction of the diagonal risk component separately from the effects of other sources of uncertainty, as e.g. the development pattern uncertainty. Therefore, we always assume in this paper that the development quotas γ_j are externally given. This allows for an evaluation of the reserve estimates and their prediction error in closed form. This would no longer be possible (see Saluz *et al.* (2014)) if one simultaneously also had to estimate the development quotas γ_j.
- Since $\mathbf{E}(\zeta_{i+j}) \equiv 0$, we assume that there is no trend in calendar year effects. Therefore, in principle, the model should be applied after any calendar year trend (e.g. economic inflation) has been removed from the data.

The innovative aspect of the ADR model is the inclusion of the calendar year effect in the random variable $\Theta_{i,t}$ which takes the form

$$\Theta_{i,t} = \eta_i + \zeta_t \quad \text{for the calendar year} \quad t = i + j,$$

thus separating it additively from the (random) accident year effect. Our assumptions imply the following correlation structure for the risk parameters

$$\frac{\mathbf{Cov}(\Theta_{it}, \Theta_{ks})}{\mathbf{Var}(\Theta_{it})} = \begin{cases} 1 & \text{if} \quad i = k,\ t = s, \\ \frac{\tau^2}{\tau^2+\chi^2} & \text{if} \quad i = k,\ t \neq s, \\ \frac{\chi^2}{\tau^2+\chi^2} & \text{if} \quad i \neq k,\ t = s, \\ 0 & \text{if} \quad i \neq k,\ t \neq s\,. \end{cases}$$

Hence, in the ADR model the correlation within the same accident year i is induced by η_i and is proportional to τ^2; the correlation within the same calendar year t is induced by ζ_t and is proportional to χ^2. It will be convenient to specify covariance properties using the indicator function

$$\mathbb{I}_{i,k} = \begin{cases} 1 & \text{if} \quad i = k, \\ 0 & \text{if} \quad i \neq k\,. \end{cases}$$

With this notation, the covariance structure for the risk parameters is given by

$$\mathbf{Cov}(\Theta_{it}, \Theta_{ks}) = \tau^2\,\mathbb{I}_{i,k} + \chi^2\,\mathbb{I}_{t,s}, \qquad 0 \leq i, k \leq I,\ 0 \leq t, s \leq I + J\,. \tag{7}$$

We want

i. to show that the ADR model is still tractable by techniques stemming from credibility theory and how stochastic diagonal effects can be incorporated.

ii. to find out whether the inclusion of stochastic diagonal effects substantially changes the estimates of the reserves and its MSEP.

To this aim it will be useful to consider Case A of dependence, i.e. a CILI model without random diagonal effects, as a "control model". In what follows the results obtained with the ADR model will be compared with the corresponding results provided by the BSCR model, as described in Wüthrich and Merz (2008).

3. Prediction

For $X_{ij} \in \mathcal{D}_I^c$, property (4) suggests to consider the predictor

$$\widehat{X}_{ij} = a_i\,\gamma_j\,(\widehat{\eta}_i + \widehat{\zeta}_{i+j})\,.$$

For all $X_{ij} \in \mathcal{D}_I^c$, there are no observations on the diagonals $t = i + j$, hence we pose $\widehat{\zeta}_{i+j} = \mathbf{E}[\zeta_{i+j}] = 0$ for $i + j > I$. Therefore, we consider the predictor

$$\widehat{X}_{ij} = a_i\,\gamma_j\,\widehat{\eta}_i\,. \tag{8}$$

Using (8), we have the following predictor for the reserve in accident year $i \in \{I - J + 1, \ldots, I\}$

$$\widehat{R}_i = \sum_{j=I-i+1}^{J} \widehat{X}_{ij} = a_i\,(1 - \beta_{I-i})\,\widehat{\eta}_i,$$

where $\beta_j := \sum_{k=0}^{j} \gamma_k,\;\; j = 0, \ldots, J$. The corresponding predictor of total reserve is then obtained by summing over the open accident years

$$\widehat{R} = \sum_{i=I-J+1}^{I} \widehat{R}_i\,.$$

We want to derive estimators for the reserve of single accident years and the total reserve, as well as estimates of the corresponding prediction error. The estimators we are looking for are in the class of the credibility estimators.

4. Credibility estimators

4.1. Incremental loss ratios

For the predictor $\widehat{\eta}_i$, we define the *incremental loss ratios* $(0 \le i \le I,\; 0 \le j \le J)$

$$Z_{ij} := \frac{X_{ij}}{a_i\,\gamma_j}\,.$$

Then, using (6), we have

$$Z_{ij} = \eta_i + \zeta_{i+j} + \frac{\sigma}{\sqrt{a_i\,\gamma_j}}\,\varepsilon_{ij}, \tag{9}$$

which, defining $\boldsymbol{\eta} = \{\eta_i;\;\; 0 \le i \le I\}$, implies

$$\mathbf{E}(Z_{ij}|\boldsymbol{\eta}) = \eta_i, \qquad \mathbf{Var}(Z_{ij}|\boldsymbol{\eta}) = \chi^2 + \frac{\sigma^2}{a_i\,\gamma_j}\,.$$

By (7) we have

$$\mathbf{Cov}(Z_{ij}, Z_{kl}) = \tau^2\, \mathbb{I}_{i,k} + \chi^2\, \mathbb{I}_{i+j,k+l} + \frac{\sigma^2}{a_i\, \gamma_j}\, \mathbb{I}_{i,k}\, \mathbb{I}_{j,l}, \quad 0 \le i, k \le I,\ 0 \le j, l \le J.$$

4.2. Covariance matrix of the observations

We shall denote by $\mathbf{\Omega}$ the covariance matrix of all $Z_{i,j}$ observed at time I. Let us introduce the double index (i, j), where i denotes an accident year and j denotes a development year. We consider the set of double indices (i, j) defined as

$$\mathcal{I} := \left\{(i, j) :\ X_{i,j} \in \mathcal{D}_I\right\} .$$

The set $\mathcal{I}$ has $N = J\,(I - J) + J\,(J + 1)/2$ elements. For any pair of observations $X_{i,j},\ X_{k,l} \in \mathcal{D}_I$, we consider the covariance

$$\omega_{(i,j),(k,l)} := \mathbf{Cov}(Z_{i,j}, Z_{k,l}),$$

then the covariance matrix of the observations is

$$\mathbf{\Omega} := \left(\omega_{(i,j),(k,l)}\right)_{(i,j),(k,l)\in\mathcal{I}} .$$

We also introduce a single index notation by defining an ordering in the observation set $\mathcal{I}$ and denoting by $\nu = 1, \ldots, N$ the νth pair in the set. Hence, any observed $Z_{i,j}$ can be also denoted as Z_ν. To fix ideas, we choose the left-to-right/top-to-bottom (LR/TB) ordering

$$\nu(i, j) := 1 + j + \sum_{k=0}^{i-1} [\iota(k) + 1], \quad 0 \le i \le I,\ 0 \le j \le \iota(i),$$

where $\iota(i) := \min(I - i, J)$. Then the covariance matrix of the observations can also be denoted as

$$\mathbf{\Omega} := \left(\omega_{\nu,\lambda}\right)_{\nu,\lambda=1,\ldots,N},$$

where

$$\omega_{\nu,\lambda} := \mathbf{Cov}(Z_\nu, Z_\lambda), \quad 1 \le \nu, \lambda \le N.$$

The single index notation allows us to use matrix notation. In order to switch from the single index to the double index notation, we also define, for $\nu = 1, \ldots, N$

$$\mathrm{ay}(\nu) = \text{accident year of } Z_\nu \quad \text{and} \quad \mathrm{dy}(\nu) = \text{development year of } Z_\nu .$$

Remark. The ordering chosen for the single index notation is irrelevant for the results below. However, the LR/TB ordering allows for a more intuitive representation (see Table 1). Observe that for $\chi^2 = 0$, the $\mathbf{\Omega}$ matrix is block diagonal.

TABLE 1

STRUCTURE OF THE Ω MATRIX WITH $I = J = 2$ (WITH $\sigma_\Theta^2 := \tau^2 + \chi^2$).

		1	2	3	4	5	6
λ	ν	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(2,0)
1	(0,0)	$\sigma_\Theta^2 + \sigma^2/w_{0,0}$	τ^2	τ^2	0	0	0
2	(0,1)	τ^2	$\sigma_\Theta^2 + \sigma^2/w_{0,1}$	τ^2	χ^2	0	0
3	(0,2)	τ^2	τ^2	$\sigma_\Theta^2 + \sigma^2/w_{0,2}$	0	χ^2	χ^2
4	(1,0)	0	χ^2	0	$\sigma_\Theta^2 + \sigma^2/w_{1,0}$	τ^2	0
5	(1,1)	0	0	χ^2	τ^2	$\sigma_\Theta^2 + \sigma^2/w_{1,1}$	χ^2
6	(2,0)	0	0	χ^2	0	χ^2	$\sigma_\Theta^2 + \sigma^2/w_{2,0}$

4.3. Best linear inhomogeneous estimators

Let $\boldsymbol{Z} := (Z_1, \ldots, Z_N)'$. The best linear inhomogeneous (BL) estimator for $\eta_i,\ 0 \le i \le I$, is defined as

$$\widehat{\eta}_i = \operatorname{argmin}_{\widehat{\eta}_i \in L(\boldsymbol{Z},1)} \mathbf{E}[(\widehat{\eta}_i - \eta_i)^2], \tag{10}$$

where $L(\boldsymbol{Z}, 1)$ is the space of all the linear combination of $Z_1, \ldots, Z_N, 1$, that is

$$L(\boldsymbol{Z}, 1) := \left\{ \widehat{\eta}_i : \widehat{\eta}_i = c_0^{(i)} + \sum_{\nu=1}^{N} c_\nu^{(i)} Z_\nu, \quad c_0^{(i)}, c_1^{(i)}, \ldots, c_N^{(i)} \in \mathbb{R} \right\}.$$

This estimator is "global" in the sense that it is based on all the observations at time I. By Corollary 3.17 in Bühlmann and Gisler (2005), $\widehat{\eta}_i$ satisfies (10) if and only if the following *normal equations* are satisfied for $0 \le i \le I$

$$\begin{aligned} &i. \quad \mathbf{E}(\widehat{\eta}_i - \eta_i) = 0, \\ &ii. \quad \mathbf{Cov}(\widehat{\eta}_i, Z_\nu) = \mathbf{Cov}(\eta_i, Z_\nu), \quad \nu = 1, \ldots, N. \end{aligned} \tag{11}$$

Note that condition (i) implies that $\widehat{\eta}_i$ is unbiased. As shown in Corollary 3.18 in Bühlmann and Gisler (2005), if the covariance matrix is non-singular, then it follows that

$$\widehat{\eta}_i = \mu_0 + \boldsymbol{\kappa}_i' \, \boldsymbol{\Omega}^{-1} (\boldsymbol{Z} - \mu_0 \, \mathbf{1}),$$

with $\mathbf{1} := (1, \ldots, 1)' \in \mathbb{R}^N$ and where

$$\boldsymbol{\kappa}_i' := (\mathbf{Cov}(\eta_i, Z_1), \ldots, \mathbf{Cov}(\eta_i, Z_N)) \;.$$

The covariances $\boldsymbol{\kappa}$ have a specific expression in the ADR model. Under our assumptions, we have

$$\boldsymbol{\kappa}_i' = \tau^2 \, \boldsymbol{\delta}_i' \quad \text{where} \quad \boldsymbol{\delta}_i := \left(\mathbb{I}_{\mathrm{ay}(1),i}, \ldots, \mathbb{I}_{\mathrm{ay}(N),i}\right)' \;;$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

hence, the BL estimator for η_i is given by

$$\widehat{\eta}_i = \mu_0 + \tau^2\, \boldsymbol{\delta}_i'\, \boldsymbol{\Omega}^{-1}(\boldsymbol{Z} - \mu_0\, \mathbf{1}), \qquad 0 \le i \le I. \tag{12}$$

Denoting by $\omega_{\nu,\lambda}^{(-1)}$, or $\omega_{(i,j),(k,l)}^{(-1)}$, the generic element of $\boldsymbol{\Omega}^{-1}$, expression (12) can also be written in double index notation

$$\widehat{\eta}_i = \mu_0 + \tau^2 \sum_{j=0}^{\iota(i)} \sum_{k=0}^{I} \sum_{l=0}^{\iota(k)} \omega_{(i,j),(k,l)}^{(-1)}\, (Z_{k,l} - \mu_0), \qquad 0 \le i \le I.$$

Observe that the $c_\nu^{(i)}$ coefficients can also be represented in double index notation as $c_{k,l}^{(i)}$, with $k = \mathrm{ay}(\nu)$ and $l = \mathrm{dy}(\nu)$.

By these results, the following fundamental theorem is immediately obtained.

Theorem 4.1 (Best linear inhomogeneous estimator). *The best linear inhomogeneous estimator for η_i, $0 \le i \le I$, is given by*

$$\widehat{\eta}_i = \mu_0\, (1 - \alpha_i) + \alpha_i\, \bar{\bar{Z}}_i, \tag{13}$$

where

$$\alpha_i := \sum_{k=0}^{I} \sum_{l=0}^{\iota(k)} c_{k,l}^{(i)} = \tau^2 \sum_{k=0}^{I} \sum_{l=0}^{\iota(k)} \sum_{j=0}^{\iota(i)} \omega_{(i,j),(k,l)}^{(-1)}, \tag{14}$$

and

$$\bar{\bar{Z}}_i := \sum_{k=0}^{I} \sum_{l=0}^{\iota(l)} \frac{c_{k,l}(i)}{\alpha_i}\, Z_{k,l} = \frac{\tau^2}{\alpha_i} \sum_{k=0}^{I} \sum_{l=0}^{\iota(k)} \sum_{j=0}^{\iota(i)} \omega_{(i,j),(k,l)}^{(-1)}\, Z_{k,l}. \tag{15}$$

Proof. Let $c_0^{(i)}, c_1^{(i)}, \ldots, c_N^{(i)}$ be the coefficients satisfying (10); that is

$$\widehat{\eta}_i = c_0^{(i)} + \sum_{\nu=1}^{N} c_\nu^{(i)}\, Z_\nu, \qquad 0 \le i \le I. \tag{16}$$

By condition (i) and (ii) of normal Equations (11), one finds (see proof of Corollary 3.18 in Bühlmann and Gisler (2005)) that $c_0^{(i)}$ can also be expressed as

$$c_0^{(i)} = \mu_0 - \boldsymbol{\kappa}_i'\, \boldsymbol{\Omega}^{-1} \mu_0\, \mathbf{1} = \mu_0 (1 - \tau^2 \boldsymbol{\delta}_i'\, \boldsymbol{\Omega}^{-1} \mathbf{1})\,;$$

therefore we also have, for $0 \le i \le I$

$$\left(c_1^{(i)}, \ldots, c_N^{(i)}\right) = \tau^2 \boldsymbol{\delta}_i'\, \boldsymbol{\Omega}^{-1} \quad \text{and} \quad c_0^{(i)} = \mu_0 \left(1 - \sum_{\nu=1}^{N} c_\nu^{(i)}\right).$$

We then obtain

$$\widehat{\eta}_i = \mu_0 \left(1 - \sum_{\nu=1}^{N} c_\nu^{(i)}\right) + \sum_{\nu=0}^{N} c_\nu^{(i)} Z_\nu, \qquad 0 \leq i \leq I,$$

which gives (13) posing $\alpha_i = \sum_{\nu=1}^{N} c_\nu^{(i)}$. The $c_\nu^{(i)}$ coefficient has the explicit form

$$c_\nu^{(i)} = \tau^2 \sum_{j=0}^{\iota(i)} \omega_{(i,j),(\mathrm{ay}(\nu),\mathrm{dy}(\nu))}^{(-1)}, \qquad 0 \leq i \leq I, \quad 1 \leq \nu \leq N, \tag{17}$$

or, using double index notation,

$$c_{k,l}^{(i)} = \tau^2 \sum_{j=0}^{\iota(i)} \omega_{(i,j),(k,l)}^{(-1)}, \qquad 0 \leq i, k \leq I, \quad 0 \leq l \leq \iota(k).$$

■

By Theorem 4.1, the BL estimator for η_i is the weighted average, with credibility weights α_i, of μ_0 and the average incremental loss ratio $\bar{\bar{Z}}_i$. Observe that in the ADR model the average $\bar{\bar{Z}}_i$ is taken over the entire trapezoid $\mathcal{D}_I$.

Remarks

- By expression (17), the $c_\nu^{(i)}$ coefficient is obtained by the matrix $\mathbf{\Omega}^{-1}$, under the LR/TB ordering, by computing the νth column sum on the band corresponding to accident year i. The coefficient α_i is obtained by summing all these column sums.
- For $\chi^2 = 0$, one has $c_\nu^{(i)} = 0$ if $\mathrm{ay}(\nu) \neq i$, since $\mathbf{\Omega}^{-1}$ is block diagonal. Therefore α_i, $\bar{\bar{Z}}_i$ and $\widehat{\eta}_i$ are based only on the observations of accident year i. It can be shown that the version of Theorem 4.1 obtained in the case $\chi^2 = 0$ provides all the main results of the classical credibility theory. This is illustrated in Appendix A.

The following lemma provides the covariance structure of the BL estimators.

Lemma 4.2. *The covariance of the BL estimators is given by*

$$\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = \tau^2 \sum_{p=0}^{I} \sum_{l=0}^{\iota(p)} c_{p,l}^{(i)} \, \mathbb{I}_{p,k} = \left(\tau^2\right)^2 \sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(k)} \omega_{(i,j),(k,l)}^{(-1)}, \qquad 0 \leq i, k \leq I. \tag{18}$$

Proof. From (16)

$$\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = \mathbf{Cov}\left(\widehat{\eta}_i,\ \left(c_0^{(k)} + \sum_{\nu=1}^{N} c_\nu^{(k)}\, Z_\nu\right)\right)$$

$$= \sum_{\nu=1}^{N} c_\nu^{(k)}\, \mathbf{Cov}\,(\widehat{\eta}_i,\, Z_\nu)$$

$$= \sum_{\nu=1}^{N} c_\nu^{(k)}\, \mathbf{Cov}\,(\eta_i,\, Z_\nu)\,,$$

where the last equality holds by condition (*ii*) of the normal equations. Then, we have

$$\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = \tau^2 \sum_{\nu=1}^{N} c_\nu^{(k)}\, \mathbb{I}_{\mathrm{ay}(\nu),i} = \tau^2 \sum_{\nu=1}^{N} c_\nu^{(i)}\, \mathbb{I}_{\mathrm{ay}(\nu),k}, \qquad 0 \le i, k \le I\,.$$

■

The variance of the BL estimators is given by

$$\mathbf{Var}(\widehat{\eta}_i) = \left(\tau^2\right)^2 \sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)}, \qquad 0 \le i \le I\,. \tag{19}$$

Since by Theorem 4.1 $\widehat{\eta}_i = \alpha_i\,(\bar{\bar{\bar{Z}}}_i - \mu_0) + \mu_0$, the covariance of the average incremental loss ratios is given by

$$\mathbf{Cov}(\bar{\bar{\bar{Z}}}_i,\, \bar{\bar{\bar{Z}}}_k) = \frac{1}{\alpha_i\,\alpha_k}\, \mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k), \qquad 0 \le i, k \le I\,.$$

For $i = k$

$$\mathbf{Var}(\bar{\bar{\bar{Z}}}_i) = \frac{1}{\alpha_i^2}\, \mathbf{Var}(\widehat{\eta}_i), \qquad 0 \le i \le I\,.$$

Remark . For $\chi^2 = 0$, since $\mathbf{\Omega}^{-1}$ is block diagonal, $\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = 0$ for $i \neq k$. In this case

$$\alpha_i = \tau^2 \sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)},$$

hence, $\mathbf{Var}(\widehat{\eta}_i) = \tau^2\, \alpha_i$ and $\mathbf{Var}(\bar{\bar{\bar{Z}}}_i) = \tau^2/\alpha_i$.

In the following lemma, we provide cross-products of losses of BL estimators.

Lemma 4.3. *For $0 \leq i, k \leq I$, the cross-product of losses of the BL estimators is given by*

$$\boldsymbol{E}[(\widehat{\eta}_i - \eta_i)(\widehat{\eta}_k - \eta_k)] = \begin{cases} \tau^2 - \boldsymbol{Var}(\widehat{\eta}_i), & i = k, \\ -\,\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k), & i \neq k. \end{cases} \tag{20}$$

Proof. For $0 \leq i, k \leq I$ one has

$$\begin{aligned} \mathbf{E}[(\widehat{\eta}_i - \eta_i)(\widehat{\eta}_k - \eta_k)] &= \mathbf{E}[(\widehat{\eta}_i - \mu_0)(\widehat{\eta}_k - \mu_0)] \\ &\quad - \mathbf{E}[(\widehat{\eta}_i - \mu_0)(\eta_k - \mu_0)] - \mathbf{E}[(\widehat{\eta}_k - \mu_0)(\eta_i - \mu_0)] \\ &\quad + \mathbf{E}[(\eta_i - \mu_0)(\eta_k - \mu_0)] \\ &= \mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) + \mathbf{Cov}(\eta_i, \eta_k) \\ &\quad - \mathbf{E}[(\widehat{\eta}_i - \mu_0)(\eta_k - \mu_0)] - \mathbf{E}[(\widehat{\eta}_k - \mu_0)(\eta_i - \mu_0)] . \end{aligned}$$

We have

$$\begin{aligned} \mathbf{E}[(\widehat{\eta}_i - \mu_0)(\eta_k - \mu_0)] &= \mathbf{E}[(\widehat{\eta}_i\, \eta_k)] - \mu_0^2 \\ &= \mu_0\, c_0^{(i)} + \sum_{\nu=1}^{N} c_\nu^{(i)}\, \mathbf{E}(Z_\nu\, \eta_k) - \mu_0^2 \\ &= \mu_0\, c_0^{(i)} + \sum_{\nu=1}^{N} c_\nu^{(i)}\, (\mu_0^2 + \tau^2\, \mathbb{I}_{\mathrm{ay}(\nu),k}) - \mu_0^2 \\ &= \mu_0^2 \left(1 - \sum_{\nu=1}^{N} c_\nu^{(i)}\right) + \mu_0^2 \sum_{\nu=1}^{N} c_\nu^{(i)} \\ &\quad + \tau^2 \sum_{\nu=1}^{N} c_\nu^{(i)}\, \mathbb{I}_{\mathrm{ay}(\nu),k} - \mu_0^2 \\ &= \tau^2 \sum_{\nu=1}^{N} c_\nu^{(i)}\, \mathbb{I}_{\mathrm{ay}(\nu),k} = \mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) . \end{aligned}$$

Hence,

$$\mathbf{E}[(\widehat{\eta}_i - \eta_i)(\widehat{\eta}_k - \eta_k)] = \mathbf{Cov}(\eta_i, \eta_k) - \mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) .$$

Therefore, for $i = k$ we obtain

$$\mathbf{E}\left[(\widehat{\eta}_i - \eta_i)^2\right] = \tau^2 - \mathbf{Var}(\widehat{\eta}_i) = \tau^2 \left(1 - \tau^2 \sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(k)} \omega_{(i,j),(i,l)}^{(-1)}\right), \tag{21}$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

and for $i \neq k$

$$\mathbf{E}[(\widehat{\eta}_i - \eta_i)(\widehat{\eta}_k - \eta_k)] = -\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = -\left(\tau^2\right)^2 \sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(k)} \omega^{(-1)}_{(i,j),(k,l)} \,. \tag{22}$$

■

4.4. Best linear homogeneous estimators

Later, we shall also use estimates of η_i with μ_0 estimated from the data. This leads to the definition of homogeneous estimators.

We have defined the inhomogeneous estimator for η_i as the best estimator in the space $L(\boldsymbol{Z}, 1)$ of all the linear combination of $Z_1, \ldots, Z_N$, 1. Let us now define the linear space

$$L_e(\boldsymbol{Z}) := \left\{\text{all linear combination of } Z_1, \ldots, Z_N, \text{ with expectation } \mu_0\right\}.$$

The best linear homogeneous estimator (BLH) for η_i is defined as

$$\widehat{\eta}_i^{\text{hom}} = \text{argmin}_{\widehat{\eta}_i \in L_e(\boldsymbol{Z})} \mathbf{E}[(\widehat{\eta}_i - \eta_i)^2], \qquad 0 \leq i \leq I. \tag{23}$$

The homogeneous estimator is obtained starting from the same linear space where the inhomogeneous estimator is defined and imposing the additional requirements of homogeneity and unbiasedness. Forcing the constant term in the inhomogeneous estimator to be zero, together with the unbiasedness condition, leads to a built-in estimator for μ_0. In addition, we have a useful property (Theorem of Pythagoras) for the quadratic loss of the homogeneous estimator. This is shown by the following fundamental theorem.

Theorem 4.4 (Best linear homogeneous estimator). *The BLH estimator for η_i, $0 \leq i \leq I$, is given by*

$$\widehat{\eta}_i^{\text{hom}} = \alpha_i \,\bar{\bar{Z}}_i + (1 - \alpha_i)\,\widehat{\mu}_0 \,. \tag{24}$$

where

$$\widehat{\mu}_0 = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha_\bullet} \,\bar{\bar{Z}}_i \,. \tag{25}$$

Moreover, the quadratic loss of $\widehat{\eta}_i^{\text{hom}}$ is given by

$$\boldsymbol{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \eta_i\right)^2\right] = \boldsymbol{E}\left[(\widehat{\eta}_i - \eta_i)^2\right] + \boldsymbol{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \widehat{\eta}_i\right)^2\right]. \tag{26}$$

Proof. We use the approach based on Hilbert spaces introduced in Bühlmann and Gisler (2005), Section 3.2.1. We consider the Hilbert space $\mathcal{L}^2$ of the real-valued square integrable random variables and for any closed affine

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

space $M \subset \mathcal{L}^2$, we denote by $\mathrm{Pro}(X|M)$ the orthogonal projection of the random variable $X \in \mathcal{L}^2$ on M. Moreover, we denote by $\|X\| := E(X^2)$ the norm of $X \in \mathcal{L}^2$. In terms of Hilbert spaces, condition (10), and then expression (13), is equivalent to

$$\widehat{\eta_i} = \mathrm{Pro}\big(\eta_i | L(\boldsymbol{Z}, 1)\big),$$

and condition (23) is equivalent to

$$\widehat{\eta}_i^{\mathrm{hom}} = \mathrm{Pro}\big(\eta_i \big| L_e(\boldsymbol{Z})\big)\,.$$

Since $L_e(\boldsymbol{Z}) \subset L(\boldsymbol{Z}, 1)$, by the iterativity of projections (see Theorem 3.13 in Bühlmann and Gisler (2005)), we have

$$\mathrm{Pro}\big(\eta_i \big| L_e(\boldsymbol{Z})\big) = \mathrm{Pro}\Big(\mathrm{Pro}\big(\eta_i \big| L(\boldsymbol{Z}, 1)\big)\Big| L_e(\boldsymbol{Z})\Big), \tag{27}$$

and

$$\begin{aligned}\big\|\eta_i - \mathrm{Pro}\big(\eta_i | L_e(\boldsymbol{Z})\big)\big\|^2 &= \big\|\eta_i - \mathrm{Pro}\big(\eta_i | L(\boldsymbol{Z}, 1)\big)\big\|^2 \\ &\quad + \big\|\mathrm{Pro}\big(\eta_i | L_e(\boldsymbol{Z})\big) - \mathrm{Pro}\big(\eta_i | L(\boldsymbol{Z}, 1)\big)\big\|^2 .\end{aligned} \tag{28}$$

Equation (28), which corresponds to the Theorem of Pythagoras, is equivalent to (26).

Using the linearity of $\widehat{\eta_i}$, (27) gives

$$\mathrm{Pro}\big(\eta_i \big| L_e(\boldsymbol{Z})\big) = \alpha_i \,\bar{\bar{Z}}_i + (1 - \alpha_i)\, \mathrm{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big)\,. \tag{29}$$

We then need an expression for $\mathrm{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big)$. Let us define the linear space

$$\begin{aligned}&L_e(\boldsymbol{Z}, \boldsymbol{\eta}) \\ &\quad := \big\{\text{all linear combination of } Z_1, \ldots, Z_N, \eta_0, \ldots, \eta_I \text{ with expectation } \mu_0\big\}\,.\end{aligned}$$

Using Theorem 3.16 in Bühlmann and Gisler (2005), we have

$$\mathrm{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z}, \boldsymbol{\eta})\big) = \frac{1}{I+1} \sum_{k=i}^{I} \eta_i := \bar{\eta}\,.$$

This can be seen as follows. The first orthogonality condition $\mathbf{E}(\bar{\eta}, \mu_0) = 0$ in the theorem is trivially fulfilled. Then, we have to prove that $\bar{\eta}$ also satisfies the second orthogonality condition

$$\mathbf{E}\left[(\bar{\eta} - \mu_0)(\bar{\eta} - \eta^*)\right] = 0 \quad \text{for all} \quad \eta^* \in L_e(\boldsymbol{Z}, \boldsymbol{\eta})\,.$$

For $\eta^* = \eta_i, \; 0 \le i \le I$, one has, by the independence of η_i,

$$\mathbf{E}\,[(\bar\eta - \mu_0)(\bar\eta - \eta_i)] = \mathbf{E}\,[(\bar\eta - \mu_0)\bar\eta] - \sum_{k=0}^{I} \frac{\mathbf{E}[(\eta_k - \mu_0)\,\eta_i]}{I+1}$$

$$= \mathbf{Var}(\bar\eta) - \frac{\mathbf{Var}(\eta_i)}{I+1}$$

$$= \frac{\tau^2}{I+1} - \frac{\tau^2}{I+1} = 0\,.$$

The same result holds for $\eta^* = Z_\nu, \; 1 \le \nu \le N$, since $\sum_{k=0}^{I} \mathbf{E}[(\eta_k - \mu_0)\, Z_\nu] = \mathbf{Var}(\eta_{\text{ay}(\nu)}) = \tau^2$.

Then, since $L_e(\boldsymbol{Z}) \subset L(\boldsymbol{Z}, \eta)$, again by the iterativity we obtain

$$\text{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big) = \text{Pro}\Big(\text{Pro}\big(\mu_0 \big| L(\boldsymbol{Z}, \boldsymbol{\eta})\big)\Big| L_e(\boldsymbol{Z})\Big)$$

$$= \text{Pro}\big(\bar\eta \big| L_e(\boldsymbol{Z})\big)$$

$$= \frac{1}{I+1} \sum_{i=0}^{I} \text{Pro}\Big(\eta_i \Big| L_e(\boldsymbol{Z})\Big)\,.$$

Hence, by (29)

$$\text{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big) = \frac{1}{I+1} \sum_{i=0}^{I} \Big[\alpha_i\, \bar{\bar{Z}}_i + (1-\alpha_i)\, \text{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big)\Big],$$

which implies

$$\widehat{\mu}_0 := \text{Pro}\big(\mu_0 \big| L_e(\boldsymbol{Z})\big) = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha_\bullet}\, \bar{\bar{Z}}_i\,.$$

Then (29) reads

$$\widehat{\eta}_i^{\text{hom}} = \alpha_i\, \bar{\bar{Z}}_i + (1-\alpha_i)\, \widehat{\mu}_0\,.$$

■

The following corollaries provide an explicit expression for the variance of $\widehat{\mu}_0$ and the quadratic loss of the BLH estimators.

Corollary 4.5. *The variance of the estimator $\widehat{\mu}_0$ is given by*

$$\boldsymbol{E}\big[(\mu_0 - \widehat{\mu}_0)^2\big] = \frac{1}{(\alpha_\bullet)^2} \sum_{i=0}^{I} \sum_{k=0}^{I} \mathbf{Cov}\,(\widehat{\eta}_i, \widehat{\eta}_k)\,. \tag{30}$$

Proof. By (25)

$$\mathbf{E}\left[(\mu_0 - \widehat{\mu}_0)^2\right] = \mathbf{E}\left[\left(\sum_{k=0}^{I} \frac{\alpha_k}{\alpha_\bullet}\left(\bar{\bar{Z}}_k - \mu_0\right)\right)^2\right] = \sum_{k=0}^{I}\sum_{l=0}^{I} \frac{\alpha_k\,\alpha_l}{(\alpha_\bullet)^2}\,\mathbf{Cov}\left(\bar{\bar{Z}}_k, \bar{\bar{Z}}_l\right).$$

Then (30) follows, since $\alpha_k\alpha_l\,\mathbf{Cov}\left(\bar{\bar{Z}}_k, \bar{\bar{Z}}_l\right) = \mathbf{Cov}\left(\widehat{\eta}_k, \widehat{\eta}_l\right)$. ■

Corollary 4.6. *The quadratic loss of the BLH estimator is given by*

$$\boldsymbol{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \eta_i\right)^2\right] = \tau^2 - \mathbf{Var}(\widehat{\eta}_i) + \left(\frac{1-\alpha_i}{\alpha_\bullet}\right)^2 \sum_{k=0}^{I}\sum_{l=0}^{I} \mathbf{Cov}\left(\widehat{\eta}_k, \widehat{\eta}_l\right). \tag{31}$$

Proof. The term $\tau^2 - \mathbf{Var}(\widehat{\eta}_i)$ is the first term on the right-hand side of (26) given by (21). For the second term we have, by (24)

$$\mathbf{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \widehat{\eta}_i\right)^2\right] = (1-\alpha_i)^2\,\mathbf{E}\left[(\mu_0 - \widehat{\mu}_0)^2\right],$$

where the variance of $\widehat{\mu}_0$ is given by Corollary 4.5. ■

4.5. Mean square error of prediction of credibility estimators

The quadratic losses for the credibility estimators $\widehat{\eta}_i$ and $\widehat{\eta}_i^{\text{hom}}$ are of fundamental importance because they provide the unconditional MSEP for the corresponding estimators.

By the first expression in (20) and by (31), we immediately conclude that in the ADR model the MSEP for the inhomogeneous credibility estimator $\widehat{\eta}_i$, $0 \le i \le I$, is given by

$$\text{msep}_{\eta_i}(\widehat{\eta}_i) := \mathbf{E}\left[(\widehat{\eta}_i - \eta_i)^2\right] = \tau^2 - \mathbf{Var}(\widehat{\eta}_i) = \tau^2\left(1 - \tau^2 \sum_{j=0}^{\iota(i)}\sum_{l=0}^{\iota(k)} \omega^{(-1)}_{(i,j),(i,l)}\right), \tag{32}$$

and the MSEP for the homogeneous credibility estimator $\widehat{\eta}_i^{\text{hom}}$, $0 \le i \le I$, is given by

$$\begin{aligned}\text{msep}(\widehat{\eta}_i^{\text{hom}}) &:= \mathbf{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \eta_i\right)^2\right] = \text{msep}(\widehat{\eta}_i) + \mathbf{E}\left[\left(\widehat{\eta}_i^{\text{hom}} - \widehat{\eta}_i\right)^2\right] \\ &= \tau^2 - \mathbf{Var}(\widehat{\eta}_i) + \left(\frac{1-\alpha_i}{\alpha_\bullet}\right)^2 \sum_{k=0}^{I}\sum_{l=0}^{I} \mathbf{Cov}\left(\widehat{\eta}_k, \widehat{\eta}_l\right).\end{aligned} \tag{33}$$

The expressions for $\mathbf{Var}\left(\widehat{\eta}_i\right)$ and $\mathbf{Cov}\left(\widehat{\eta}_i, \widehat{\eta}_k\right)$ are given by (19) and (18).

Remark. If $\chi^2 = 0$, all the previous expressions reduce to the corresponding expression in the classical Bühlmann–Straub model.

5. Credibility estimate of reserves

5.1. Single accident year and total reserve estimate

Let us consider the inhomogeneous estimates. For accident year $i \in \{0, \dots, I\}$, we have

$$\widehat{\eta}_i = \alpha_i \, \bar{\bar{Z}}_i + (1 - \alpha_i) \, \mu_0,$$

then, returning to the X-scale we obtain, for $X_{ij} \in \mathcal{D}_I^c$

$$\widehat{X}_{ij} = \alpha_i \, a_i \, \bar{\bar{Z}}_i \, \gamma_j + (1 - \alpha_i) \, a_i \, \mu_0 \, \gamma_j \, .$$

Thus, the following predictor results for the reserve in AY $i \in \{I - J + 1, \dots, I\}$

$$\widehat{R}_i = \alpha_i \, a_i \, \bar{\bar{Z}}_i \, (1 - \beta_{I-i}) + (1 - \alpha_i) \, a_i \, \mu_0 \, (1 - \beta_{I-i}) \, . \tag{34}$$

The corresponding total reserve estimate is then obtained by summing over the open accident years

$$\widehat{R} = \sum_{i=I-J+1}^{I} \widehat{R}_i \, .$$

The corresponding expressions for the homogeneous case are obtained by replacing the estimator $\widehat{\mu}_0$ for μ_0. We have, for $i \in \{I - J + 1, \dots, I\}$

$$\widehat{R}_i^{\text{hom}} = \alpha_i \, a_i \, \bar{\bar{Z}}_i \, (1 - \beta_{I-i}) + (1 - \alpha_i) \, a_i \, \widehat{\mu}_0 \, (1 - \beta_{I-i}), \tag{35}$$

and

$$\widehat{R}^{\text{hom}} = \sum_{i=I-J+1}^{I} \widehat{R}_i^{\text{hom}} \, .$$

5.2. Credibility decomposition of reserve estimates

Formulae (34) and (35) show that our reserve estimates have two components which are then mixed with the credibility weights α_i and $1 - \alpha_i$. These weights do change for each accident year. For $i \in \{I - J + 1, \dots, I\}$, (34) and (35) can be written, respectively

$$\begin{aligned} \widehat{R}_i &= \alpha_i \, \widehat{R}_i^{\text{pro}} + (1 - \alpha_i) \, \widehat{R}_i^{\text{all}}, \\ \widehat{R}_i^{\text{hom}} &= \alpha_i \, \widehat{R}_i^{\text{pro}} + (1 - \alpha_i) \, \widehat{R}_i^{\text{all/hom}}, \end{aligned} \tag{36}$$

where the mixture components are defined as follows.

a. *Projective component*
The estimator $\widehat{R}_i^{\text{pro}}$ has the form

$$\widehat{R}_i^{\text{pro}} := a_i\, \bar{\bar{Z}}_i\, (1 - \beta_{I-i})\,.$$

It projects the final reserve from the observed data of *all accident years*.
In the special case of the BSCR model (ADR model with $\chi^2 = 0$), $\bar{\bar{Z}}_i$ reduces to

$$\bar{Z}_i = \frac{C_{i,\iota(i)}}{a_i\, \beta_{\iota(i)}}, \qquad i = 0, \ldots, I,$$

and we have

$$\widehat{R}_i^{\text{pro}} = \frac{1 - \beta_{I-i}}{\beta_{I-i}}\, C_{i,I-i}, \qquad i = I - J + 1, \ldots, I\,.$$

That is all information needed is contained in $C_{i,I-i}$, the (*latest*) *diagonal element* of accident year i in the triangle/trapezoid of observed cumulative data, which is then multiplied by the projection factor $(1 - \beta_{I-i})/\beta_{I-i}$.
In the general case of the ADR model, it is convenient to introduce a *generalized*, or *adjusted*, *diagonal element*

$$\bar{\bar{C}}_{i,\iota(i)} := a_i\, \beta_{\iota(i)}\, \bar{\bar{Z}}_i\,, \qquad i = 0, \ldots, I, \tag{37}$$

which permits to write analogously

$$\widehat{R}_i^{\text{pro}} := \frac{1 - \beta_{I-i}}{\beta_{I-i}}\, \bar{\bar{C}}_{i,I-i}, \qquad i = I - J + 1, \ldots, I\,.$$

Remarks .

- The quantity $\bar{\bar{Z}}_i$ depends on all observations and can be interpreted as a *correction factor* applied to the *a priori* value on the latest diagonal which is $a_i \beta_{I-i}$. In established reserving techniques the Cape Cod method calculates similarly

$$\bar{Z}^{CC} := \frac{\sum_{k=0}^{I} C_{k,\iota(k)}}{\sum_{k=0}^{I} a_k\, \beta_{\iota(k)}}, \tag{38}$$

  which is a correction factor *identical for all accident years*. All the correction factors $\bar{\bar{Z}}_i$, $\bar{Z}_i$, $\bar{Z}^{CC}$ are average incremental loss ratios, where the average is taken over different observation sets using different weights.
- If the development pattern $(\beta_j)_{j=0,\ldots,J}$ is calculated from the Chain Ladder factors $(f_j)_{j=0,\ldots,J-1}$ by $\beta_j = 1/(f_j\, f_{j+1} \cdots f_{J-1})$, we have

$$\frac{1 - \beta_j}{\beta_j} = f_j\, f_{j+1} \cdots f_{J-1} - 1, \tag{39}$$

  which is the Chain Ladder projection factor. We make this remark to emphasize that the Chain Ladder is the prototype of the projective reserve.

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

b. *Allocative Component*
The estimators

$$\widehat{R}_i^{\text{all}} := a_i\,\mu_0\,(1-\beta_{I-i}),$$

$$\widehat{R}_i^{\text{all/hom}} := a_i\,\widehat{\mu}_0\,(1-\beta_{I-i}),$$

allocate the *a priori* estimate $a_i\mu_0$ or $a_i\widehat{\mu}_0$, respectively, to the development years still to come. In the inhomogeneous case, this is the Bornhuetter–Ferguson reserve estimate with *a priori* values $a_i\mu_0$. In the homogeneous case, the *a priori* values are adjusted by experience to $a_i\widehat{\mu}_0$, with $\widehat{\mu}_0$ given by (25).

Remark. Here the procedure is also similar to that based on the Cape Cod method, where one sets $\widehat{\mu}_0^{\text{CC}} = \bar{Z}^{\text{CC}}$.

6. Mean square error of prediction of reserves

We have that for a single accident year $i = I-J+1,\ldots,I$ the open liability in the ADR model is given by

$$R_i = \sum_{j=I-i+1}^{J} X_{ij} = \sum_{j=I-i+1}^{J} \left[a_i\,\gamma_j\,(\eta_i+\zeta_{i+j}) + \sqrt{a_i\,\gamma_j}\,\sigma\,\varepsilon_{ij}\right],$$

and the inhomogeneous and homogeneous reserve estimators are

$$\widehat{R}_i = \sum_{j=I-i+1}^{J} \widehat{X}_{ij} = \sum_{j=I-i+1}^{J} a_i\,\gamma_j\,\widehat{\eta}_i = a_i\,(1-\beta_{\iota(i)})\,\widehat{\eta}_i,$$

$$\widehat{R}_i^{\text{hom}} = \sum_{j=I-i+1}^{J} \widehat{X}_{ij}^{\text{hom}} = \sum_{j=I-i+1}^{J} a_i\,\gamma_j\,\widehat{\eta}_i = a_i\,(1-\beta_{\iota(i)})\,\widehat{\eta}_i^{\text{hom}}.$$

For the aggregated accident years, we have the corresponding estimators

$$\widehat{R} = \sum_{i=I-J+1}^{I} \widehat{R}_i, \qquad \widehat{R}^{\text{hom}} = \sum_{i=I-J+1}^{I} \widehat{R}_i^{\text{hom}}.$$

For all these estimators, we derive expressions for the (unconditional) MSEP.

6.1. MSEP for the reserve estimate, single accident year

The MSEP for the inhomogeneous and homogeneous reserve estimators for accident year $i = I-J+1,\ldots,I$, is defined as, respectively

$$\text{msep}_{R_i}(\widehat{R}_i) := \mathbf{E}\left[(R_i-\widehat{R}_i)^2\right], \qquad \text{msep}_{R_i}(\widehat{R}_i^{\text{hom}}) := \mathbf{E}\left[\left(R_i-\widehat{R}_i^{\text{hom}}\right)^2\right].$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

The corresponding expressions are provided by the following theorem.

Theorem 6.1. *In the ADR model, the MSEP for the inhomogeneous and homogeneous reserve estimators for accident year $i = I - J + 1, \ldots, I$, are given by*

$$\mathrm{msep}_{R_i}(\widehat{R}_i) = a_i^2\,(1-\beta_{I-i})^2\,\mathrm{msep}_{\eta_i}(\widehat{\eta}_i) + \chi^2\,a_i^2 \sum_{j=I-i+1}^{J} \gamma_j^2 + a_i\,(1-\beta_{I-i})\,\sigma^2, \tag{40}$$

and

$$\begin{aligned}\mathrm{msep}_{R_i}(\widehat{R}_i^{\mathrm{hom}}) &= a_i^2\,(1-\beta_{I-i})^2\,\mathrm{msep}_{\eta_i}(\widehat{\eta}_i^{\mathrm{hom}}) \\ &\quad + \chi^2\,a_i^2 \sum_{j=I-i+1}^{J} \gamma_j^2 + a_i\,(1-\beta_{I-i})\,\sigma^2,\end{aligned} \tag{41}$$

respectively.

The expression for $\mathrm{msep}_{\eta_i}(\widehat{\eta}_i)$ and $\mathrm{msep}_{\eta_i}(\widehat{\eta}_i^{\mathrm{hom}})$ is given in (32) and (33) respectively.

Proof.
Inhomogeneous case. If we consider the difference

$$R_i - \widehat{R}_i = \underbrace{a_i\,(1-\beta_{I-i})\,(\eta_i - \widehat{\eta}_i)}_{A_i'} + \underbrace{\sum_{j=I-i+1}^{J} (a_i\gamma_j\,\zeta_{i+j} + \sqrt{a_i\gamma_j}\,\sigma\,\varepsilon_{i,j})}_{A_i''},$$

the terms A_i' and A_i'' are independent as $\widehat{\eta}_i$ depends only on past ζ and ε variables. Hence,

$$\begin{aligned}\mathbf{E}\left[(R_i - \widehat{R}_i)^2\right] &= \mathbf{E}\left(A_i'^2\right) + \mathbf{E}\left(A_i''^2\right) = a_i^2\,(1-\beta_{I-i})^2\,\mathbf{E}\left[(\eta_i - \widehat{\eta}_i)^2\right] \\ &\quad + \chi^2\,a_i^2 \sum_{j=I-i+1}^{J} \gamma_j^2 + a_i\,(1-\beta_{I-i})\,\sigma^2,\end{aligned}$$

which gives (40).

Homogeneous case. By the decomposition property of quadratic losses (26) in Theorem 4.4, we have

$$\mathbf{E}\left[\left(R_i - \widehat{R}_i^{\mathrm{hom}}\right)^2\right] = \mathbf{E}\left[\left(R_i - \widehat{R}_i\right)^2\right] + \mathbf{E}\left[\left(\widehat{R}_i - \widehat{R}_i^{\mathrm{hom}}\right)^2\right], \tag{42}$$

and, by Corollary 4.6

$$\mathrm{msep}_{\eta_i}\left(\widehat{\eta}_i^{\mathrm{hom}}\right) = \mathrm{msep}_{\eta_i}\left(\widehat{\eta}_i\right) + (1-\alpha_i)^2\,\mathbf{E}\left[(\mu_0 - \widehat{\mu}_0))^2\right].$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

Since

$$\widehat{R}_i - \widehat{R}_i^{\text{hom}} = a_i\,(1-\beta_{I-i})\,(1-\alpha_i)\,(\mu_0 - \widehat{\mu}_0),$$

we get

$$\mathbf{E}\Big[\big(R_i - \widehat{R}_i^{\text{hom}}\big)^2\Big] = a_i^2\,(1-\beta_{I-i})^2\,(1-\alpha_i)^2\,\mathbf{E}\left[(\mu_0 - \widehat{\mu}_0))^2\right],$$

which inserted in (42) gives (41). ■

Remark. For $\chi^2 = 0$, these are the BSCR formulae in Corollary 4.60 in Wüthrich and Merz, 2008.

6.2. MSEP for the total reserve

For the aggregated accident years, we consider the (unconditional) MSEP for the inhomogeneous and homogeneous total reserve estimator

$$\text{msep}_R(\widehat{R}) := \mathbf{E}\left[(R-\widehat{R})^2\right], \qquad \text{msep}_R(\widehat{R}^{\text{hom}}) := \mathbf{E}\left[(R-\widehat{R}^{\text{hom}})^2\right].$$

In this case, we have to take into account that the ζ_{i+j} variables generate dependence between accident years.

6.2.1. *MSEP for the total reserve, inhomogeneous case.*

Theorem 6.2. *In the ADR model, the MSEP for the inhomogeneous reserve estimator for the aggregated accident years is given by*

$$\begin{aligned}\text{msep}_R(\widehat{R}) = &\sum_{i=I-J+1}^{I} \text{msep}_{R_i}(\widehat{R}_i)\\ &-2\sum_{I-J+1\le i<k\le I} a_i\,a_k\,(1-\beta_{I-i})\,(1-\beta_{I-k})\,\mathbf{Cov}(\widehat{\eta}_i,\widehat{\eta}_k)\\ &+2\,\chi^2\sum_{I-J+1\le i<k\le I} a_i\,a_k \sum_{j=I-(i\wedge k)+1}^{J} \gamma_j\,\gamma_{j-|k-i|}\,.\end{aligned} \tag{43}$$

The expression for $\mathbf{Cov}(\widehat{\eta}_i,\widehat{\eta}_k)$ is given by Lemma 4.2.

Proof. We consider the difference

$$R-\widehat{R} = \underbrace{\sum_{i=I-J+1}^{I} a_i\,(1-\beta_{I-i})\,(\eta_i - \widehat{\eta}_i)}_{A'} + \underbrace{\sum_{i=I-J+1}^{I}\sum_{j=I-i+1}^{J} (a_i\gamma_j\,\zeta_{i+j} + \sqrt{a_i\gamma_j}\,\sigma\,\varepsilon_{i,j})}_{A''},$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

where the terms A' and A'' are still independent. Hence, $\mathbf{E}[(R-\widehat{R})^2] = \mathbf{E}\left(A'^2\right) + \mathbf{E}\left(A''^2\right)$. For the first expectation, we have

$$\mathbf{E}\left(A'^2\right) = \overbrace{\sum_{i=I-J+1}^{I} a_i^2\,(1-\beta_{I-i})^2\,\mathrm{msep}(\widehat{\eta}_i)}^{B} \\ + \underbrace{2\sum_{I-J+1\le i<k\le I} a_i\,a_k\,(1-\beta_{I-i})\,(1-\beta_{I-k})\,\mathbf{E}\left[(\widehat{\eta}_i-\eta_i)(\widehat{\eta}_k-\eta_k)\right]}_{C}\,.$$

For the second expectation, we obtain

$$\begin{aligned}\mathbf{E}\left(A''^2\right) &= \mathbf{E}\left[\left(\sum_{i=I-J+1}^{I}\sum_{j=I-i+1}^{J}(a_i\gamma_j\,\zeta_{i+j} + \sqrt{a_i\gamma_j}\,\sigma\,\varepsilon_{i,j})\right)^2\right] \\ &= \mathbf{E}\left[\sum_{i,k=I-J+1}^{I}\sum_{j=I-i+1}^{J} a_i\gamma_j\,\zeta_{i+j}\sum_{l=I-k+1}^{J} a_k\gamma_l\,\zeta_{k+l}\right] \\ &\quad + \mathbf{E}\left[\left(\sum_{i=I-J+1}^{I}\sum_{j=I-i+1}^{J}\sqrt{a_i\gamma_j}\,\sigma\,\varepsilon_{i,j}\right)^2\right] \\ &= \underbrace{\chi^2\sum_{i,k=I-J+1}^{I} a_i\,a_k\sum_{j=I-(i\wedge k)+1}^{J}\gamma_j\,\gamma_{j-|k-i|}}_{D} + \underbrace{\sum_{i=I-J+1}^{I} a_i(1-\beta_{I-i})\,\sigma^2}_{E},\end{aligned}$$

where the D term can be decomposed as

$$D = \underbrace{\chi^2\sum_{i=I-J+1}^{I} a_i^2\sum_{j=I-i+1}^{J}\gamma_j^2}_{D'} + \underbrace{2\,\chi^2\sum_{I-J+1\le i<k\le I} a_i\,a_k\sum_{j=I-(i\wedge k)+1}^{J}\gamma_j\,\gamma_{j-|k-i|}}_{D''}\,.$$

Then (43) is obtained as $\operatorname{msep}_R(\widehat{R}) = (B + D' + E) + C + D''$, observing that

$$B + D' + E = \sum_{i=I-J+1}^{I} \left(a_i^2 (1-\beta_{I-i})^2 \operatorname{msep}_{\eta_i}(\widehat{\eta}_i) + \chi^2 a_i^2 \sum_{j=I-i+1}^{J} \gamma_j^2 + a_i (1-\beta_{I-i}) \sigma^2 \right)$$
$$= \sum_{i=I-J+1}^{I} \operatorname{msep}_{R_i}(\widehat{R}_i),$$

and recalling that $\mathbf{E}[(\widehat{\eta}_i - \eta_i)(\widehat{\eta}_k - \eta_k)] = -\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k)$, by (22). ■

Remark. If $\chi^2 = 0$, one has $\operatorname{msep}_R(\widehat{R}) = \sum_i \operatorname{msep}_{R_i}(\widehat{R}_i)$ since $\mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) = 0$ for $i \neq k$.

6.2.2. *MSEP for the total reserve, homogeneous case.*

Theorem 6.3. *In the ADR model, the MSEP for the homogeneous reserve estimator for the aggregated accident years is given by*

$$\operatorname{msep}_R(\widehat{R}^{\text{hom}}) = \operatorname{msep}_R(\widehat{R}) + \left(\sum_{i=I-J+1}^{I} \frac{a_i (1-\alpha_i)}{\alpha_\bullet} (1-\beta_{I-i}) \right)^2 \sum_{i=0}^{I} \sum_{k=0}^{I} \mathbf{Cov}(\widehat{\eta}_i, \widehat{\eta}_k) \,. \tag{44}$$

Proof. For the total reserve, property (26) still holds, then

$$\operatorname{msep}_R(\widehat{R}^{\text{hom}}) = \operatorname{msep}_R(\widehat{R}) + \mathbf{E}\left[\left(\widehat{R} - \widehat{R}^{\text{hom}} \right)^2 \right].$$

Since

$$\widehat{R} - \widehat{R}^{\text{hom}} = \sum_{i=I-J+1}^{I} a_i (1-\beta_{I-i}) (1-\alpha_i) (\mu_0 - \widehat{\mu}_0),$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

we have, by Corollary 4.5,

$$\mathbf{E}\left[\left(R-\widehat{R}^{\mathrm{hom}}\right)^2\right]=\left(\sum_{i=I-J+1}^{I} a_i\,(1-\beta_{I-i})\,(1-\alpha_i)\right)^2\mathbf{E}\left[(\mu_0-\widehat{\mu}_0)^2\right]$$
$$=\left(\sum_{i=I-J+1}^{I}\frac{a_i\,(1-\alpha_i)}{\alpha_\bullet}\,(1-\beta_{I-i})\right)^2\sum_{i=0}^{I}\sum_{k=0}^{I}\mathbf{Cov}\,(\widehat{\eta}_i,\widehat{\eta}_k)$$

■

Corollary 6.4. *In the ADR model, the MSEP for the homogeneous reserve estimator for the aggregated accident years is given by*

$$\begin{aligned}\mathrm{msep}_R(\widehat{R}^{\mathrm{hom}})&=\sum_{i=I-J+1}^{I}\mathrm{msep}_{R_i}(\widehat{R}_i^{\mathrm{hom}})\\&-2\sum_{I-J+1\le i<k\le I}a_i\,a_k\,(1-\beta_{I-i})\,(1-\beta_{I-k})\,\mathbf{Cov}(\widehat{\eta}_i,\widehat{\eta}_k)\\&+2\,\chi^2\sum_{I-J+1\le i<k\le I}a_i\,a_k\sum_{j=I-(i\wedge k)+1}^{J}\gamma_j\,\gamma_{j-|k-i|}\,\cdot\\&+2\sum_{i=0}^{I}\sum_{k=0}^{I}\mathbf{Cov}\,(\widehat{\eta}_i,\widehat{\eta}_k)\\&\times\sum_{I-J+1\le i<k\le I}a_i\,a_k\,\frac{(1-\alpha_i)(1-\alpha_k)}{\alpha_\bullet^2}(1-\beta_{I-i})\,(1-\beta_{I-k})\,.\end{aligned}$$

Proof. This is immediately obtained by (44), using (43) and observing that, for $I-J+1\le i\le I$

$$\mathrm{msep}_{R_i}(\widehat{R}_i)-\mathrm{msep}_{R_i}(\widehat{R}_i^{\mathrm{hom}})=a_i^2\,(1-\beta_{I-i})^2\left(\mathrm{msep}_{\eta_i}(\widehat{\eta}_i)-\mathrm{msep}_{\eta_i}(\widehat{\eta}_i^{\mathrm{hom}})\right)$$
$$=-a_i^2(1-\beta_{I-i})^2\left(\frac{1-\alpha_i}{\alpha_\bullet}\right)^2\sum_{i=0}^{I}\sum_{k=0}^{I}\mathbf{Cov}\,(\widehat{\eta}_i,\widehat{\eta}_k)\,.$$

■

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

Remark . For $\chi^2 = 0$, one has $\mathbf{Cov}\,(\widehat{\eta}_i, \widehat{\eta}_k) = \mathbb{I}_{i,k}\,\tau^2\alpha_i$. Then

$$\begin{aligned}\mathbf{E}\left[\left(R - \widehat{R}^{\text{hom}}\right)^2\right] &= \left(\sum_{i=I-J+1}^{I} \frac{a_i\,(1-\alpha_i)}{\alpha_\bullet}\,(1-\beta_{I-i})\right)^2 \sum_{i=0}^{I}\sum_{k=0}^{I} \mathbf{Cov}\,(\widehat{\eta}_i, \widehat{\eta}_k) \\ &= \left(\sum_{i=I-J+1}^{I} \frac{a_i\,(1-\alpha_i)}{\alpha_\bullet}\,(1-\beta_{I-i})\right)^2 \sum_{i=0}^{I}\sum_{k=0}^{I} \mathbb{I}_{i,k}\,\tau^2\alpha_i \\ &= \frac{\tau^2}{\alpha_\bullet}\left(\sum_{i=I-J+1}^{I} a_i\,(1-\alpha_i)\,(1-\beta_{I-i})\right)^2 .\end{aligned}$$

In the BSCR model, Corollary 6.4 gives

$$\begin{aligned}\text{msep}_R(\widehat{R}^{\text{hom}}) = &\sum_{i=I-J+1}^{I} \text{msep}_{R_i}(\widehat{R}_i^{\text{hom}}) \\ &+ 2\,\frac{\tau^2}{\alpha_\bullet} \sum_{I-J+1\le i<k\le I} a_i\,a_k\,(1-\alpha_i)(1-\alpha_k)(1-\beta_{I-i})\,(1-\beta_{I-k}).\end{aligned}$$

7. Parameter estimation

Recall that in this paper, the prior values a_i and the development quotas γ_j are assumed to be known. In addition, we have the parameters:

- collective correction factor: μ_0,
- variance components: σ^2, τ^2, χ^2,

which we now want to estimate. The estimates for these parameters then need to be inserted into the formulae obtained in the earlier sections.

7.1. Assessing the value of μ_0

The role of the collective correction factor μ_0 is that of adjusting the prior values a_i to the "correct" level. We distinguish two cases.

a. *Inhomogeneous case*

   We believe that the *a priori* values a_i are already on the "correct" level. Hence, we use $\mu_0 = 1$. The resulting estimators for the reserves and the corresponding MSEP are the inhomogeneous estimators $\widehat{R}_i$, $\widehat{R}$ and $\text{msep}_{R_i}(\widehat{R}_i)$, $\text{msep}_R(\widehat{R})$.

b. *Homogeneous case*

   We adjust the *a priori* values a_i by a common correction factor $\widehat{\mu}_0$ which is estimated from the data using the estimator in Theorem 4.4. The corresponding estimators for the reserves and the MSEP are the homogeneous estimators $\widehat{R}_i^{\text{hom}}$, $\widehat{R}^{\text{hom}}$ and $\text{msep}_{R_i}(\widehat{R}_i^{\text{hom}})$, $\text{msep}_R(\widehat{R}^{\text{hom}})$.

7.2. Estimation of variance components

For the estimation of the variance parameters τ^2, χ^2, σ^2, we adopt a method which is typical in the analysis of variance. Some types of *sum of square errors* (SS) are taken from the data and, given the model assumptions, the expectation $\mathbf{E}(\text{SS})$ of each of these sums is expressed as a function $f(\tau^2, \chi^2, \sigma^2)$ of the variance parameters. If the expectation $\mathbf{E}(\text{SS})$ is replaced by the corresponding observed value SS^*, the equations $\text{SS}^* = f(\tau^2, \chi^2, \sigma^2)$ can provide sufficient constraints to identify (i.e. estimate) the variance parameters. We apply this method considering three different SS taken on the incremental loss ratios in the observed trapezoid $\mathcal{D}_I$. Given the assumptions in our model, the functions f are linear, and replacing the expectations with the corresponding observed values, we obtain a system of three independent linear equations. We then obtain estimates for τ^2, χ^2, σ^2 by solving this system. Because of the linearity of the equations, these estimates are *unbiased.*

In applying this estimation method, we consider two alternative approaches. In a first approach, the SS are taken directly on the trapezoid $\mathcal{D}_I$, where the paid losses are organized by accident year and development year. An alternative approach is obtained by rearranging the paid losses by accident year and calendar year and taking the SS on this AY/CY data. Since the two approaches lead to two different systems of independent linear equations, two distinct sets of estimators for the variance parameters are available; these estimators are both unbiased but are possibly characterized by different efficiency properties. In the following sections, we provide the estimation equation systems under the AY/DY and AY/CY approaches. The details of the derivations are given in Appendix B. The efficiency of the two sets of estimators is explored numerically in Section 8.2, where we observe that the AY/CY approach performs better than the AY/DY approach.

7.2.1. *Variance parameters estimates with the AY/DY approach.* The SS computation involves weighted sums taken over the data, thus including only cells $(i, j) \in \mathcal{D}_I$. To simplify notation, we define the new weights

$$q_{i,j} = \begin{cases} w_{i,j} = a_i \gamma_j & \text{if} \quad (i, j) \in \mathcal{D}_I, \\ 0 & \text{if} \quad (i, j) \in \mathcal{D}_I^c. \end{cases} \tag{45}$$

Hence, all the weighted sums over i can be extended from $i = 0$ to $i = I$ and all the weighted sums over j can be extended from $j = 0$ to $j = J$. We shall denote by $q_{\bullet j}$, $q_{i \bullet}$, $q_{\bullet\bullet}$ the sums $\sum_{i=0}^{I} q_{i,j}$, $\sum_{j=0}^{J} q_{i,j}$, $\sum_{i=0}^{I} \sum_{j=0}^{J} q_{i,j}$ respectively. Moreover, we denote by J_i, $i \in \{0, \ldots, I\}$, the number of nonzero $q_{i,j}$-weights of accident year i, and we denote $J_\bullet = \sum_{i=0}^{I} J_i$. In this Section 7 treating the parameter estimation, we will also use the following notation

$$\bar{Z}_i^* := \sum_{j=0}^{J} \frac{q_{i,j}}{q_{i\bullet}} Z_{ij}, \quad 0 \le i \le I, \qquad \bar{Z}_j^* := \sum_{i=0}^{I} \frac{q_{i,j}}{q_{\bullet j}} Z_{ij}, \quad 0 \le j \le J,$$

$$\bar{\bar{Z}}^* := \sum_{i=0}^{I} \frac{q_{i\bullet}}{q_{\bullet\bullet}} \bar{Z}_i,$$

which do not depend on the structural parameters.

As shown in Appendix B, with these notations an unbiased estimate for τ^2, χ^2 and σ^2 can be obtained, as the solution of the following system of linear equations

$$\begin{cases} & q_{\bullet\bullet}\left(1-\overline{h}_{(i)}^{(j)}\right)\widehat{\chi}^2 & +\,\widehat{\sigma}^2\,(J_\bullet - I - 1) & = \sum_{i,j} q_{i,j}\,(Z_{i,j} - \bar{Z}_i^*)^2 \\ q_{\bullet\bullet}\left(1-\overline{h}_{(j)}^{(i)}\right)\widehat{\tau}^2 & +\,q_{\bullet\bullet}\left(1-\overline{h}_{(j)}^{(i)}\right)\widehat{\chi}^2 & +\,\widehat{\sigma}^2\,(J_\bullet - J - 1) & = \sum_{i,j} q_{i,j}\,(Z_{i,j} - \bar{Z}_j^*)^2 \\ q_{\bullet\bullet}\left(1-h^{(i)}\right)\widehat{\tau}^2 & +\,q_{\bullet\bullet}\left(\overline{h}_{(i)}^{(j)} - h^{(d)}\right)\widehat{\chi}^2 & +\,\widehat{\sigma}^2\, I & = \sum_{i} q_{i\bullet}\,(\bar{Z}_i^* - \bar{\bar{Z}}^*)^2 \end{cases} \tag{46}$$

where

$$\overline{h}_{(i)}^{(j)} = \sum_i \frac{q_{i\bullet}}{q_{\bullet\bullet}} h_i^{(j)} \quad \text{with} \quad h_i^{(j)} := \sum_j \frac{q_{i,j}^2}{q_{i\bullet}^2},$$

$$\overline{h}_{(j)}^{(i)} = \sum_j \frac{q_{\bullet j}}{q_{\bullet\bullet}} h_j^{(i)} \quad \text{with} \quad h_j^{(i)} := \sum_i \frac{q_{i,j}^2}{q_{\bullet j}^2},$$

and

$$h^{(i)} := \sum_i \frac{q_{i\bullet}^2}{q_{\bullet\bullet}^2}, \qquad h^{(d)} := \sum_{t=0}^{I} \frac{\left(\sum_{i+j=t} q_{i,j}\right)^2}{q_{\bullet\bullet}^2}.$$

We observe that $J_\bullet = (J+1)(I+1-J/2)$ is the total number of nonzero weights of all accident years.

Numerical Rule. If a solution of (46) for either variance component turns out negative, it is set equal to zero.

Remark . The coefficients $h^{(i)}, h^{(d)}, \overline{h}_{(j)}^{(i)}, \overline{h}_{(i)}^{(j)}$ can be interpreted as Herfindahl indices or weighted averages of Herfindahl indices. As it is well-known, Herfindahl indices are measures of concentration. Referring to the distribution $\{x_1, \ldots, x_n\}$ of a given quantity x, the Herfindahl index

$$h := \sum_{k=1}^{n} \frac{x_k^2}{x_\bullet^2},$$

provides a measure of how much the distribution of x is concentrated. One has minimum concentration $h = 1/n$ for $x_k \equiv x_0$ and maximum concentration $h = 1$ if only one of the x_k is different from zero.

7.2.2. *Variance parameters estimates with the AY/CY approach.* An alternative approach to the estimation of variance components is obtained by arranging the data by accident year and calendar year. The calendar year is indexed as $t = 0, \ldots, T$, where $T = I$. Now, we transpose our definitions for AY/DY indices (i, j) to AY/CY indices (i, t). We have

$$\widetilde{q}_{i,t} = q_{i,i+j} \text{ and } \widetilde{Z}_{it} = Z_{i,i+j} \quad \text{for} \quad i + j \leq I,$$

and

$$\widetilde{q}_{i,t} = 0 \text{ and } \widetilde{Z}_{it} = 0 \quad \text{for} \quad t \leq i - 1 \text{ or } t \geq i + J + 1\,.$$

Hence, all our summations can be taken from $t = 0$ to $t = T$. Let T_i, $i \in \{0, \ldots, I\}$, the number of nonzero weights in AY i and $T_\bullet = \sum_{i=0}^{i} T_i$. Obviously $T_i = J_i$ as defined in *7.2.1*. For simplicity of notation, we write from here on $q_{i,t}$ for $\widetilde{q}_{i,t}$ and $Z_{i,t}$ for $\widetilde{Z}_{i,t}$.

Unbiased estimates of τ^2, χ^2 and σ^2 can then be obtained as the solution of the following system of linear equations

$$\begin{cases} & q_{\bullet\bullet}\left(1 - \overline{h}^{(t)}_{(i)}\right)\widehat{\chi}^2 & + \widehat{\sigma}^2\,(T_\bullet - I - 1) & = \sum_{i,t} q_{it}\,(Z_{it} - \bar{Z}^*_i)^2 \\ q_{\bullet\bullet}\left(1 - \overline{h}^{(i)}_{(t)}\right)\widehat{\tau}^2 & & + \widehat{\sigma}^2\,(T_\bullet - T - 1) & = \sum_{i,t} q_{it}\,(Z_{it} - \bar{Z}^*_t)^2 \\ q_{\bullet\bullet}\,(1 - h^{(i)})\,\widehat{\tau}^2 & + q_{\bullet\bullet}\left(\overline{h}^{(t)}_{(i)} - h^{(t)}\right)\widehat{\chi}^2 & + \widehat{\sigma}^2\, I & = \sum_i q_{i\bullet}\,(\bar{Z}^*_i - \bar{\bar{Z}}^*)^2 \end{cases} \tag{47}$$

where

$$\overline{h}^{(t)}_{(i)} = \sum_i \frac{q_{i\bullet}}{q_{\bullet\bullet}}\, h^{(t)}_i \quad \text{with} \quad h^{(t)}_i := \sum_t \frac{q^2_{i,t}}{q^2_{i\bullet}},$$

$$\overline{h}^{(i)}_{(t)} = \sum_t \frac{q_{\bullet t}}{q_{\bullet\bullet}}\, h^{(i)}_t \quad \text{with} \quad h^{(i)}_t := \sum_i \frac{q^2_{i,t}}{q^2_{\bullet t}},$$

and

$$h^{(t)} := \sum_t \frac{q^2_{\bullet t}}{q^2_{\bullet\bullet}}\,.$$

Since, the equalities hold

$$\begin{gathered} h^{(t)} = h^{(d)}, \quad h^{(t)}_i = h^{(j)}_i, \quad \overline{h}^{(t)}_{(i)} = \overline{h}^{(j)}_{(i)}, \quad T_\bullet = J_\bullet, \\ \sum_{i,t} q_{it}\,(Z_{it} - \bar{Z}^*_i)^2 = \sum_{i,j} q_{ij}\,(Z_{ij} - \bar{Z}^*_i)^2, \end{gathered} \tag{48}$$

the first and the third equations in the estimation system (47) are the same as in the estimation system (46) and the only difference in the two systems is given by the second equation.

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

Remarks

• The system (47) is similarly obtained as (46), see Appendix B, and the numerical rule applies equally to (47). Also the coefficients $h_t^{(i)}$ and $\overline{h}_{(t)}^{(i)}$ can be interpreted in terms of Herfindahl indexes.

• For $\chi^2 = 0$, the first and third equation in estimation system (46) or (47) provide the classical Bühlmann–Straub estimators for τ^2 and σ^2.

8. Numerical results

8.1. Applying the model to reference data

In Wüthrich and Merz (2008), Example 4.63, a numerical illustration of the BSCR model is provided. Since in this paper we refer to BSCR as a "control model", we use the same data for our numerical example concerning the ADR model. The triangle of observed incremental claims, with $I = J = 9$, and the corresponding *a priori* estimates a_i are reported in Table 2.

As previously emphasized, this paper concentrates on showing how stochastic diagonal effects influence reserve estimates and their prediction error in the — theoretical — situation where the quotas γ_j are known. The known γ_j are also those of Example 4.63 in Wüthrich and Merz (2008), obtained by Chain Ladder estimators according to (39). This also allows us to compare the results of the ADR model with the corresponding figures provided by a standard stochastic Chain Ladder model. The numerical values of the γ_j quotas are given in Table 3. The weights $w_{i,j}$ are then computed as $w_{ij} = a_i \, \gamma_j$.

For the estimation of the variance parameters, we adopted the AY/CY approach (the motivations of this choice will be apparent in Section 8.2). The values of the relevant Herfindahl indexes, which determine the values of the coefficients in Equations (47), are as follows

$$h^{(i)} = 0.1015845, \quad h^{(t)} = 0.1014692, \quad \bar{h}_{(t)}^{(i)} = 0.4959558, \quad \bar{h}_{(i)}^{(t)} = 0.4958533.$$

Solving Equations (47), we obtained the following estimates for the variance parameters

$$\widehat{\tau} = 0.0496097, \quad \widehat{\chi} = 0.0575456, \quad \widehat{\sigma} = 83.233023\,. \tag{49}$$

The estimates for the credibility weights are reported in the first column of Table 4. The corresponding homogeneous estimator of μ_0 is $\widehat{\mu}_0 = 0.8820442$. In Table 4, the reserves and the square root of the MSEPs (as well as the percentage coefficients of variation) in the inhomogeneous and the homogeneous case are also reported.

For the purpose of comparison, we also computed the reserves and the MSEPs estimated with the BSCR model and with the Time Series Chain Ladder (TSCL) model (see Buchwalder *et al.* (2006)). For the BSCR, we use the estimators for the variance parameters τ^2 and σ^2 and the homogeneous estimator for

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

TABLE 2

A PRIORI ESTIMATES OF THE ULTIMATE CLAIM AND OBSERVED INCREMENTAL CLAIMS.

i	a_i	$X_{i,0}$	$X_{i,1}$	$X_{i,2}$	$X_{i,3}$	$X_{i,4}$	$X_{i,5}$	$X_{i,6}$	$X_{i,7}$	$X_{i,8}$	$X_{i,9}$
0	11,653,101	5,946,975	3,721,237	895,717	207,761	206,704	62,124	65,813	14,850	11,129	15,814
1	11,367,306	6,346,756	3,246,406	723,221	151,797	67,824	36,604	52,752	11,186	11,646	
2	10,962,965	6,269,090	2,976,223	847,053	262,768	152,703	65,445	53,545	8,924		
3	10,616,762	5,863,015	2,683,224	722,532	190,653	132,975	88,341	43,328			
4	11,044,881	5,778,885	2,745,229	653,895	273,395	230,288	105,224				
5	11,480,700	6,184,793	2,828,339	572,765	244,899	104,957					
6	11,413,572	5,600,184	2,893,207	563,114	225,517						
7	11,126,527	5,288,066	2,440,103	528,042							
8	10,986,548	5,290,793	2,357,936								
9	11,618,437	5,675,568									

TABLE 3

CHAIN LADDER DEVELOPMENT PATTERN.

γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9
0.5900	0.2904	0.0684	0.0217	0.0144	0.0069	0.0051	0.0011	0.0010	0.0014

TABLE 4

RESERVES AND MSEP IN THE ADR MODEL.

i	α_i	$\widehat{R}_i$	$\text{msep}^{1/2}_{R_i}(\widehat{R}_i)$	(%)	$\widehat{R}^{\text{hom}}_i$	$\text{msep}^{1/2}_{R_i}(\widehat{R}^{\text{hom}}_i)$	(%)
0	0.4405	0	0	.	0	0	.
1	0.4090	15,155	10,620	70.1	14,031	10,623	75.7
2	0.3952	26,683	13,743	51.5	24,757	13,749	55.5
3	0.3867	36,544	16,219	44.4	33,825	16,229	48.0
4	0.3848	91,926	26,089	28.4	85,000	26,132	30.7
5	0.3829	170,354	35,945	21.1	157,395	36,054	22.9
6	0.3769	320,635	50,801	15.8	295,551	51,088	17.3
7	0.3668	511,867	67,539	13.2	468,989	68,170	14.5
8	0.3487	1,208,764	113,536	9.4	1,107,452	115,623	10.4
9	0.3047	4,620,160	316,789	6.9	4,229,107	327,843	7.8
Total		7,002,087	407,426	5.8	6,416,109	426,609	6.6

μ_0 as standard in the Bühlmann–Straub model and obtain

$$\widehat{\mu}_0^{\text{BS}} = 0.8810151, \quad \widehat{\tau}^{\text{BS}} = 0.0595243, \quad \widehat{\sigma}^{\text{BS}} = 104.01929\,.$$

The results in the BSCR model are reported in Table 5. The reserves and the MSEPs with the TSCL model are reported in Table 6. Since the uncertainty in the development pattern is missing both in the ADR and the BSCR, the MSEP in these models should be compared with the process error component in the TSCL model. To this aim, we calculated the decomposition of the MSEP of the TSCL model in the process error and the estimation error part.

In order to discuss the results for the reserve estimates, it is useful to consult the numerical details of the credibility decomposition introduced in Section 5.2. As explained there, the reserve estimates in both the ADR and the BSCR model are credibility mixtures of a projective-type reserve $\widehat{R}_i^{\text{pro}}$, based only on the triangle data, and a BF-type reserves $\widehat{R}_i^{\text{all}}$, based on external information (see expression (36)). Since the development pattern used in our numerical example is derived by the CL method, the projection factors $(1-\beta_{I-i})/\beta_{I-i}$ are equal to the CL projection factors $\prod_{j=I-i}^{J-1} f_j$. Then, in the BSCR model the projective reserves have the usual CL form $\widehat{R}_i^{\text{pro}} = \widehat{R}_i^{\text{CL}} = C_{i,I-i} \prod_{j=I-i}^{J-1} \widehat{f}_j$, while in the

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

TABLE 5

RESERVES AND MSEP IN THE BSCR MODEL.

i	α_i	$\widehat{R}_i$	$\text{msep}^{1/2}_{R_i}(\widehat{R}_i)$	(%)	$\widehat{R}_i^{\text{hom}}$	$\text{msep}^{1/2}_{R_i}(\widehat{R}_i^{\text{hom}})$	(%)
0	0.7924	0	0	.	0	0	.
1	0.7880	15,338	13,216	86.2	14,931	13,216	88.5
2	0.7817	26,419	17,108	64.8	25,718	17,109	66.5
3	0.7760	35,219	20,191	57.3	34,217	20,192	59.0
4	0.7819	87,511	32,243	36.8	85,035	32,246	37.9
5	0.7873	161,074	44,160	27.4	156,568	44,167	28.2
6	0.7838	298,051	61,499	20.6	289,272	61,520	21.3
7	0.7756	477,205	80,460	16.9	461,874	80,507	17.4
8	0.7600	1,109,352	125,486	11.3	1,071,689	125,669	11.7
9	0.6917	4,202,908	276,469	6.6	4,027,964	278,257	6.9
Total		6,413,077	326,040	5.1	6,167,268	329,031	5.3

TABLE 6

RESERVES AND MSEP IN THE TSCL MODEL.

		$\text{msep}^{1/2}_{R_i}(\widehat{R}_i)$					
i	$\widehat{R}_i$	Process	(%)	Estimation	(%)	Prediction	(%)
1	15,126	191	1.3	187	1.2	268	1.8
2	26,257	742	2.8	535	2.0	915	3.5
3	34,538	2,669	7.7	1,493	4.3	3,059	8.9
4	85,302	6,832	8.0	3,392	4.0	7,628	8.9
5	156,494	30,478	19.5	13,517	8.6	33,341	21.3
6	286,121	68,212	23.8	27,286	9.5	73,467	25.7
7	449,167	80,076	17.8	29,675	6.6	85,398	19.0
8	1,043,242	126,960	12.2	43,903	4.2	134,337	12.9
9	3,950,815	389,783	9.9	129,770	3.3	410,818	10.4
Total	6,047,064	424,380	7.0	185,026	3.1	462,961	7.7

ADR model they are obtained as $\widehat{R}_i^{\text{pro}} = \bar{\bar{C}}_{i,I-i} \prod_{j=I-i}^{J-1} \widehat{f}_j$, where $\bar{\bar{C}}_{i,I-i}$ is the generalized, or adjusted, diagonal defined by (37). The values of the observed diagonal and the ADR adjusted diagonal are illustrated in Figure 1, where the vertical bars refer to the observed values $C_{i,I-i}$ and the continuous line with dots refers to the adjusted values $\bar{\bar{C}}_{i,I-i}$. For comparison, the dotted line with stars illustrates the values of the adjusted diagonal in the Cape Cod method, as defined in (38) (for the Cape Cod correction factor we obtained $\bar{Z}^{CC} = 0.883973$). This graphical illustration clearly shows that the ADR adjustment has a strong smoothing effect on the observed diagonal values.

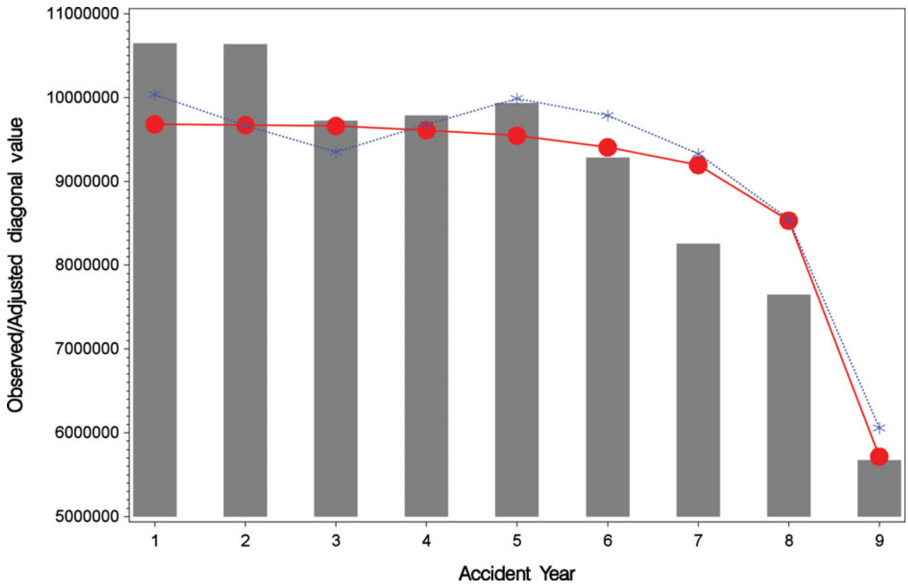

FIGURE 1: Observed and adjusted diagonal values in the ADR and the Cape Cod model (Color online).

The projective reserve estimates $\widehat{R}_i^{\text{pro}}$ corresponding to the diagonal values illustrated in Figure 1 are reported in the first part of Table 7. In the second part of Table 7, the allocative reserve estimates $\widehat{R}_i^{\text{all}}$ are also reported. Here, we provide the Bornhuetter–Ferguson reserves computed with $\mu_0 = 1$ and with the homogeneous μ_0 estimate both in the ADR and the BSCR model:

$$\widehat{R}_i^{\text{BF}/1} = a_i\,(1-\beta_{I-i}), \quad \widehat{R}_i^{\text{BF}/\widehat{\mu}_0} = a_i\,(1-\beta_{I-i})\,\widehat{\mu}_0, \quad \widehat{R}_i^{\text{BF}/\widehat{\mu}_0^{\text{BS}}} = a_i\,(1-\beta_{I-i})\,\widehat{\mu}_0^{\text{BS}}.$$

Given the reserve components reported in Table 7, the results on the reserve estimates in Tables 4 and 5 can be commented as follows. The inhomogeneous and homogeneous reserve estimates in the ADR model are a credibility mix of the R_i^{pro} reserves in the first column of Table 7 and the R_i^{all} reserves in the fourth and the fifth column, respectively, of the same table. Since the BF/1 reserves are higher than the BF/$\widehat{\mu}_0$ reserves, the inhomogeneous ADR reserves are higher than the homogeneous ADR reserves. A similar argument holds for the BSCR model. The difference between the reserves in the ADR model and the BSCR model, both in the inhomogeneous and homogeneous case, can be explained by the level of the credibility weights. Due to the additional uncertainty generated by the diagonal random effects, the credibility weights in ADR model are substantially smaller than in BSCR model. Hence, ADR provides reserves closer to the corresponding BF-type reserves and therefore closer to the *a priori* values. The difference between the total reserves based on ADR and those based on BSCR is 9.2% in the inhomogeneous case and 4.0% in the homogeneous case.

TABLE 7

PROJECTIVE AND ALLOCATIVE COMPONENTS OF THE RESERVE ESTIMATES.

	Projective Reserves $\widehat{R}_i^{\text{pro}}$			Allocative Reserves $\widehat{R}_i^{\text{all}}$		
i	ADR	CL	CC	BF/1	BF/$\widehat{\mu}_0$	BF/$\widehat{\mu}_0^{\text{BS}}$
1	13,754	15,126	14,254	16,125	14,223	14,206
2	23,878	26,257	23,866	26,999	23,814	23,786
3	34,317	34,538	33,216	37,576	33,143	33,105
4	83,778	85,302	84,361	95,434	84,177	84,079
5	150,349	156,494	157,369	178,024	157,025	156,842
6	289,942	286,121	301,705	341,306	301,047	300,696
7	500,276	449,167	507,480	574,090	506,373	505,782
8	1,163,742	1,043,242	1,165,647	1,318,646	1,163,104	1,161,747
9	3,979,360	3,950,815	4,215,123	4,768,385	4,205,926	4,201,019
Total	6,239,396	6,047,064	6,503,021	7,356,584	6,488,832	6,481,262

The main difference between ADR and BSCR concerns the prediction error. We find that the MSEP is substantially bigger for ADR than for BSCR, the difference between the square roots being 25.0% in the inhomogeneous case and 30.0% in the homogeneous case.

Comparing the MSEP based on the ADR model and the MSEP based on the TSCL model, we find that the process error component of MSEP in TSCL is higher than the MSEP in the inhomogeneous ADR and slightly lower in the homogeneous ADR, the difference between the square roots being 4.2% and −0.5% in the inhomogeneous and the homogeneous case, respectively.

8.2. Testing the estimation approaches by simulation

In order to empirically test the efficiency of the estimation approaches for μ_0 as well as the variance components described in Section 7, we applied the estimation procedures to simulated data. For the simulation, we used the data in Section 8.1 as reference data. Then we generated new run-off triangles of incremental payments by Equation (6), taking the a_i priors from Table 2 and the γ_j quotas from Table 3 and using the variance components estimates as in (49). The random variables η, ζ, ε in (6) were generated as independent and normally distributed, i.e.

$$\eta_i \sim \mathrm{N}(\mu_0, \tau^2), \quad \zeta_{i+j} \sim \mathrm{N}(0, \chi^2), \quad \varepsilon_{i,j} \sim \mathrm{N}(0, 1), \qquad 0 \le i + j \le I\,.$$

We also set $\mu_0 = 1$.

With this choice of the true values for the parameters μ_0, τ^2, χ^2, σ^2, we generated 100,000 10 × 10 triangles of simulated data and applied the estimation Equations (46) and (47) to each triangle[1]. The results of the estimation under the AY/DY approach (Equations 46) are summarized in the first part of Table

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

8. The corresponding results under the AY/CY approach (Equations 47) are reported in the second part of Table 8. These results are reported under the heading "ADR". Under the heading "BSCR" we also provide, for comparison, the estimates for μ_0, τ and σ obtained with the BSCR model. As it is also the case for the Bühlmann–Straub estimators, also the AY/DY and AY/CY approaches can provide negative values for the variance parameters estimates. In our sample statistics, these values have not been set at zero. The number of negative estimation results is also reported in the table.

The results provided in the table show that both the AY/DY and AY/CY estimation procedure reproduce *unbiased* estimates of the true values of the parameters (sample mean is correct). However, the AY/CY approach is *more efficient* (smaller sample variance). Moreover, if we use the estimation procedure from BSCR (i.e. the Bühlmann–Straub procedure), we miss χ and we have great bias for τ and σ.

Table 9 illustrates the results obtained by posing $\chi = 0$ (i.e. assuming that BSCR is the true model) and using the same values for τ and σ. As expected, the Bühlmann–Straub estimation procedure results to be the correct approach in this case. However, also the estimation procedures from ADR provide satisfactory results, producing a negligible value of χ and little loss of efficiency for the other variance components.

Summarizing the results of the simulation study

I. If the true model includes stochastic diagonal effects:
- both estimation procedures AY/DY and AY/CY derived from the ADR model provide *unbiased* parameter estimates but AY/CY is more *efficient*;
- the estimation procedure derived from the BSCR model performs badly and produces *biased* parameter estimates.

II. If the true model includes no stochastic diagonal effects:
- the procedure derived from BSCR is best;
- the procedure AY/CY derived from ADR is still *unbiased* and loses *little* efficiency in comparison with best estimators derived from BSCR.

Conclusions

The theoretical results and the results of the numerical example lead to the following conclusions:

1. The ADR model, which includes stochastic diagonal effects, is tractable in closed form using the credibility theory approach. Since diagonal effects induce correlation within calendar years, the covariance matrix of the incremental loss ratios is not block diagonal and the credibility formulae are more involved than in the classical theory. However, in the case of given development pattern, closed form expressions can be obtained for the reserve and MSEP estimates, both in the inhomogeneous and the homogeneous case.

TABLE 8

ESTIMATION RESULTS UNDER THE AY/DY AND THE AY/CY APPROACH (100,000 SIMULATIONS).

	μ_0		τ		χ	σ	
True Values:	1		0.0496097		0.0575456	83.233023	
	AY/DY Approach						
	ADR	BSCR	ADR	BSCR	ADR	ADR	BSCR
# Negatives:	0	0	13,824	979	16,152	20	0
Sample Mean:	0.9994591	1.0001122	0.0495926	0.0594948	0.0574970	83.29557	104.03776
Sample CoVa:	0.1273799	0.0255935	0.9775685	0.6123772	1.0681165	0.5142962	0.2734148
	AY/CY Approach						
	ADR	BSCR	ADR	BSCR	ADR	ADR	BSCR
# Negatives:	0	0	7,585	979	3,227	256	0
Sample Mean:	1.0001112	1.0001122	0.0496043	0.0594948	0.0574661	83.32062	104.03776
Sample CoVa:	0.0255554	0.0255935	0.8002220	0.6123772	0.6815610	0.3292820	0.2734148

TABLE 9

ESTIMATION RESULTS UNDER THE AY/DY AND THE AY/CY APPROACH WITH $\chi^2 = 0$.

	μ_0		τ		χ	σ	
True Values:	1		0.0496097		0	83.233023	
	AY/DY Approach						
	ADR	BSCR	ADR	BSCR	ADR	ADR	BSCR
# Negatives:	0	0	4,954	842	45,478	1	0
Sample Mean:	1.0016314	1.0000119	0.0496336	0.0496142	0.0024269	83.284829	83.243255
Sample CoVa:	0.4867062	0.0176713	0.696665	0.6016151	−452.216	0.4980504	0.2111396
	AY/CY Approach						
	ADR	BSCR	ADR	BSCR	ADR	ADR	BSCR
# Negatives:	0	0	1,013	842	52,653	2	0
Sample Mean:	1.000013	1.0000119	0.049639	0.0496142	0.0027436	83.296379	83.243255
Sample CoVa:	0.0176809	0.0176713	0.6098596	0.6016151	−104.0526	0.2496164	0.2111396

2. Alternative sets of unbiased estimators for the variance parameters of the ADR model can be obtained using two different procedures, the AY/DY and the AY/CY procedure. A simulation exercise suggests that the AY/CY approach provides more efficient estimators.
3. The credibility decomposition of the reserve estimate into a projective and an allocative component — which is typical in the BSCR model — can be extended to the ADR model through the definition of a "generalized diagonal".
4. The inclusion of random diagonal effects gives more weight to the allocative reserve estimates of the Bornhuetter–Ferguson type.
5. More importantly, such inclusion significantly increases the MSEP, both for the inhomogeneous and the homogeneous reserve estimates.

ACKNOWLEDGEMENT

This paper has undergone a thorough process of evolution, which was triggered by the Editor. His comments and suggestions have helped us to largely improve the contents of the article. We also thank two anonymous referees for their useful comments.

NOTE

1. Referring to a generic parameter π (i.e. μ_0, τ^2, χ^2 or σ^2), let us denote by $\widetilde{\pi}_k$ the estimate of π in the kth simulation. With n simulations, the following sample statistics have been computed

$$\text{sample mean:} \qquad m = \frac{1}{n}\sum_{k=1}^{n}\widetilde{\pi}_k$$

$$\text{sample variance:} \qquad q^2 = \frac{1}{n-1}\sum_{k=1}^{n}(\widetilde{\pi}_k - m)^2$$

$$\text{standard error:} \qquad \epsilon = \frac{q}{\sqrt{n-1}}$$

$$\text{relative standard error:} \qquad \frac{\epsilon}{m}$$

$$\text{coefficient of variation:} \qquad \text{CoVa} = \frac{\epsilon}{m}\sqrt{n-1} = \frac{q}{m}$$

REFERENCES

BORNHUETTER, R.L. and FERGUSON, R.E. (1972) The actuary and IBNR. *Proc. CAS*, **LIX**, pp. 181–195.

BUCHWALDER, M., BÜHLMANN, H., MERZ, M. and WÜTHRICH, M.V. (2006) The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin,* **36**(2), 521–542.

BÜHLMANN, H. (1983) Estimation of IBNR reserves by the methods chain ladder, Cape Cod and complementary loss ratio. *International Summer School 1983,* unpublished.

BÜHLMANN, H. and GISLER, A. (2005) *A Course in Credibility Theory and its Applications*. Berlin-Heidelberg: Springer-Verlag.

HENDERSON, H.V. and SEARLE, S.R. (1980) *On Deriving the Inverse of a Sum of Matrices*. Biometrics Unit Series, BU-647-M, Biometrics Unit, Ithaca, New York: Cornell University.

JESSEN, A.H. and RIETDORF, N. (2011) Diagonal effects in claims reserving. *Scandinavian Actuarial Journal*, **2011**(1), 21–37.

SALUZ, A., BÜHLMANN, H., GISLER, A. and MORICONI, F. (2014) *Bornhuetter-Ferguson reserving method with repricing. Presented to the 1st EAJ Conference*, Lausanne, September 2012.

SHI, P., BASU, S. and MEYERS, G. (2012) A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal,* **16**(1), 29–51.

SCHMIDT, K.D. (2006) Methods and models for loss reserving based on run-off triangles: A unifying survey. CAS Forum (Fall), 269–317.

WÜTHRICH, M.V. (2012) Discussion paper on "A Bayesian log-normal model for multivariate loss reserving". *North American Actuarial Journal,* **16**(1), 398–401.

WÜTHRICH, M.V. (2013) *Calendar year dependence modeling in run-off triangles.* ASTIN Colloquium May 2013*,* The Hague.

WÜTHRICH, M.V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Wiley Finance.

WÜTHRICH, M.V. and MERZ, M. (2012) Full and one-year Runoff risk in the credibility-based additive loss reserving method. *Journal of Applied Stochastic Models in Business and Industry,* **28**(4), 362–380.

HANS BÜHLMANN
Department of Mathematics
ETH Zurich
8092 Zurich, Switzerland
E-Mail: hbuhl@math.ethz.ch

FRANCO MORICONI (Corresponding author)
Dipartimento di Economia
Università di Perugia
Via A. Pascoli, 1
06123 Perugia, Italy
E-Mail: franco.moriconi1@unipg.it

APPENDIX A

For $\chi^2 = 0$, Theorem 4.1 provides the classical credibility theory. Without diagonal effects the covariance matrix $\boldsymbol{\Omega}$, with the LR/TB ordering, is a block diagonal matrix, where each block corresponds to an accident year. Then all elements $\omega_{(i,j),(k,l)}$ with $i \neq k$ are zero and the blocks have the form

$$B_i = \left(\omega_{(i,j),(i,l)}\right)_{0 \leq j,l \leq \iota(i)}, \quad 0 \leq i \leq I,$$

with

$$\omega_{(i,j),(i,l)} = \tau^2 + \frac{\sigma^2}{w_{i,j}} \mathbb{I}_{j,l}.$$

As it is well-known, if $\boldsymbol{\Omega}$ is a block diagonal matrix with blocks B_i, the inverse matrix $\boldsymbol{\Omega}^{-1}$ is also a block diagonal, with blocks B_i^{-1}. A convenient expression for B_i^{-1} is obtained by observing that B_i can be given by the sum

$$B_i = D_i + \tau^2 \mathbf{1}_i \mathbf{1}_i',$$

where $\mathbf{1}_i := (1, \ldots, 1)' \in \mathbb{R}^{\iota(i)+1}$ and D_i is the diagonal matrix

$$D_i = (\sigma^2/w_{i,0}, \ldots, \sigma^2/w_{i,\iota(i)}) \, \mathrm{I}_i,$$

where I_i is the identity matrix with size $\iota(i) + 1$. It can be shown (see e.g. identity (3) in Henderson and Searle (1980)) that

$$B_i^{-1} = D_i^{-1} - \frac{\tau^2}{1 + \tau^2 \mathbf{1}_i' D_i^{-1} \mathbf{1}_i} D_i^{-1} \mathbf{1}_i \mathbf{1}_i' D_i^{-1}.$$

This expression is also referred to as Sherman–Morrison formula. Let us introduce the precisions

$$\pi_0 := \frac{1}{\tau^2}, \qquad \pi_{i,j} := \frac{1}{\chi^2 + \sigma^2/w_{i,j}}, \qquad 0 \leq i \leq I, \, 0 \leq j \leq \iota(i).$$

For $\chi^2 = 0$, $\pi_{i,j}$ simplifies to

$$\pi_{i,j} := \frac{w_{i,j}}{\sigma^2}, \quad 0 \leq i \leq I, \, 0 \leq j \leq \iota(i);$$

hence, we have

$$D_i^{-1} = \boldsymbol{\pi}_i' \, \mathrm{I}_i, \quad \text{with} \quad \boldsymbol{\pi}_i := (\pi_{i,0}, \ldots, \pi_{i,\iota(i)})',$$

and

$$B_i^{-1} = \boldsymbol{\pi}_i' \, \mathrm{I}_i - \frac{1}{\pi_0 + \pi_{i,[\iota(i)]}} (\boldsymbol{\pi}_i' \, \mathrm{I}_i) \mathbf{1}_i \mathbf{1}_i' (\boldsymbol{\pi}_i' \, \mathrm{I}_i),$$

or

$$B_i^{-1} = \boldsymbol{\pi}_i' \, \mathrm{I}_i - \frac{C_i}{\pi_0 + \pi_{i,[\iota(i)]}},$$

where $C_i := (\pi_{i,j} \, \pi_{i,l})_{0 \leq i,l \leq \iota(i)}$. Hence, the generic element of block B_i^{-1}, $0 \leq i \leq I$, has the form

$$\omega^{(-1)}_{(i,j),(i,l)} = \pi_{i,j} \, \mathbb{I}_{j,l} - \frac{\pi_{i,j} \, \pi_{i,l}}{\pi_0 + \pi_{i,[\iota(i)]}}, \qquad 0 \leq j, l \leq \iota(i).$$

Summing over the column (or the line) $l = 0, \ldots, \iota(i)$ of block B_i^{-1}, one obtains

$$\sum_{l=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)} = \pi_{i,j} \, \frac{\pi_0}{\pi_0 + \pi_{i,[\iota(i)]}},$$

and summing over the block

$$\sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)} = \frac{\pi_0 \, \pi_{i,[\iota(i)]}}{\pi_0 + \pi_{i,[\iota(i)]}} .$$

In classical credibility theory, the credibility weights are defined as

$$\alpha_i := \frac{\tau^2}{\tau^2 + \sigma^2 / w_{i,[\iota(i)]}}, \qquad 0 \le i \le I,$$

which, in terms of precisions, is equal to

$$\alpha_i = \frac{\pi_{i,[\iota(i)]}}{\pi_0 + \pi_{i,[\iota(i)]}}, \qquad 0 \le i \le I .$$

Then for the sum over the block B_i^{-1}, we have

$$\sum_{j=0}^{\iota(i)} \sum_{l=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)} = \frac{\alpha_i}{\tau^2}, \qquad 0 \le i \le I,$$

and the sum over $j = 0, \ldots, \iota(i)$ in block B_i^{-1} is given by

$$\sum_{j=0}^{\iota(i)} \omega^{(-1)}_{(i,j),(i,l)} = \frac{\alpha_i}{\tau^2} \, \frac{\pi_{i,j}}{\pi_{i,[\iota(i)]}} = \frac{\alpha_i}{\tau^2} \, \frac{w_{i,j}}{w_{i,[\iota(i)]}}, \qquad 0 \le i \le I .$$

Then, we are led to the standard expression for the inhomogeneous credibility estimator

$$\widehat{\eta}_i = \mu_0 \, (1 - \alpha_i) + \alpha_i \, \bar{Z}_i \quad \text{with} \quad \bar{Z}_i := \sum_{j=0}^{\iota(i)} \frac{w_{j,l}}{w_{i,[\iota(i)]}} Z_{i,j} .$$

More formally, Theorem 4.1 provides the classical results for the inhomogeneous case in credibility theory if one poses, for $0 \le i, k \le I$, $0 \le j \le \iota(i)$ and $0 \le l \le \iota(k)$

$$\omega_{(i,j),(k,l)} = \left(\pi_{i,j} \, \mathbb{I}_{j,l} - \frac{\pi_{i,j} \, \pi_{i,l}}{\pi_0 + \pi_{i,[\iota(i)]}} \right) \mathbb{I}_{i,k} .$$

In this case, the classical results for the homogeneous case are immediately provided by Theorem 4.4, which gives

$$\widehat{\eta}_i^{\text{hom}} = \widehat{\mu}_0 \, (1 - \alpha_i) + \alpha_i \, \bar{Z}_i \quad \text{with} \quad \widehat{\mu}_0 = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha_\bullet} \, \bar{Z}_i .$$

APPENDIX B

We adopt here the definition (45) of the weights $q_{i,j}$. Hence, the weighted sums over i are extended from 0 to I, the weighted sums over j are extended from 0 to J and the weighted sums over t are extended from 0 to T.

AY/DY approach

To derive the first equation in (46), we consider the sums of squares, $0 \leq i \leq I$

$$\mathrm{SS}_i := \sum_j q_{i,j}\,(Z_{i,j} - \bar{Z}_i^*)^2 \quad \text{where} \quad \bar{Z}_i^* := \sum_j \frac{q_{i,j}}{q_{i\bullet}}\, Z_{ij}\,.$$

By (9), given the properties of the η, ζ, ε variables and the independence assumptions, we have for $i + j \leq I$

$$\mathbf{E}(Z_{i,j}) = \mathbf{E}(\bar{Z}_i^*) = \mu_0, \quad \mathbf{Var}(Z_{i,j}) = \tau^2 + \chi^2 + \frac{\sigma^2}{q_{i,j}}, \quad \mathbf{Var}(\bar{Z}_i^*) = \tau^2 + \chi^2\, h_i^{(j)} + \frac{\sigma^2}{q_{i\bullet}}\,.$$

(The explicit expressions of the Herfindahl indices are provided in the text). The sums of squares can be written as

$$\mathrm{SS}_i = \sum_j q_{i,j}\,(Z_{i,j} - \mu_0)^2 - q_{i\bullet}\,(\bar{Z}_i^* - \mu_0)^2,$$

then taking the expectation

$$\begin{aligned}
\mathbf{E}(\mathrm{SS}_i) &= \sum_j q_{i,j}\,\mathbf{E}[(Z_{i,j} - \mu_0)^2] - q_{i\bullet}\,\mathbf{E}[(\bar{Z}_i^* - \mu_0)^2] \\
&= \sum_j q_{i,j}\,\mathbf{Var}(Z_{i,j}) - q_{i\bullet}\,\mathbf{Var}(\bar{Z}_i^*) \\
&= \sum_j q_{i,j}\left(\tau^2 + \chi^2 + \frac{\sigma^2}{q_{i,j}}\right) - q_{i\bullet}\left(\tau^2 + \chi^2\, h_i^{(j)} + \frac{\sigma^2}{q_{i\bullet}}\right) \\
&= q_{i\bullet}\left(1 - h_i^{(j)}\right)\chi^2 + \sigma^2\,(J_i - 1)\,.
\end{aligned}$$

Summing over i

$$\mathbf{E}\left(\sum_i \mathrm{SS}_i\right) = q_{\bullet\bullet}\left(1 - \sum_i \frac{q_{i,\bullet}}{q_{\bullet\bullet}}\, h_i^{(j)}\right)\chi^2 + \sigma^2\,(J_\bullet - I - 1)\,.$$

The first of the estimation Equation (46) is then obtained by substituting the observed value of $\sum_i \mathrm{SS}_i$ to the expectation.

For the second equation in (46), we consider the sums of squares, $0 \leq j \leq J$

$$\mathrm{SS}_j := \sum_i q_{i,j}\,(Z_{i,j} - \bar{Z}_j^*)^2 \quad \text{where} \quad \bar{Z}_j^* := \sum_i \frac{q_{i,j}}{q_{\bullet j}}\, Z_{ij}\,.$$

https://doi.org/10.1017/asb.2015.3 Published online by Cambridge University Press

Now we have, for $i + j \leq I$

$$\mathbf{E}(\bar{Z}_j^*) = \mu_0, \quad \mathbf{Var}(\bar{Z}_j^*) = (\tau^2 + \chi^2)\, h_j^{(i)} + \frac{\sigma^2}{q_{\bullet j}}.$$

Since

$$\mathrm{SS}_j = \sum_i q_{i,j}\,(Z_{i,j} - \mu_0)^2 - q_{\bullet j}\,(\bar{Z}_j^* - \mu_0)^2,$$

taking the expectation, we obtain

$$\begin{aligned}
\mathbf{E}(\mathrm{SS}_j) &= \sum_i q_{i,j}\,\mathbf{Var}(Z_{i,j}) - q_{\bullet j}\,\mathbf{Var}(\bar{Z}_j^*) \\
&= \sum_i q_{i,j}\left(\tau^2 + \chi^2 + \frac{\sigma^2}{q_{i,j}}\right) - q_{\bullet j}\left((\tau^2 + \chi^2)\, h_j^{(i)} + \frac{\sigma^2}{q_{\bullet j}}\right) \\
&= q_{\bullet j}\left(1 - h_j^{(i)}\right)(\tau^2 + \chi^2) + \sigma^2\,(I_j - 1),
\end{aligned}$$

where I_j denotes the number of nonzero elements of DY j. Summing over j

$$\mathbf{E}\left(\sum_j \mathrm{SS}_j\right) = q_{\bullet\bullet}\left(1 - \sum_j \frac{q_{\bullet j}}{q_{\bullet\bullet}}\, h_j^{(i)}\right)(\tau^2 + \chi^2) + \sigma^2\,(I_\bullet - J - 1).$$

The second estimation equation in (46) is then obtained by equating the expectation to the observed value of $\sum_j \mathrm{SS}_j$ (and observing that $I_\bullet = J_\bullet$).

The third equation in (46) is obtained by considering the total sums of squares

$$\mathrm{TSS} := \sum_i q_{i\bullet}\,(\bar{Z}_i^* - \bar{\bar{Z}}^*)^2 \quad \text{where} \quad \bar{\bar{Z}}^* := \sum_i \sum_j \frac{q_{i,j}}{q_{\bullet\bullet}}\, Z_{ij}.$$

Now we have, for $i + j \leq I$,

$$\mathbf{E}(\bar{\bar{Z}}^*) = \mu_0, \quad \mathbf{Var}(\bar{\bar{Z}}^*) = \tau^2\, h^{(i)} + \chi^2\, h^{(d)}.$$

Since

$$\mathrm{TSS} = \sum_i q_{i\bullet}\,(\bar{Z}_i^* - \mu_0)^2 - q_{\bullet\bullet}\,(\bar{\bar{Z}}^* - \mu_0)^2,$$

taking the expectation, we obtain

$$\begin{aligned}
\mathbf{E}(\mathrm{TSS}) &= \sum_i q_{i\bullet}\,\mathbf{Var}(\bar{Z}_i^*) - q_{\bullet\bullet}\,\mathbf{Var}(\bar{\bar{Z}}^*) \\
&= \sum_i q_{i\bullet}\left(\tau^2 + \chi^2\, h_i^{(j)} + \frac{\sigma^2}{q_{i\bullet}}\right) - q_{\bullet\bullet}\left(\tau^2\, h^{(i)} + \chi^2\, h^{(d)}\right) \\
&= q_{\bullet\bullet}\left(1 - h^{(i)}\right)\tau^2 + q_{\bullet\bullet}\left(\bar{h}^{(i)} - h^{(d)}\right)\chi^2 + \sigma^2\, I.
\end{aligned}$$

The last equation of system (46) is then obtained by substituting the expectation with the observed value of TSS. ∎

AY/CY approach

For the AY/CY approach, data are reorganized by accident year and calendar year. In this new language, we have for $t \leq I$

$$Z_{i,t} = \eta_i + \zeta_t + \frac{\sigma}{\sqrt{q_{i,t}}}\, \varepsilon_{i,t}\,.$$

Since the values on the right-hand side of the first and the third equation in (47) are clearly the same as in (46), in the AY/CY approach the first and third estimation equations are the same as the corresponding equations in (46) written in the new language (the correspondence between the two languages is summarized by equalities (48) in the text). Then, we have only to prove the second equation in (47). To this aim, we consider the sum of squares

$$\mathrm{SS}_t := \sum_t q_{i,t}\,(Z_{i,t} - \bar{Z}_t^*)^2 = \sum_t q_{i,t}\,(Z_{i,t} - \mu_0)^2 - q_{i\bullet}\,(\bar{Z}_t^* - \mu_0)^2\,.$$

Since, for $t \leq I$

$$\mathbf{E}(Z_{i,t}) = \mathbf{E}(\bar{Z}_t^*) = \mu_0, \quad \mathbf{Var}(Z_{i,t}) = \tau^2 + \chi^2 + \frac{\sigma^2}{q_{i,t}}, \quad \mathbf{Var}(\bar{Z}_t^*) = \tau^2\, h_t^{(i)} + \chi^2 + \frac{\sigma^2}{q_{i\bullet}},$$

by taking the expectation, we obtain

$$\begin{aligned}\mathbf{E}(\mathrm{SS}_t) &= \sum_t q_{i,t}\,\mathbf{Var}(Z_{i,t}) - q_{i\bullet}\,\mathbf{Var}(\bar{Z}_t^*)\\ &= \sum_t q_{i,t}\left(\tau^2 + \chi^2 + \frac{\sigma^2}{q_{i,t}}\right) - q_{i\bullet}\left(\tau^2\, h_t^{(i)} + \chi^2 + \frac{\sigma^2}{q_{i\bullet}}\right)\\ &= q_{i\bullet}\left(1 - h_t^{(i)}\right)\,\tau^2 + \sigma^2\,(T_i - 1)\,.\end{aligned}$$

Summing over i, we get

$$\mathbf{E}\left(\sum_t \mathrm{SS}_t\right) = q_{\bullet\bullet}\left(1 - \sum_t \frac{q_{t\bullet}}{q_{\bullet\bullet}}\, h_t^{(i)}\right)\,\tau^2 + \sigma^2\,(T_\bullet - T - 1),$$

which finally leads to the second equation in (47). ■