Stochastic Claims Reserving with Credible Priors

Franco Moriconi
University of Perugia

SAA Annual Meeting – Fribourg, September 2, 2016
The Claims Reserving Problem: standard setup

We have the rectangle of *incremental claims* $X_{i,j}$ where:

- $i = 0, 1, \ldots, I$ is the index of the *accident year* (AY)
- $j = 0, 1, \ldots, J$ is the index of the *development year* (DY)

(the index $t = i + j$ denotes the *calendar year*)

The *cumulative claims* of AY $i$ up to DY $j$ are defined as:

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k},$$

We assume that all claims are settled after development year $J$ where $J \leq I$.

Assume we are at time $I$ and let us consider the sets:

$$D_I = \{X_{i,j}; i + j \leq I, j \leq J\}, \quad D_I^c = \{X_{i,j}; i + j > I, i \leq I, j \leq J\}.$$  

⇒

- the random variables $X_{ij} \in D_I$ are *observed*  
- the random variables $X_{ij} \in D_I^c$ – the outstanding claims – must be *predicted*.
Denoting by $\hat{X}_{i,j}$ the prediction of $X_{i,j} \in D_I^c$ we then have the reserve estimates:

$$\hat{R}_i = \sum_{j=I-i+1}^{J} \hat{X}_{i,j}, \quad i = I - J + 1, \ldots, I,$$

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

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

---

<table>
<thead>
<tr>
<th>$X_{0,0}$</th>
<th>$X_{0,1}$</th>
<th>...</th>
<th>.</th>
<th>...</th>
<th>$X_{0,J-1}$</th>
<th>$X_{0,J}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{1,0}$</td>
<td>$X_{1,1}$</td>
<td>...</td>
<td>.</td>
<td>...</td>
<td>$X_{1,J-1}$</td>
<td>$X_{1,J}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{I-J,0}$</td>
<td>$X_{I-J,1}$</td>
<td>...</td>
<td>.</td>
<td>...</td>
<td>$X_{I-J,J-1}$</td>
<td>$X_{I-J,J}$</td>
</tr>
<tr>
<td>$X_{I-J+1,0}$</td>
<td>$X_{I-J+1,1}$</td>
<td>...</td>
<td>.</td>
<td>...</td>
<td>$X_{I-J+1,J-1}$</td>
<td>$X_{I-J+1,J}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{i,0}$</td>
<td>$X_{i,1}$</td>
<td>...</td>
<td>$X_{i,I-i}$</td>
<td>$X_{i,I-i+1}$</td>
<td>...</td>
<td>$X_{i,J-1}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>.</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{I,0}$</td>
<td>$X_{I,1}$</td>
<td>...</td>
<td>.</td>
<td>...</td>
<td>$X_{I,J-1}$</td>
<td>$X_{I,J}$</td>
</tr>
</tbody>
</table>
Simplified notations

For semplicity sake we assume $I = J$ hereafter ($D_I$ is a “triangle”). Moreover we denote ($0 \leq i \leq I$):

- $D_i := C_{i, I-i}$: cumulative claims on the last observed diagonal
- $U_i := C_{i, J}$: ultimate cost

Hence: $\hat{R}_i = \hat{U}_i - D_i$

We also use the notation: $x_{[n]} := \sum_{k=0}^{n} x_k$
Loss Development Methods

Let us define \((0 \leq i \leq I):\)

- loss development quota \(\gamma_{i,j}\): percentage of \(U_i\) paid in DY \(j = 0, 1, \ldots, J\),
- cumulative loss development quota \(\beta_{i,j} := \gamma_{i,[j]}\).

Of course \(\beta_{i,J} \equiv \gamma_{i,[J]} \equiv 1\).

Basic assumption

There exists a loss development pattern independent of the accident year:

\[ \gamma := \{\gamma_j; 0 \leq j \leq J\} . \]

The cumulative quotas are given by \(\beta_j := \gamma_{[j]}\), with \(\beta_J = 1\).

By the previous definitions, we can derive a number of logical relations which can be used as the starting point for defining different stochastic models.

- E.g. \(X_{i,j} = U_i \gamma_j, \ 0 \leq i, j \leq I\), suggests a cross classified model.
Moreover, let us consider the relation \((0 \leq i \leq I)\):

\[
U_i = \frac{D_i}{\beta_{I-i}}
\]

- By adding and subtracting \(D_i\):

\[
U_i = D_i + D_i \frac{1 - \beta_{I-i}}{\beta_{I-i}} \quad \rightarrow \quad \text{Projective Reserve: } R_i = D_i \frac{1 - \beta_{I-i}}{\beta_{I-i}}
\]

\(\rightarrow (1 - \beta_{I-i})/\beta_{I-i}: \text{projection factor}\)

- Also:

\[
U_i = D_i + U_i (1 - \beta_{I-i}) \quad \rightarrow \quad \text{Allocative Reserve: } R_i = U_i (1 - \beta_{I-i})
\]

\(\rightarrow (1 - \beta_{I-i}): \text{“still to come percentage”}\)

! All these representations are equivalent in a deterministic setting, but under uncertainty they originate a number of different stochastic models and estimation methods!
**Projective (or multiplicative) methods**

- Chain-Ladder (CL), and all its versions:

\[
1 - \frac{\hat{\beta}_j^{CL}}{\hat{\beta}_j^{CL}} = \prod_{k=j}^{J-1} \hat{f}_k^{CL} - 1, \quad 0 \leq j \leq J - 1,
\]

where:

\[
\hat{f}_j^{CL} = \frac{C_{[I-i+1],j+1}}{C_{[I-i+1],j}},
\]

is the CL estimator of the *development factor* from DY \(j\) to DY \(j + 1\).

Chain Ladder reserving is the prototype of the projective reserve.

- Models based on estimates of *development factors* \(\hat{f}_j := \hat{\beta}_{j+1}/\hat{\beta}_j\) different from CL.
Allocative (or additive) methods

- Bornhuetter-Ferguson (BF):
  \[ \hat{R}_i = a_i (1 - \beta_{I-i}) , \]
  where \( a_i \) is a prior estimate of \( U_i \).
  BF reserving is the prototype of the allocative reserve

- Cape Cod (CC):
  \[ \hat{R}_i = \hat{\kappa}^{CC} P_i (1 - \beta_{I-i}) , \]
  where \( P_i \) are the premiums received and \( \hat{\kappa}^{CC} \) is an average loss ratio estimate (e.g. \( \hat{\kappa}^{CC} = D_{[I]}/\sum_i P_i \beta_{I-i} \)).
  However CC can also be considered as a projection method applied to a “robusted diagonal” \( \hat{D}^{CC}_i := \hat{\kappa}^{CC} P_i \beta_{I-i} \).

- Additive Loss Reserving (ALR).
Under uncertainty, the fundamental question is:

*What kind of information is the estimation of \( U_i \) and \( \gamma_j \) based on?*

Generally:

- Projective methods only use data from “triangles“ (eventually only \( D_i \)).
  The development pattern can also be obtained by the triangle.
  How additional information can be incorporated into the model?

- Allocative methods use *prior estimates* \( a_i \) for the ultimate loss \( U_i \), but need additional information concerning the development pattern.
  The priors \( a_i \) are determined by *exogenous information* (external expert opinion, pricing, ...).
  How the “triangle” information can be incorporated?

  Correct theoretical answer: *Bayesian Approach*

However for practical reasons “suboptimal” solutions are often used.
Hybrid models: some examples

- Naïve (classical) approach to BF model: development pattern derived by Chain-Ladder. Widely used, but intrinsecally inconsistent: prior information not used to estimate $\hat{\gamma}_j$.

- Hybrid Chain Ladder, by Arbenz and Salzmann (2012): weighted average of stochastic CL and BF.

- ALR model (Schmidt, 2006), BF-type model by Mack (2008). The estimate $\hat{\gamma}_j$ uses mixed information:

$$
\hat{\gamma}_j = \frac{X_{[I-j],j}}{a_{[I-j]}}.
$$

We shall denote by:

$$
\hat{\gamma}^{RE}_j = \frac{X_{[I-j],j}}{a_{[I-j]}} \left( \sum_{k=0}^{J} \frac{X_{[I-k],j}}{a_{[I-k]}} \right)^{-1},
$$

the normalized version of this “raw estimator”.

- BF-type model by Saluz, Gisler and Wüthrich (2011).
• *Linear hybrid model*

\[ \hat{R}_i = \alpha_i D_i \frac{1 - \beta_{I-i}}{\beta_{I-i}} + (1 - \alpha_i) a_i (1 - \beta_{I-i}) \]

How are the weights \( \alpha_i \) determined?

*Ad hoc* solution: \( \alpha_i \) should increase with the development of \( C_{i,j} \) since we obtain better information on \( U_i \) with increasing time \( j \)

\[ \Rightarrow \text{Benktander(1976)-Hovinen(1981): } \alpha_i := \beta_{I-i}. \]

*Can we adopt a more Bayesian attitude?*

**Additional assumptions**

In order to introduce a link between the projective and the allocative point of view we introduce a property common to all the development years:

\[ \begin{align*} &! \text{ The risk characteristics of accident year } i \text{ are described by a } \textit{latent variable} \Theta_i \\ &! \text{ Conditionally, given } \Theta_i, \text{ the incremental losses } X_{i,j} \text{ are } \textit{independent} \end{align*} \]

\[ \Rightarrow \text{ Conditionally Independent Loss Increments (CILI) models.} \]
• The approach we consider to the claims reserving problem with CILI models is based on the *linear credibility methods*, which restrict the search for the best estimators to the class of estimators which are *linear functions of the observations*.

• The credibility approach can be considered as an *approximation* of the fully Bayesian approach. In particular cases (Exponential Dispersion Family with its associates conjugates) credibility estimators are also *exact* Bayesian estimators. 
(A review of fully Bayesian methods in claims reserving can be found in M.V. Wüthrich, M-Merz, 2008, sec. 4.3-4.4. See also R. Verral, 2007)

• The most representative example of a CILI model for claims reserving is the **Bühlmann-Straub Credibility Reserving Model** (BSCR).
The Bühlmann-Straub
Credibility Reserving Model

Basic references:
Model assumptions

**A1.** Let $\Theta := \{\Theta_i; \ 0 \leq i \leq I\}$. There exist positive parameters $a_0, \ldots, a_I, \gamma_0, \ldots, \gamma_J,$ and $\sigma^2$, with $\sum_{j=0}^{J} \gamma_j = 1$, such that for $0 \leq i \leq I$ and $0 \leq j \leq J$:

$$E(X_{ij}|\Theta) = a_i \gamma_j \Theta_i,$$

and:

$$\text{Var}(X_{ij}|\Theta) = a_i \gamma_j \sigma^2.$$

**A2.** Let $X_i = \{X_{i,j}; \ 0 \leq j \leq J\}$. For $0 \leq i \leq I$ the pairs $\{\Theta_i, X_i\}$ are independent. Moreover, all $\Theta$ variables are independent, with:

$$E(\Theta_i) = \mu_0, \quad \text{Var}(\Theta_i) = \tau^2, \quad 0 \leq i \leq I.$$

A1, A2 are the usual assumptions in the classical Bühlmann-Straub model, reformulated here in the claims reserving context:

- the parameters $\{a_i; \ 0 \leq i \leq I\}$ are the given priors,
- the development pattern $\gamma = \{\gamma_j; \ 0 \leq j \leq J\}$ is assumed to be known.

The parameters $\mu_0, \tau^2, \sigma^2$ must be estimated from the data.
The previous assumptions can be equivalently formulated in terms of \textit{Incremental Loss Ratios} (i.e. normalized incremental losses):

\[ Z_{ij} := \frac{X_{ij}}{a_i \gamma_j}, \quad 0 \leq i \leq I, \quad 0 \leq j \leq J. \]

An estimator \( \hat{\Theta}_i \) of \( \Theta_i \) is an estimator of \( \mathbb{E}(Z_{i,j}|\Theta_i) \).

\textbf{Prediction}

For \( X_{ij} \in \mathcal{D}_I \) (the outstanding claims) we consider the predictor:

\[ \hat{X}_{ij} = a_i \gamma_j \hat{\Theta}_i. \]

Then we have the following predictor for the reserve in accident year \( i = 1, \ldots, I \) and for the total reserve:

\[ \hat{R}_i = \sum_{j=I-i+1}^{J} \hat{X}_{ij} = a_i (1 - \beta_{I-i}) \hat{\Theta}_i, \quad \hat{R} = \sum_{i=1}^{I} \hat{R}_i. \]

- We want to derive estimates for the reserve of single accident years and the total reserve, as well as estimates of the corresponding prediction error (\textit{mean square error of prediction}, MSEP). The estimators we are looking for are in the class of the credibility estimators.
Credible priors
If $\hat{\Theta}_i$ is a credibility estimator, we define $\hat{a}_i := a_i \hat{\Theta}_i$ the credible prior of the ultimate loss $U_i$ of accident year $i$. One has:

$$\hat{X}_{ij} = \hat{a}_i \gamma_j, \quad \hat{R}_i = \hat{a}_i (1 - \beta_{I-i}).$$

Credibility Estimators
If the prior mean $\mu_0 = E(\Theta_i)$ is known (typically $\mu_0 = 1$), we have:

- **Inhomogeneous Bühlmann-Straub estimator.** The best linear inhomogeneous estimator for $\Theta_i$ ($0 \leq i \leq I$) is given by:

$$\hat{\Theta}_{i}^{inh} = \alpha_i \overline{Z}_i + (1 - \alpha_i) \mu_0,$$

where $\alpha_i$ is the credibility weight of accident year $i$:

$$\alpha_i = \frac{a_i \beta_{I-i}}{a_i \beta_{I-i} + \sigma^2/\tau^2},$$

and $\overline{Z}_i$ is the weighted average of the incremental loss ratios observed in AY $i$:

$$\overline{Z}_i = \sum_{j=0}^{I-i} \frac{\gamma_j}{\beta_{I-i}} Z_{i,j} = \frac{D_i}{a_i \beta_{I-i}}.$$
If $\mu_0$ is not known one has:

- **Homogeneous Bühlmann-Straub estimator**. The best linear homogeneous estimator for $\Theta_i$ ($0 \leq i \leq I$) is given by:

  \[
  \hat{\Theta}_i^{\text{hom}} = \alpha_i \bar{Z}_i + (1 - \alpha_i) \hat{\mu}_0,
  \text{ where } \hat{\mu}_0 = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha[I]} \bar{Z}_i.
  \]

### Credibility decomposition of reserve estimates

In the homogeneous case, the credible prior for accident year $i$ is:

\[
\hat{\alpha}_i^{\text{hom}} := a_i \hat{\Theta}_i^{\text{hom}}
= \alpha_i a_i \bar{Z}_i + (1 - \alpha_i) a_i \hat{\mu}_0
= \alpha_i \frac{D_i}{\beta_{I-i}} + (1 - \alpha_i) a_i \hat{\mu}_0.
\]

$\implies$ credibility mixture of projective and allocative reserve:

\[
\hat{R}_i^{\text{hom}} := \sum_{j=I-i+1}^{J} \hat{X}_{i,j}^{\text{hom}} = \alpha_i \frac{D_i (1 - \beta_{I-i})}{\beta_{I-i}} + (1 - \alpha_i) a_i \hat{\mu}_0 (1 - \beta_{I-i}).
\]

\[
\begin{align*}
\text{Projective Reserve} & \quad \text{Allocative Reserve}
\end{align*}
\]
• If $\tau^2 = 0$ one has $\alpha_i \equiv 0$ hence:

  $\Rightarrow$ In the inhomogeneous case (with $\mu_0 = 1$) one obtains the classical Bornhuetter-
  Ferguson model:

  $$\hat{R}_{i}^{inh} = a_i (1 - \beta_{I-i}).$$

  $\Rightarrow$ In the homogeneous case one has:

  $$\hat{R}_{i}^{hom} = a_i \hat{\mu}_0 (1 - \beta_{I-i}) = a_i \kappa^{CC} (1 - \beta_{I-i}),$$

  where:

  $$\kappa^{CC} := \hat{\mu}_0 = \frac{\sum_{i=0}^{I} D_i}{\sum_{i=0}^{I} a_i \beta_{I-i}},$$

  is the Cape Cod estimate of the overall loss ratio.

  Also projective point of view by defining the robusted diagonal: $D^{CC} := a_i \hat{\mu}_0 \beta_{I-i}$.
**Numerical example**

Data (Example 4.63 in Wüthrich and Merz, 2008)

A priori estimates $a_i$ of the ultimate claim and observed incremental claims $X_{i,j}$

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

Results

**Estimated development patterns**

<table>
<thead>
<tr>
<th>$j$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{j}^{CL}$</td>
<td>1.4925</td>
<td>1.0778</td>
<td>1.02288</td>
<td>1.0148</td>
<td>1.0070</td>
<td>1.0052</td>
<td>1.0011</td>
<td>1.0011</td>
<td>1.0014</td>
<td></td>
</tr>
<tr>
<td>$\hat{c}_{j}^{CL}$</td>
<td>0.5900</td>
<td>0.2904</td>
<td>0.0684</td>
<td>0.0217</td>
<td>0.0144</td>
<td>0.0069</td>
<td>0.0051</td>
<td>0.0011</td>
<td>0.0010</td>
<td>0.0014</td>
</tr>
<tr>
<td>$\hat{c}_{j}^{RE}$</td>
<td>0.5860</td>
<td>0.2906</td>
<td>0.0694</td>
<td>0.0224</td>
<td>0.0151</td>
<td>0.0073</td>
<td>0.0055</td>
<td>0.0012</td>
<td>0.0011</td>
<td>0.0015</td>
</tr>
</tbody>
</table>
Credibility estimators (d.p.: $\hat{\gamma}^{CL}$)

By the usual BS estimators for the structural parameters:

$$\hat{\tau} = 0.0595, \hat{\sigma} = 104.01929, \hat{\mu}_0 = 0.88102.$$ 

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\alpha_i$</th>
<th>$\bar{Z}_i$</th>
<th>$\hat{\Theta}_{ih}^i$</th>
<th>$\hat{\Theta}_{ih}^i$</th>
<th>$a_i$</th>
<th>$\hat{a}_{ih}^i$</th>
<th>$\hat{a}_{ih}^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.7924</td>
<td>0.9567</td>
<td>0.9657</td>
<td>0.9410</td>
<td>11,653,101</td>
<td>11,252,979</td>
<td>10,965,073</td>
</tr>
<tr>
<td>1</td>
<td>0.7880</td>
<td>0.9381</td>
<td>0.9512</td>
<td>0.9260</td>
<td>11,367,306</td>
<td>10,812,560</td>
<td>10,525,829</td>
</tr>
<tr>
<td>2</td>
<td>0.7817</td>
<td>0.9725</td>
<td>0.9785</td>
<td>0.9526</td>
<td>10,962,965</td>
<td>10,727,703</td>
<td>10,442,964</td>
</tr>
<tr>
<td>3</td>
<td>0.7760</td>
<td>0.9192</td>
<td>0.9373</td>
<td>0.9106</td>
<td>10,616,762</td>
<td>9,950,833</td>
<td>9,667,869</td>
</tr>
<tr>
<td>4</td>
<td>0.7819</td>
<td>0.8938</td>
<td>0.9170</td>
<td>0.8910</td>
<td>11,044,881</td>
<td>10,127,949</td>
<td>9,841,357</td>
</tr>
<tr>
<td>5</td>
<td>0.7873</td>
<td>0.8791</td>
<td>0.9048</td>
<td>0.8795</td>
<td>11,480,700</td>
<td>10,387,587</td>
<td>10,097,017</td>
</tr>
<tr>
<td>6</td>
<td>0.7838</td>
<td>0.8383</td>
<td>0.8733</td>
<td>0.8475</td>
<td>11,413,572</td>
<td>9,967,090</td>
<td>9,673,507</td>
</tr>
<tr>
<td>7</td>
<td>0.7756</td>
<td>0.7824</td>
<td>0.8312</td>
<td>0.8045</td>
<td>11,126,527</td>
<td>9,248,782</td>
<td>8,951,648</td>
</tr>
<tr>
<td>8</td>
<td>0.7600</td>
<td>0.7911</td>
<td>0.8413</td>
<td>0.8127</td>
<td>10,986,548</td>
<td>9,242,775</td>
<td>8,928,979</td>
</tr>
<tr>
<td>9</td>
<td>0.6917</td>
<td>0.8285</td>
<td>0.8814</td>
<td>0.8447</td>
<td>11,618,437</td>
<td>10,240,620</td>
<td>9,814,360</td>
</tr>
</tbody>
</table>

All $\hat{\Theta}$ estimates are less then 1 $\Rightarrow$ priors are pessimistic: $\hat{a}_i < a_i$.

Moreover, since $\hat{\mu}_0 < \mu_0 = 1$, homogeneous reserves will be lower than inhomogeneous.
Summary of reserve estimates with different methods

Known development pattern: $\gamma^{CL}$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\text{BSCR}^{inh}$</th>
<th>$\text{BSCR}^{hom}$</th>
<th>BF</th>
<th>BH</th>
<th>CC</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15,338</td>
<td>14,931</td>
<td>16,125</td>
<td>15,128</td>
<td>14,254</td>
<td>15,126</td>
</tr>
<tr>
<td>2</td>
<td>26,419</td>
<td>25,718</td>
<td>26,999</td>
<td>26,259</td>
<td>23,866</td>
<td>26,257</td>
</tr>
<tr>
<td>3</td>
<td>35,219</td>
<td>34,217</td>
<td>37,576</td>
<td>34,549</td>
<td>33,216</td>
<td>34,538</td>
</tr>
<tr>
<td>4</td>
<td>87,511</td>
<td>85,035</td>
<td>95,434</td>
<td>85,389</td>
<td>84,361</td>
<td>85,302</td>
</tr>
<tr>
<td>5</td>
<td>161,074</td>
<td>156,568</td>
<td>178,024</td>
<td>156,828</td>
<td>157,369</td>
<td>156,494</td>
</tr>
<tr>
<td>6</td>
<td>298,051</td>
<td>289,272</td>
<td>341,306</td>
<td>287,771</td>
<td>301,705</td>
<td>286,121</td>
</tr>
<tr>
<td>7</td>
<td>477,205</td>
<td>461,874</td>
<td>574,090</td>
<td>455,613</td>
<td>507,480</td>
<td>449,167</td>
</tr>
<tr>
<td>8</td>
<td>1,109,352</td>
<td>1,071,689</td>
<td>1,318,646</td>
<td>1,076,297</td>
<td>1,165,647</td>
<td>1,043,242</td>
</tr>
<tr>
<td>9</td>
<td>4,202,908</td>
<td>4,027,964</td>
<td>4,768,385</td>
<td>4,286,358</td>
<td>4,215,123</td>
<td>3,950,815</td>
</tr>
<tr>
<td>Total</td>
<td>6,413,076</td>
<td>6,167,268</td>
<td>7,356,584</td>
<td>6,424,193</td>
<td>6,503,021</td>
<td>6,047,064</td>
</tr>
</tbody>
</table>

Total, $\hat{\gamma}^{RE}$

|       | 6,573,961        | 6,319,544         | 7,505,461 | 6,452,322 | 6,644,053 |

(BH: Benktander-Hovinen $\rightarrow \alpha_i = \beta_{I-i}$)
Mean Square Error of Prediction of the Reserves

The expressions for the MSEPs of the single accident year reserves are derived in Wüthrich and Merz (2008), Corollary 4.60.

The following expressions for the MSEPs of the total reserve are derived in Bühlmann and M. (2015), as a special case of Theorem 6.2 and Corollary 6.4:

\[
\text{msep}_R \left( \hat{R}^{inh} \right) = \sum_{i=1}^{I} \text{msep}_{R_i} \left( \hat{R}_i^{inh} \right),
\]

\[
\text{msep}_R \left( \hat{R}^{hom} \right) = \text{msep}_R \left( \hat{R}^{inh} \right) + \frac{\tau^2}{\alpha[I]} \left( \sum_{i=1}^{I} \alpha_i (1 - \alpha_i) (1 - \beta_{I-i}) \right)^2
\]

\[
= \sum_{i=1}^{I} \text{msep}_{R_i} \left( \hat{R}_i^{hom} \right) + 2 \frac{\tau^2}{\alpha[I]} \sum_{1 \leq i < k \leq I} a_i a_k (1 - \alpha_i)(1 - \alpha_k)(1 - \beta_{I-i})(1 - \beta_{d_k}).
\]

The last equality shows that in the homogeneous case the reserve estimates of different accident years are positively correlated.
Mean square errors of prediction in the BSCR model (known $\gamma = \gamma^{CL}$)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\hat{\bar{R}}_{i}^{inh}$</th>
<th>mse$^{1/2}(\hat{\bar{R}}_{i}^{inh})$ (%)</th>
<th>$\hat{\bar{R}}_{i}^{hom}$</th>
<th>mse$^{1/2}(\hat{\bar{R}}_{i}^{hom})$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15,338</td>
<td>13,216</td>
<td>86.2</td>
<td>14,931</td>
</tr>
<tr>
<td>2</td>
<td>26,419</td>
<td>17,108</td>
<td>64.8</td>
<td>25,718</td>
</tr>
<tr>
<td>3</td>
<td>35,219</td>
<td>20,191</td>
<td>57.3</td>
<td>34,217</td>
</tr>
<tr>
<td>4</td>
<td>87,511</td>
<td>32,243</td>
<td>36.8</td>
<td>85,035</td>
</tr>
<tr>
<td>5</td>
<td>161,074</td>
<td>44,160</td>
<td>27.4</td>
<td>156,568</td>
</tr>
<tr>
<td>6</td>
<td>298,051</td>
<td>61,499</td>
<td>20.6</td>
<td>289,272</td>
</tr>
<tr>
<td>7</td>
<td>477,205</td>
<td>80,460</td>
<td>16.9</td>
<td>461,874</td>
</tr>
<tr>
<td>8</td>
<td>1,109,352</td>
<td>125,486</td>
<td>11.3</td>
<td>1,071,689</td>
</tr>
<tr>
<td>9</td>
<td>4,202,908</td>
<td>276,469</td>
<td>6.6</td>
<td>4,027,964</td>
</tr>
<tr>
<td>Total</td>
<td>6,413,076</td>
<td>326,040</td>
<td>5.1</td>
<td>6,167,268</td>
</tr>
</tbody>
</table>
Reserves and MSEPs in the Time Series Chain Ladder model
(Buchwalder, Bühlmann, Merz and Wüthrich, 2006)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\hat{R}_i$</th>
<th>Process (%)</th>
<th>Estimation (%)</th>
<th>Prediction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15,126</td>
<td>191</td>
<td>187</td>
<td>268</td>
</tr>
<tr>
<td>2</td>
<td>26,257</td>
<td>742</td>
<td>535</td>
<td>915</td>
</tr>
<tr>
<td>3</td>
<td>34,538</td>
<td>2,669</td>
<td>1,493</td>
<td>3,059</td>
</tr>
<tr>
<td>4</td>
<td>85,302</td>
<td>6,832</td>
<td>3,392</td>
<td>7,628</td>
</tr>
<tr>
<td>5</td>
<td>156,494</td>
<td>30,478</td>
<td>13,517</td>
<td>33,341</td>
</tr>
<tr>
<td>6</td>
<td>286,121</td>
<td>68,212</td>
<td>27,286</td>
<td>73,467</td>
</tr>
<tr>
<td>7</td>
<td>449,167</td>
<td>80,076</td>
<td>29,675</td>
<td>85,398</td>
</tr>
<tr>
<td>8</td>
<td>1,043,242</td>
<td>126,960</td>
<td>43,903</td>
<td>134,337</td>
</tr>
<tr>
<td>9</td>
<td>3,950,815</td>
<td>389,783</td>
<td>129,770</td>
<td>410,818</td>
</tr>
<tr>
<td>Total</td>
<td>6,047,064</td>
<td>424,380</td>
<td>185,026</td>
<td>462,961</td>
</tr>
</tbody>
</table>
MSEP under the one-year view

For solvency purposes the MSEP s should be computed under a “one-year view”, i.e. the random variable to be considered should be the *Year-End Obligations* (next-diagonal payments plus residual reserve), or equivalently the one-year Claims Development Result (CDR).

In credibility-based claims reserving models, closed form expressions for MSEP s under the one-year view have been derived:

- in Bühmann, De Felice, Gisler, M. and Wüthrich (2008) for the Credibility Chain Ladder model,

- in Merz and Wüthrich (2011) for the Credibility-Based Additive Loss Leserving model.

A similar approach can be used in the BSCR model. The basic idea is to express the year-end estimate $\alpha_{i+1}^{I}$ of the credibility weight as a linear updating of the current estimate $\alpha_{i}^{I}$. 
BSCR model with unknown development pattern
(Work in progress, jointly with Hans Bühlmann and Mario Wüthrich)

If we relax the assumption of known development pattern, the quotas \( \gamma \) must be estimated from the data.

**Estimation of the development pattern**

We propose an iterative procedure. The basic idea is that if the parameters \( \Theta \) would be known, then the best linear unbiased estimator of \( \gamma_j \) would be:

\[
\gamma^{raw}_j := \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} a_i \Theta_i}, \quad j = 0, \ldots, J.
\]

Since the \( \Theta \) variables are unknown they are replaced by some estimates. Moreover, some kind of normalization is needed.

For the inhomogeneous/homogeneous case we choose the following estimator:

\[
\tilde{\gamma}_{inh/hom}^j = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} a_i \tilde{\Theta}_{inh/hom}^j} \left( \sum_{l=0}^{J} \frac{\sum_{i=0}^{I-l} X_{i,l}}{\sum_{i=0}^{I-l} a_i \tilde{\Theta}_{inh/hom}^l} \right)^{-1}, \quad j = 0, \ldots, J. \tag{1}
\]
a) The estimated pattern \( \hat{\gamma} \) will depend on \( \hat{\Theta} \), that is \( \hat{\gamma} = \hat{\gamma}(\hat{\Theta}) \);
b) in turn, \( \hat{\Theta} \) depends itself by \( \hat{\gamma} \), that is \( \hat{\Theta} = \hat{\Theta}(\hat{\gamma}) \).

**Definition.** We call \((\tilde{\Theta}, \tilde{\gamma})\) a **compatible pair** if \( \hat{\Theta}(\tilde{\gamma}) = \tilde{\Theta} \) and \( \hat{\gamma}(\tilde{\Theta}) = \tilde{\gamma} \).

**Iterative estimation algorithm** (homogeneous case):

0. **Initialization.** We choose an initial development pattern \( \gamma^{(0)} \) and pose \( \hat{\gamma} = \gamma^{(0)} \).

1. **Estimation of the structural parameters.** Using \( \hat{\gamma} \), estimates of the parameters \( \hat{\mu}_0, \hat{\tau}^2, \hat{\sigma}^2 \) are obtained by the data.

2. **Estimation of \( \Theta \).** Using \( \hat{\gamma} \) and the parameter estimates obtained in step 1, the estimates \( \hat{\Theta} \) are derived.

3. **Estimation of \( \gamma \).** Using \( \hat{\Theta} \), a new d.p. estimate \( \hat{\gamma}' \) is computed.

4. **Iteration.** If a given convergence criterion is not fulfilled, we return to step 1 assuming for \( \hat{\gamma} \) the development pattern \( \hat{\gamma}' \). Otherwise the iteration is stopped and we adopt the resulting estimates \( \hat{\Theta} \) and \( \hat{\gamma} \) as the compatible pair \((\tilde{\Theta}, \tilde{\gamma})\).

**convergence criterion:** \[
\max \left\{ \sqrt{ \sum_{j=0}^{J} (\hat{\gamma}_j^{(n)} - \hat{\gamma}_j^{(n-1)})^2 }, \sqrt{ \sum_{i=0}^{I} (\hat{\Theta}_i^{(n)} - \hat{\Theta}_i^{(n-1)})^2 } \right\} < \epsilon,
\]
Reserve estimates and MSEPs with unknown $\gamma$

*Inhomogeneous case.* After 5 iterations (with $\epsilon = 10^{-7}$) we obtained:

$$\hat{\tau} = 0.05926, \quad \hat{\sigma} = 103.76437.$$

*Homogeneous case.* After 4 iterations:

$$\hat{\tau} = 0.05926, \quad \hat{\sigma} = 103.76583, \quad \hat{\mu}_0 = 0.88133.$$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\alpha_i$</th>
<th>$\hat{R}_i^{inh}$</th>
<th>$\text{msep}_{\hat{R}_i}^{1/2}$ (%)</th>
<th>$\hat{R}_i^{hom}$</th>
<th>$\text{msep}_{\hat{R}_i}^{1/2}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.7917</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.7873</td>
<td>15,596</td>
<td>13,294</td>
<td>85.2</td>
<td>15,181</td>
</tr>
<tr>
<td>2</td>
<td>0.7810</td>
<td>26,844</td>
<td>17,203</td>
<td>64.1</td>
<td>26,128</td>
</tr>
<tr>
<td>3</td>
<td>0.7753</td>
<td>35,797</td>
<td>20,306</td>
<td>56.7</td>
<td>34,775</td>
</tr>
<tr>
<td>4</td>
<td>0.7812</td>
<td>88,896</td>
<td>32,416</td>
<td>36.5</td>
<td>86,373</td>
</tr>
<tr>
<td>5</td>
<td>0.7866</td>
<td>163,437</td>
<td>44,372</td>
<td>27.1</td>
<td>158,852</td>
</tr>
<tr>
<td>6</td>
<td>0.7831</td>
<td>301,931</td>
<td>61,742</td>
<td>20.4</td>
<td>293,014</td>
</tr>
<tr>
<td>7</td>
<td>0.7747</td>
<td>482,521</td>
<td>80,699</td>
<td>16.7</td>
<td>466,986</td>
</tr>
<tr>
<td>8</td>
<td>0.7590</td>
<td>1,117,632</td>
<td>125,627</td>
<td>11.2</td>
<td>1,079,625</td>
</tr>
<tr>
<td>9</td>
<td>0.6904</td>
<td>4,217,905</td>
<td>276,202</td>
<td>6.5</td>
<td>4,042,181</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6,450,559</td>
<td>326,035</td>
<td>5.1</td>
<td>6,203,114</td>
</tr>
<tr>
<td>Total, $\hat{\gamma}^{CL}$</td>
<td></td>
<td>6,413,076</td>
<td>326,040</td>
<td>5.1</td>
<td>6,167,268</td>
</tr>
</tbody>
</table>
Deriving the estimation error

If the development pattern is unknown, we need also an assessment of the prediction error on the reserve estimates due to the estimation error of $\gamma$.

We define this estimation error for the reserve estimate for accident year $i = 1, \ldots, I$ as:

$$ EE_i^{(\gamma)} := E \left( \hat{R}_i^{(\gamma)} - \tilde{R}_i \right)^2, $$

- $\hat{R}_i^{(\gamma)}$ is the reserve estimate obtained with the "true value" of $\gamma$,
- $\tilde{R}_i$ is the reserve estimate derived by the iterative algorithm.

For the total reserve we pose:

$$ EE^{(\gamma)} := \left( \sum_{i=1}^{I} E \left( \hat{R}_i^{(\gamma)} - \tilde{R}_i \right) \right)^2. $$

In order to derive an estimate of $EE_i^{(\gamma)}$ and $EE^{(\gamma)}$ we follow a parametric bootstrap simulation method partially similar to that used in Saluz, Bühlmann, Gisler and M. (2014).
The bootstrap simulation procedure

In the bootstrap simulation a large number of “pseudo-triangles” $\mathcal{D}_I$ is simulated and for each pseudo-triangle two reserve estimates are computed for each one of these pseudo-data:

- one estimate is obtained using as development pattern the estimate $\tilde{\gamma}$ resulting by the original iterative procedure (which is assumed to provide the “true development pattern”),

- the other estimate is obtained by a new iterative estimation procedure (therefore we run the estimation algorithm in each simulation), without re-estimation of the structural parameters.

An assessment of the estimation error is then obtained as the average of the square differences of the two reserve estimate.

Remark. Differently from the approach used in Saluz, Bühlmann, Gisler and M. (2014), in the bootstrap simulations we use the original estimates for the structural parameters in all the iterations in order to avoid to include the estimation error of these hyperparameters into the estimation error of $\gamma$. 
Let us denote by $\tilde{\tau}^2, \tilde{\sigma}^2, \tilde{\mu}_0$ the parameter estimates obtained for the compatible pair $(\tilde{\Theta}, \tilde{\gamma})$.

In the $s$-th simulation ($s = 1, \ldots, S$) we have the following steps:

1. The random variables $\Theta, \varepsilon$ are generated as independent and normally distributed, with:
   $$\Theta_i^{(s)} \sim N(\tilde{\mu}_0, \tilde{\tau}^2), \quad \varepsilon_{i,j}^{(s)} \sim N(1, 0), \quad 0 \leq i + j \leq I.$$

2. A pseudo-trapezoid of incremental claims is generated as:
   $$X_{i,j}^{(s)} = a_i \tilde{\gamma}_j \Theta_i^{(s)} + \sqrt{a_i \tilde{\gamma}_j} \tilde{\sigma} \varepsilon_{i,j}^{(s)}, \quad 0 \leq i + j \leq I.$$

3. The reserve estimates $\hat{R}_i^{(s)}$ and $\hat{R}^{(s)} = \sum_i \hat{R}_i^{(s)}$, are obtained by this data assuming the original development pattern $\tilde{\gamma}$ and using the original parameter estimates $\tilde{\tau}^2, \tilde{\sigma}^2, \tilde{\mu}_0$.

4. The previous iterative estimation procedure is applied with $\gamma^{(0)} = \tilde{\gamma}$ and where step 1 is skipped, so that the original parameter estimates $\tilde{\tau}^2, \tilde{\sigma}^2, \tilde{\mu}_0$ are used in each iteration. The simulated reserve estimates $\tilde{R}_i^{(s)}$ and $\tilde{R}^{(s)} = \sum_i \tilde{R}_i^{(s)}$ are then derived by the compatible pair $(\tilde{\Theta}^{(s)}, \tilde{\gamma}^{(s)})$ provided by the iterations.

5. The squared errors are computed:
   $$\Delta_i^{(s)} = (\hat{R}_i^{(s)} - \tilde{R}_i^{(s)})^2, \quad i = 1, \ldots, I, \quad \Delta^{(s)} = \left( \sum_{i=1}^{I} (\hat{R}_i^{(s)} - \tilde{R}_i^{(s)}) \right)^2.$$

The estimation errors are estimated as $\hat{EE}_i^{(\gamma)} = \frac{1}{S} \sum_{s=1}^{S} \Delta_i^{(s)}$ and $\hat{EE}^{(\gamma)} = \frac{1}{S} \sum_{s=1}^{S} \Delta^{(s)}$. 
MSEP in the BSCR model including $\gamma$ estimation error

$S = 10000, \ 3 \leq \text{num\_iterations} \leq 5$

<table>
<thead>
<tr>
<th>$i$</th>
<th>msep$^{1/2}(\hat{R}_i^{inh})$ (%)</th>
<th>msep$^{1/2}(\hat{R}_i^{hom})$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19,072 122.3</td>
<td>18,783 123.7</td>
</tr>
<tr>
<td>2</td>
<td>23,140 86.2</td>
<td>22,829 87.4</td>
</tr>
<tr>
<td>3</td>
<td>26,067 72.8</td>
<td>25,754 74.1</td>
</tr>
<tr>
<td>4</td>
<td>38,856 43.7</td>
<td>38,516 44.1</td>
</tr>
<tr>
<td>5</td>
<td>51,524 31.5</td>
<td>51,166 32.2</td>
</tr>
<tr>
<td>6</td>
<td>69,385 23.0</td>
<td>69,025 23.6</td>
</tr>
<tr>
<td>7</td>
<td>88,730 18.4</td>
<td>88,386 18.9</td>
</tr>
<tr>
<td>8</td>
<td>135,231 12.1</td>
<td>134,973 12.5</td>
</tr>
<tr>
<td>9</td>
<td>291,912 6.9</td>
<td>292,987 7.2</td>
</tr>
<tr>
<td>Total</td>
<td>395,536 6.1</td>
<td>395,910 6.4</td>
</tr>
<tr>
<td>Total no EE</td>
<td>326,035 5.1</td>
<td>329,048 5.3</td>
</tr>
</tbody>
</table>
The convergence issue

In all the numerical examples we considered, the iterative estimation algorithm ever converged to a unique point \((\tilde{\Theta}, \tilde{\gamma})\), independently of the initial d.p. \(\gamma^{(0)}\). However the convergence of the algorithm should be studied (and possibly proven) using proper theoretical arguments. Technically, one should prove that the algorithm is a contraction.

We studied the case without re-estimation of the structural parameters.

A proof of convergence is available (provided by M. Wüthrich) for the case \(I = J = 1\) (the 1-dimensional case, since \(\gamma_0 + \gamma_1 = 1\)).

For \(I, J > 1\) the functions \(\Theta(\gamma)\) and \(\gamma(\Theta)\) are difficult to be studied analitically and, at the moment, we can oly conjecture that the convergence to a unique pair holds in the general case.
Adding stochastic diagonal effects:

the BSCR Model with Additive Diagonal Risk (ADR Model)

Basic references:
Model assumptions

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

\[
E(X_{ij}|\Theta) = a_i \gamma_j (\eta_i + \zeta_{i+j})
\]

and:

\[
\text{Var}(X_{ij}|\Theta) = a_i \gamma_j \sigma^2.
\]

**A2.** All \( \eta, \zeta \) variables are independent, with:

\[
\begin{align*}
E(\eta_i) &= \mu_0, & \text{Var}(\eta_i) &= \tau^2, & 0 \leq i \leq I, \\
E(\zeta_{i+j}) &= 0, & \text{Var}(\zeta_{i+j}) &= \chi^2, & 0 \leq i \leq I, 0 \leq j \leq J.
\end{align*}
\]

We interpret:

\cdot \( \eta_i \): random effect of accident year \( i \),

\cdot \( \zeta_{i+j} \): random effect of calendar year \( t = i + j \) (random diagonal effect).

As usual the parameters \( a_i \) are the given priors, the development quotas \( \gamma \) are known and the parameters \( \mu_0, \tau^2, \chi^2, \sigma^2 \) must be estimated from the data.
Innovative aspects of the ADR model:

- A *calendar year effect* is included in the latent variables $\Theta$ which takes the form:
  \[ \Theta_{i,t} = \eta_i + \zeta_t \]
  for the calendar year $t = i + j$,
  thus the calendar year effect is *separated additively* from the (random) accident year effect.

- *Independence between accident years is relaxed*. Our assumptions imply the following correlation structure for the risk parameters:
  \[ \text{Cov}(\Theta_{it}, \Theta_{ks}) = \tau^2 I_{i,k} + \chi^2 I_{t,s}, \quad 0 \leq i, k \leq I, \quad 0 \leq t, s \leq I + J, \]
  with $I_{i,k} = 1$ if $i = k$ and 0 elsewhere.
  The correlation within the same accident year $i$ is induced by $\eta_i$ and is proportional to $\tau^2$; the correlation within the same calendar year $t$ is induced by $\zeta_t$ and is proportional to $\chi^2$.

! If $\chi^2 = 0$ the ADR model reduces to the BSCR model (with $\Theta_i \equiv \eta_i$).
Using a full Bayesian approach, diagonal risk effects are also treated by Shi, Basu and Meyer (2012) and by Wüthrich (2012, 2013).

Remark. The ADR assumptions implicitly provide an extension of the classical Bühlmann-Straub model that can be used also in applications different from claims reserving. For example, for the classical applications in *premium rating*:

- rewrite the assumptions for $Z_{i,t} := X_{i,t}/a_i \gamma_{t-i}$,
- interpret $Z_{i,t}$ as the loss ratio of company $i$ (risk $i$ in a *collective*) observed in year $t$,
- use $w_{i,t} := a_i \gamma_{t-i}$ as the associated weight.
Prediction
For $X_{ij} \in D^c_I$, we now consider the predictor $\hat{X}_{ij} = a_i \gamma_j \hat{\eta}_i$.

Again, the estimators we are looking for are in the class of the credibility estimators.

**Credible priors:** $\hat{a}_i := a_i \hat{\eta}_i \implies \hat{X}_{ij} = \hat{a}_i \gamma_j$, $\hat{R}_i = \hat{a}_i (1 - \beta_{I-i})$.

However, since the accident years are no more independent, credibility formulae are more complex than in the classical case.

**Covariance matrix of the observations**

For any pair of observations $X_{i,j}, X_{k,l} \in D_I$, we consider the covariance between the incremental loss ratios:

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

With this double-index notation, the covariance matrix of the observations is:

$$\Omega := (\omega_{(i,j),(k,l)})_{(i,j),(k,l) \in D_I}.$$

To fix ideas, we choose for $\Omega$ the left-to-right/top-to-bottom ordering.

Under the model assumptions one has:

$$\text{Cov}(Z_{ij}, Z_{kl}) = \tau^2 \mathbb{I}_{i,k} + \chi^2 \mathbb{I}_{i+j,k+l} + \frac{\sigma^2}{a_i \gamma_j} \mathbb{I}_{i,k} \mathbb{I}_{j,l}, \quad 0 \leq i, k \leq I, \ 0 \leq j, l \leq J.$$
### Structure of the $\Omega$ matrix with $I = J = 2$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\nu$</th>
<th>1 (0,0)</th>
<th>2 (0,1)</th>
<th>3 (0,2)</th>
<th>4 (1,0)</th>
<th>5 (1,1)</th>
<th>6 (2,0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(0,0)</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_0\gamma_0}$</td>
<td>$\tau^2$</td>
<td>$\tau^2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>(0,1)</td>
<td>$\tau^2$</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_0\gamma_1}$</td>
<td>$\tau^2$</td>
<td>$\chi^2$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>(0,2)</td>
<td>$\tau^2$</td>
<td>$\tau^2$</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_0\gamma_2}$</td>
<td>0</td>
<td>$\chi^2$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>4</td>
<td>(1,0)</td>
<td>0</td>
<td>$\chi^2$</td>
<td>0</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_1\gamma_0}$</td>
<td>$\tau^2$</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>(1,1)</td>
<td>0</td>
<td>0</td>
<td>$\chi^2$</td>
<td>$\tau^2$</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_1\gamma_1}$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>6</td>
<td>(2,0)</td>
<td>0</td>
<td>0</td>
<td>$\chi^2$</td>
<td>0</td>
<td>$\chi^2$</td>
<td>$\tau^2 + \chi^2 + \frac{\sigma^2}{a_2\gamma_0}$</td>
</tr>
</tbody>
</table>

- If $\chi^2 = 0$ (BSCR model) the covariance matrix $\Omega$ is **block diagonal**, where each block corresponds to an accident year.

  If $\Omega$ is a block diagonal matrix with blocks $B_i$, the inverse matrix $\Omega^{-1}$ is also block diagonal, with blocks $B_i^{-1}$.

  By this property, in the classical credibility theory (hence in the BSCR) all the relevant expressions of accident year $i$ (e.g. expressions for $\alpha_i, \overline{Z}_i$ and $\hat{\eta}_i$) are based only on the observations of accident year $i$.

- If $\chi^2 > 0$ (ADR model) $\Omega$ is not block diagonal, hence the relevant expressions for accident year $i$ depend now on all the observations.
Theorem 4.1. The best linear inhomogeneous estimator for \( \eta_i \), \( 0 \leq i \leq I \), is given by:

\[
\hat{\eta}_i^{inh} = \alpha_i \bar{Z}_i + \mu_0 (1 - \alpha_i)
\]

where:

\[
\alpha_i := \tau^2 \sum_{k=0}^{I} \sum_{l=0}^{I-k} \sum_{j=0}^{I-i} \omega_{(i,j),(k,l)}^{(-1)}, \quad \bar{Z}_i := \tau^2 \sum_{k=0}^{I} \sum_{l=0}^{I-k} \sum_{j=0}^{I-i} \omega_{(i,j),(k,l)}^{(-1)} Z_{k,l}.
\]

The \( \Omega^{-1} \) matrix with \( I = J = 2 \)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\multicolumn{2}{|c|}{\lambda} & \multicolumn{6}{|c|}{\nu} \\
\hline
& & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
1 & (0,0) & & & & & & \\
2 & (0,1) & & & & & & \\
3 & (0,2) & & & & & & \\
4 & (1,0) & & & & & & \\
5 & (1,1) & & & & & & \\
6 & (2,0) & & & & & & \\
\hline
\end{tabular}
\end{table}

\( \alpha_i \) is obtained by \( \Omega^{-1} \) by computing the sum of all the elements on the band corresponding to accident year \( i \) (and then multiplying by \( \tau^2 \)).
Theorem 4.4. The best linear homogeneous estimator for $\eta_i$, $0 \leq i \leq I$, is given by:

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

with

$$\hat{\mu}_0 = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha_i} \bar{Z}_i.$$

Credibility decomposition of reserve estimates

$$\hat{R}_{i}^{\text{hom}} := \sum_{j=I-i+1}^{J} \hat{X}_{i,j}^{\text{hom}} = \alpha_i \bar{Z}_i (1 - \beta_{I-i}) + (1 - \alpha_i) \hat{\mu}_0 (1 - \beta_{I-i})$$

Projective Reserve

Allocative Reserve

For $\tau^2 = 0$ we again obtain Bornhuetter-Ferguson model and Cape Cod model as special cases.

Mean Square Error of Prediction of the Reserves

Closed form expressions are obtained both in the inhomogeneous and the homogeneous case.
Parameter estimation

Estimates of $\tau^2, \chi^2, \sigma^2$ are obtained as the solution of a system of 3 linear equations involving sums of square errors.

A method typical in the analysis of variance:
Some types of sum of square errors (SS) are taken from the data and the expectation $E(\text{SS})$ of each of these sums is expressed as a function $f(\tau^2, \chi^2, \sigma^2)$. If the expectation $E(\text{SS})$ is replaced by the corresponding observed value $\text{SS}^*$, 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 $D_I$. Given model assumptions, the $f$ functions are linear, and we are led to a system of three independent linear equations. We obtain estimates for $\tau^2, \chi^2, \sigma^2$ by solving this system.

Because of the linearity of the equations these estimates are unbiased.

If $\chi^2 = 0$ this method provides the same estimators of the classical BS model.
Reserves and MSEP in the ADR model (known $\gamma = \gamma^{CL}$)

*Structural parameters.* $\hat{\tau} = 0.04961$, $\chi = 0.05755$, $\hat{\sigma} = 63.233$, $\hat{\mu}_0 = 0.88204$. 

Due to the additional uncertainty, the $\alpha_i$ in ADR are substantially smaller than in BSCR $\implies$ ADR reserves are closer to the BF-type reserves

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\alpha_i$</th>
<th>$\hat{R}^{inh}_i$</th>
<th>msep$^{1/2}(\hat{R}^{inh}_i)$</th>
<th>(%)</th>
<th>$\hat{R}^{hom}_i$</th>
<th>msep$^{1/2}(\hat{R}^{hom}_i)$</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.4405</td>
<td>0</td>
<td>0</td>
<td>.</td>
<td>0</td>
<td>0</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>0.4090</td>
<td>15,155</td>
<td>10,620</td>
<td>70.1</td>
<td>14,031</td>
<td>10,623</td>
<td>75.7</td>
</tr>
<tr>
<td>2</td>
<td>0.3952</td>
<td>26,683</td>
<td>13,743</td>
<td>51.5</td>
<td>24,757</td>
<td>13,749</td>
<td>55.5</td>
</tr>
<tr>
<td>3</td>
<td>0.3867</td>
<td>36,544</td>
<td>16,219</td>
<td>44.4</td>
<td>33,825</td>
<td>16,229</td>
<td>48.0</td>
</tr>
<tr>
<td>4</td>
<td>0.3848</td>
<td>91,926</td>
<td>26,089</td>
<td>28.4</td>
<td>85,000</td>
<td>26,132</td>
<td>30.7</td>
</tr>
<tr>
<td>5</td>
<td>0.3829</td>
<td>170,354</td>
<td>35,945</td>
<td>21.1</td>
<td>157,395</td>
<td>36,054</td>
<td>22.9</td>
</tr>
<tr>
<td>6</td>
<td>0.3769</td>
<td>320,635</td>
<td>50,801</td>
<td>15.8</td>
<td>295,551</td>
<td>51,088</td>
<td>17.3</td>
</tr>
<tr>
<td>7</td>
<td>0.3668</td>
<td>511,867</td>
<td>67,539</td>
<td>13.2</td>
<td>468,989</td>
<td>68,170</td>
<td>14.5</td>
</tr>
<tr>
<td>8</td>
<td>0.3487</td>
<td>1,208,764</td>
<td>113,536</td>
<td>9.4</td>
<td>1,107,452</td>
<td>115,623</td>
<td>10.4</td>
</tr>
<tr>
<td>9</td>
<td>0.3047</td>
<td>4,620,160</td>
<td>316,789</td>
<td>6.9</td>
<td>4,229,107</td>
<td>327,843</td>
<td>7.8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>7,002,087</td>
<td>407,426</td>
<td>5.8</td>
<td>6,416,109</td>
<td>426,609</td>
<td>6.6</td>
</tr>
<tr>
<td>Total in BSCR</td>
<td></td>
<td>6,413,077</td>
<td>326,040</td>
<td>5.1</td>
<td>6,167,268</td>
<td>329,031</td>
<td>5.3</td>
</tr>
</tbody>
</table>
ADR model with unknown development pattern
(Work in progress, jointly with Hans Bühlmann and Mario Wüthrich)

The approach we are following is essentially the same we applied to BSCR model. Assuming that an estimate \( \hat{\eta}_{inh/hom} \) has been obtained by the previous formulae, we adopt for \( \gamma_j \) the normalized estimator:

\[
\hat{\gamma}_{j, inh/hom} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} a_i \hat{\eta}_i} \left( \sum_{l=0}^{J} \frac{\sum_{i=0}^{I-l} X_{i,l}}{\sum_{i=0}^{I-l} a_i \hat{\eta}_i} \right)^{-1}, \quad j = 0, \ldots, J.
\]

Using these estimators, we can introduce the definition of compatible pair \((\tilde{\eta}, \tilde{\gamma})\) and specify an iterative estimation algorithm essentially similar to that used in the BSCR model.

Estimation error of development pattern
A parametric bootstrap simulation method is used essentially similar to that used for BSCR.
Reserves and MSEP in the ADR model including $\gamma$ estimation error

$S = 10000$, $4 \leq \text{num\_iterations} \leq 7$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\tilde{R}_{i\text{inh}}$</th>
<th>msep$<em>{\tilde{R}</em>{i\text{inh}}}^{1/2}$</th>
<th>(%)</th>
<th>$\tilde{R}_{i\text{hom}}$</th>
<th>msep$<em>{\tilde{R}</em>{i\text{hom}}}^{1/2}$</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15,315</td>
<td>15,401</td>
<td>100.6</td>
<td>14,146</td>
<td>14,744</td>
<td>104.2</td>
</tr>
<tr>
<td>2</td>
<td>27,383</td>
<td>18,740</td>
<td>68.4</td>
<td>25,354</td>
<td>18,040</td>
<td>71.2</td>
</tr>
<tr>
<td>3</td>
<td>37,840</td>
<td>21,127</td>
<td>55.8</td>
<td>34,957</td>
<td>20,437</td>
<td>58.5</td>
</tr>
<tr>
<td>4</td>
<td>95,862</td>
<td>31,706</td>
<td>33.1</td>
<td>88,478</td>
<td>30,967</td>
<td>35.0</td>
</tr>
<tr>
<td>5</td>
<td>177,582</td>
<td>42,348</td>
<td>23.8</td>
<td>163,790</td>
<td>41,612</td>
<td>25.4</td>
</tr>
<tr>
<td>6</td>
<td>333,265</td>
<td>57,968</td>
<td>17.4</td>
<td>306,705</td>
<td>57,409</td>
<td>18.7</td>
</tr>
<tr>
<td>7</td>
<td>529,862</td>
<td>75,392</td>
<td>14.2</td>
<td>484,771</td>
<td>75,235</td>
<td>15.5</td>
</tr>
<tr>
<td>8</td>
<td>1,241,745</td>
<td>123,760</td>
<td>10.0</td>
<td>1,136,370</td>
<td>125,278</td>
<td>11.0</td>
</tr>
<tr>
<td>9</td>
<td>4,701,896</td>
<td>335,611</td>
<td>7.1</td>
<td>4,301,428</td>
<td>347,447</td>
<td>8.1</td>
</tr>
<tr>
<td>Total</td>
<td>7,160,751</td>
<td>471,848</td>
<td>6.6</td>
<td>6,555,999</td>
<td>488,943</td>
<td>7.5</td>
</tr>
<tr>
<td>Total no EE</td>
<td>7,002,087</td>
<td>411,900</td>
<td>5.8</td>
<td>6,416,109</td>
<td>432,646</td>
<td>6.6</td>
</tr>
</tbody>
</table>
The convergence issue

As concerning the convergence of the estimation algorithm, however, the situation is different from the BSCR case.

- Counterexamples, i.e. examples of not convergence, can be found also for the case without re-estimation of the structural parameters.
- However this happens in rather extreme/unrealistic cases and it seems that we have convergence in most practical situations.

\[ \text{convergence fails for very large values of the ratio } \frac{\chi^2}{\sigma^2}? \]