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:

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

$\Longrightarrow$

- the random variables $X_{ij} \in \mathcal{D}_I$ are *observed*
- the random variables $X_{ij} \in \mathcal{D}_I^c$ – the outstanding claims – must be *predicted*.

$X_{0,0}$	$X_{0,1}$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$X_{0,J-1}$	$X_{0,J}$
$X_{1,0}$	$X_{1,1}$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$X_{1,J-1}$	$X_{1,J}$
$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$
$X_{I-J,0}$	$X_{I-J,1}$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$X_{I-J,J-1}$	$X_{I-J,J}$
$X_{I-J+1,0}$	$X_{I-J+1,1}$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$X_{I-J+1,J-1}$	$X_{I-J+1,J}$
$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$
$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$
$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$
$X_{i,0}$	$X_{i,1}$	$\cdots$	$X_{i,I-i}$	$X_{i,I-i+1}$	$\cdots$	$X_{i,J-1}$	$X_{i,J}$
$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$
$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$	$\vdots\vdots\vdots$	$\vdots$	$\vdots$
$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$
$X_{I,0}$	$X_{I,1}$	$\cdots$	$\cdot$	$\cdot$	$\cdots$	$X_{I,J-1}$	$X_{I,J}$

Denoting by $\widehat{X}_{i,j}$ the prediction of $X_{i,j} \in \mathcal{D}_I^c$ we then have the *reserve estimates*:

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

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

Simplified notations

For semplicity sake we assume $I = J$ hereafter ($\mathcal{D}_I$ 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: $\widehat{R}_i = \widehat{U}_i - D_i$

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

Loss Development Methods

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

$\cdot$ *loss development quota* $\gamma_{i,j}$: percentage of U_i paid in DY $j = 0, 1, \ldots, J$,

$\cdot$ *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*:

$$\boldsymbol{\gamma} := \{\gamma_j;\ 0 \leq j \leq J\} \ .$$

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:

$$U_i = D_i + D_i \frac{1-\beta_{I-i}}{\beta_{I-i}} \quad \longrightarrow \quad \textit{Projective Reserve:} \quad R_i = D_i \frac{1-\beta_{I-i}}{\beta_{I-i}}$$

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

- Also:

$$U_i = D_i + U_i\,(1-\beta_{I-i}) \quad \longrightarrow \quad \textit{Allocative Reserve:} \quad R_i = U_i\,(1-\beta_{I-i})$$

$\rightarrow (1-\beta_{I-i})$: "*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:

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

  where:

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

  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 $\widehat{f}_j := \widehat{\beta}_{j+1}/\widehat{\beta}_j$ different from CL.

Allocative (or additive) methods

- Bornhuetter-Ferguson (BF):

$$\widehat{R}_i = a_i \left(1 - \beta_{I-i}\right),$$

where a_i is a prior estimate of U_i.

BF reserving is the prototype of the allocative reserve

- Cape Cod (CC):

$$\widehat{R}_i = \widehat{\kappa}^{CC} \, P_i \left(1 - \beta_{I-i}\right),$$

where P_i are the premiums received and $\widehat{\kappa}^{CC}$ is an average loss ratio estimate (e.g. $\widehat{\kappa}^{CC} = D_{[I]} / \sum_i P_i \beta_{I-i}$).

However CC can also be considered as a projection method applied to a "robusted diagonal" $\widehat{D}_i^{CC} := \widehat{\kappa}^{CC} \, P_i \, \beta_{I-i}$.

- Additive Loss Reserving (ALR).

Under uncertainty, the fundamental question is:

What kind of information is the estimation of U_i and γ_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 $\widehat{\gamma}_j$.

- 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 $\widehat{\gamma}_j$ uses mixed information:

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

We shall denote by:

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

the normalized version of this "raw estimator".

- BF-type model by Saluz, Gisler and Wüthrich (2011).

- ***Linear* hybrid model**

$$\widehat{R}_i = \alpha_i \, D_i \, \frac{1-\beta_{I-i}}{\beta_{I-i}} + (1-\alpha_i) \, a_i \, (1-\beta_{I-i})$$

How are the weights α_i determined?

Ad hoc solution: α_i should increase with the development of $C_{i,j}$ since we obtain better information on U_i with increasing time j

$\Longrightarrow$ 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:

! The risk characteristics of accident year i are described by a *latent variable* Θ_i

! Conditionally, given Θ_i, the incremental losses $X_{i,j}$ are *independent*

$\Longrightarrow$ *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:

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

Bühlmann, H. and Moriconi, F. (2015), *Credibility Claims Reserving with Stochastic Diagonal Effects.* ASTIN Bulletin **45(2)**, 309-353.

Saluz, A., Bühlmann, H., Gisler, A. and Moriconi, F. (2014), *Bornhuetter-Ferguson Reserving Method with Repricing*, March 17. Available at SSRN: http://ssrn.com/abstract=2697167.

Model assumptions

A1. Let $\boldsymbol{\Theta} := \{\Theta_i;\ 0 \le i \le I\}$. 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 $0 \le j \le J$:

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

and:

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

A2. Let $\boldsymbol{X}_i = \{X_{i,j};\ 0 \le j \le J\}$. For $0 \le i \le I$ the pairs $\{\Theta_i, \boldsymbol{X}_i\}$ are independent. Moreover, all Θ variables are independent, with:

$$\mathbf{E}(\Theta_i) = \mu_0\,, \qquad \mathbf{Var}(\Theta_i) = \tau^2\,, \qquad 0 \le i \le I\,.$$

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

the parameters $\{a_i;\, 0 \le i \le I\}$ are the **given priors**,

the development pattern $\boldsymbol{\gamma} = \{\gamma_j;\, 0 \le j \le J\}$ is **assumed to be known**.

The parameters μ_0, τ^2, σ^2 must be estimated from the data.

The previous assumptions can be equivalently formulated in terms of *Incremental Loss Ratios* (i.e. normalized incremental losses):

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

An estimator $\widehat{\Theta}_i$ of Θ_i is an estimator of $\mathbf{E}(Z_{i,j}|\Theta_i)$.

Prediction

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

$$\widehat{X}_{ij} = a_i \, \gamma_j \, \widehat{\Theta}_i \,.$$

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

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

• 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 (*mean square error of prediction*, MSEP). The estimators we are looking for are in the class of the credibility estimators.

Credible priors

If $\widehat{\Theta}_i$ is a credibility estimator, we define $\widehat{a}_i := a_i \widehat{\Theta}_i$ the *credible prior* of the ultimate loss U_i of accident year i. One has:

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

Credibility Estimators

If the prior mean $\mu_0 = \mathbf{E}(\Theta_i)$ 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:

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

where α_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 μ_0 is not known one has:

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

$$\widehat{\Theta}_i^{hom} = \alpha_i \, \overline{Z}_i + (1 - \alpha_i) \, \widehat{\mu}_0 \,, \qquad \text{where} \qquad \widehat{\mu}_0 = \sum_{i=0}^{I} \frac{\alpha_i}{\alpha_{[I]}} \, \overline{Z}_i \,.$$

Credibility decomposition of reserve estimates

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

$$\begin{aligned} \widehat{a}_i^{hom} &:= a_i \, \widehat{\Theta}_i^{hom} \\ &= \alpha_i \, a_i \, \overline{Z}_i + (1 - \alpha_i) \; a_i \, \widehat{\mu}_0 \\ &= \alpha_i \, \frac{D_i}{\beta_{I-i}} + (1 - \alpha_i) \, a_i \, \widehat{\mu}_0 \,. \end{aligned}$$

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

$$\widehat{R}_i^{hom} := \sum_{j=I-i+1}^{J} \widehat{X}_{i,j}^{hom} = \alpha_i \; \underbrace{D_i \, \frac{(1 - \beta_{I-i})}{\beta_{I-i}}}_{Projective\ Reserve} + (1 - \alpha_i) \; \underbrace{a_i \, \widehat{\mu}_0 \, (1 - \beta_{I-i})}_{Allocative\ Reserve} \,.$$

- If $\tau^2 = 0$ one has $\alpha_i \equiv 0$ hence:

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

$$\widehat{R}_i^{inh} = a_i \, (1 - \beta_{I-i}) \, .$$

$\Longrightarrow$ In the homogeneous case one has:

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

where:

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

is the Cape Cod estimate of the overall loss ratio.

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

Numerical example

Data (Example 4.63 in Wüthrich and Merz, 2008)

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

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									

Results

Estimated development patterns

j	0	1	2	3	4	5	6	7	8	9
$\widehat{f}_j^{CL}$	1.4925	1.0778	1.02288	1.0148	1.0070	1.0052	1 .0011	1.0011	1.0014	
$\widehat{\gamma}_j^{CL}$	0.5900	0.2904	0.0684	0.0217	0.0144	0.0069	0.0051	0.0011	0.0010	0.0014
$\widehat{\gamma}_j^{RE}$	0.5860	0.2906	0.0694	0.0224	0.0151	0.0073	0.0055	0.0012	0.0011	0.0015

Credibility estimators (d.p.: $\widehat{\gamma}^{CL}$)

By the usual BS estimators for the structural parameters:

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

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

All $\widehat{\Theta}$ estimates are less then 1 $\Rightarrow$ priors are pessimistic: $\widehat{a}_i < a_i$.
Moreover, since $\widehat{\mu}_0 < \mu_0 = 1$, homogeneous reserves will be lower than inhomogeneous.

Summary of reserve estimates with different methods

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

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

$$\mathrm{msep}_R\left(\widehat{R}^{inh}\right) = \sum_{i=1}^{I} \mathrm{msep}_{R_i}\left(\widehat{R}_i^{inh}\right),$$

$$\begin{aligned}\mathrm{msep}_R\left(\widehat{R}^{hom}\right) &= \mathrm{msep}_R\left(\widehat{R}^{inh}\right) + \frac{\tau^2}{\alpha_{[I]}}\left(\sum_{i=1}^{I} \alpha_i(1-\alpha_i)\,(1-\beta_{I-i})\right)^2 \\ &= \sum_{i=1}^{I} \mathrm{msep}_{R_i}\left(\widehat{R}_i^{hom}\right) \\ &\quad + 2\,\frac{\tau^2}{\alpha_{[I]}} \sum_{1\le i<k\le I} a_i a_k\,(1-\alpha_i)(1-\alpha_k)(1-\beta_{I-i})(1-\beta_{d_k}).\end{aligned}$$

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 $\boldsymbol{\gamma} = \boldsymbol{\gamma}^{CL}$)

i	$\widehat{R}_i^{inh}$	$\text{msep}_{R_i}^{1/2}\big(\widehat{R}_i^{inh}\big)$	(%)	$\widehat{R}_i^{hom}$	$\text{msep}_{R_i}^{1/2}\big(\widehat{R}_i^{hom}\big)$	(%)
1	15,338	13,216	86.2	14,931	13,216	88.5
2	26,419	17,108	64.8	25,718	17,109	66.5
3	35,219	20,191	57.3	34,217	20,192	59.0
4	87,511	32,243	36.8	85,035	32,246	37.9
5	161,074	44,160	27.4	156,568	44,167	28.2
6	298,051	61,499	20.6	289,272	61,520	21.3
7	477,205	80,460	16.9	461,874	80,507	17.4
8	1,109,352	125,486	11.3	1,071,689	125,669	11.7
9	4,202,908	276,469	6.6	4,027,964	278,257	6.9
Total	6,413,076	326,040	5.1	6,167,268	329,031	5.3

Reserves and MSEPs in the Time Series Chain Ladder model

(Buchwalder, Bühlmann, Merz and Wüthrich, 2006)

		$\mathrm{msep}^{1/2}_{R_i}(\widehat{R}_i)$					
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

MSEP under the one-year view

For solvency purposes the MSEPs 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 MSEPs under the one-year view have been derived:

- in Bühlmann, 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 α_i^{I+1} of the credibility weight as a linear updating of the current estimate α_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 $\boldsymbol{\gamma}$ must be estimated from the data.

Estimation of the development pattern

We propose an iterative procedure. The basic idea is that if the parameters $\boldsymbol{\Theta}$ would be known, then the best linear unbiased estimator of γ_j would be:

$$\gamma_j^{raw} := \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} a_i \, \Theta_i}, \quad j = 0, \ldots, J\,.$$

Since the Θ 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:

$$\widehat{\gamma}_j^{inh/hom} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} a_i \, \widehat{\Theta}_i^{inh/hom}} \left(\sum_{l=0}^{J} \frac{\sum_{i=0}^{I-l} X_{i,l}}{\sum_{i=0}^{I-l} a_i \, \widehat{\Theta}_i^{inh/hom}} \right)^{-1}, \quad j = 0, \ldots, J\,. \qquad (1)$$

a) The estimated pattern $\widehat{\boldsymbol{\gamma}}$ will depend on $\widehat{\boldsymbol{\Theta}}$, that is $\widehat{\boldsymbol{\gamma}} = \widehat{\boldsymbol{\gamma}}(\widehat{\boldsymbol{\Theta}})$;

b) in turn, $\widehat{\boldsymbol{\Theta}}$ depends itself by $\widehat{\boldsymbol{\gamma}}$, that is $\widehat{\boldsymbol{\Theta}} = \widehat{\boldsymbol{\Theta}}(\widehat{\boldsymbol{\gamma}})$.

Definition. *We call* $(\widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{\gamma}})$ *a compatible pair if* $\widehat{\boldsymbol{\Theta}}(\widetilde{\boldsymbol{\gamma}}) = \widetilde{\boldsymbol{\Theta}}$ *and* $\widehat{\boldsymbol{\gamma}}(\widetilde{\boldsymbol{\Theta}}) = \widetilde{\boldsymbol{\gamma}}$

$\Longrightarrow$

Iterative estimation algorithm (homogeneous case):

0. *Initialization.* We choose an initial development pattern $\boldsymbol{\gamma}^{(0)}$ and pose $\widehat{\boldsymbol{\gamma}} = \boldsymbol{\gamma}^{(0)}$.
1. *Estimation of the structural parameters.* Using $\widehat{\boldsymbol{\gamma}}$, estimates of the parameters $\widehat{\mu}_0, \widehat{\tau}^2, \widehat{\sigma}^2$ are obtained by the data.
2. *Estimation of* $\boldsymbol{\Theta}$. Using $\widehat{\boldsymbol{\gamma}}$ and the parameter estimates obtained in step 1, the estimates $\widehat{\boldsymbol{\Theta}}$ are derived.
3. *Estimation of* $\boldsymbol{\gamma}$. Using $\widehat{\boldsymbol{\Theta}}$, a new d.p. estimate $\widehat{\boldsymbol{\gamma}}'$ is computed.
4. *Iteration.* If a given convergence criterion is not fulfilled, we return to step 1 assuming for $\widehat{\boldsymbol{\gamma}}$ the development pattern $\widehat{\boldsymbol{\gamma}}'$. Otherwise the iteration is stopped and we adopt the resulting estimates $\widehat{\boldsymbol{\Theta}}$ and $\widehat{\boldsymbol{\gamma}}$ as the compatible pair $(\widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{\gamma}})$.

$$\text{convercence criterion:} \qquad \max\left\{\sqrt{\sum_{j=0}^{J}\left(\widehat{\gamma}_j^{(n)} - \widehat{\gamma}_j^{(n-1)}\right)^2},\ \sqrt{\sum_{i=0}^{I}\left(\widehat{\Theta}_i^{(n)} - \widehat{\Theta}_i^{(n-1)}\right)^2}\right\} < \epsilon\,,$$

Reserve estimates and MSEPs with unknown γ

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

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

Homogeneous case. After 4 iterations:

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

i	α_i	$\widehat{R}_i^{inh}$	$\text{msep}_{R_i}^{1/2}$	(%)	$\widehat{R}_i^{hom}$	$\text{msep}_{R_i}^{1/2}$	(%)
0	0.7917	0	0	.	0	0	.
1	0.7873	15,596	13,294	85.2	15,181	13,294	87.6
2	0.7810	26,844	17,203	64.1	26,128	17,202	65.8
3	0.7753	35,797	20,306	56.7	34,775	20,306	58.4
4	0.7812	88,896	32,416	36.5	86,373	32,419	37.5
5	0.7866	163,437	44,372	27.1	158,852	44,378	27.9
6	0.7831	301,931	61,742	20.4	293,014	61,762	21.1
7	0.7747	482,521	80,699	16.7	466,986	80,746	17.3
8	0.7590	1,117,632	125,627	11.2	1,079,625	125,811	11.7
9	0.6904	4,217,905	276,202	6.5	4,042,181	278,000	6.9
Total		6,450,559	326,035	5.1	6,203,114	329,048	5.3
Total, $\widehat{\gamma}^{CL}$		6,413,076	326,040	5.1	6,167,268	329,031	5.3

Deriving the estimation error

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

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

$$\mathrm{EE}_i^{(\boldsymbol{\gamma})} := \mathbf{E}\left(\widehat{R}_i^{(\boldsymbol{\gamma})} - \widetilde{R}_i\right)^2 ,$$

· $\widehat{R}_i^{(\boldsymbol{\gamma})}$ is the reserve estimate obtained with the "true value" of γ,

· $\widetilde{R}_i$ is the reserve estimate derived by the iterative algorithm.

For the total reserve we pose:

$$\mathrm{EE}^{(\boldsymbol{\gamma})} := \left(\sum_{i=1}^{I} \mathbf{E}\left(\widehat{R}_i^{(\boldsymbol{\gamma})} - \widetilde{R}_i\right)\right)^2 .$$

In order to derive an estimate of $\mathrm{EE}_i^{(\boldsymbol{\gamma})}$ and $\mathrm{EE}^{(\boldsymbol{\gamma})}$ 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 $\widetilde{\boldsymbol{\gamma}}$ 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 structutal 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 $\boldsymbol{\gamma}$.

Let us denote by $\widetilde{\tau}^2, \widetilde{\sigma}^2, \widetilde{\mu}_0$ the parameter estimates obtained for the compatible pair $(\widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{\gamma}})$.

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

1. The random variables Θ, ε are generated as independent and normally distributed, with:

$$\Theta_i^{(s)} \sim \mathrm{N}(\widetilde{\mu}_0, \widetilde{\tau}^2)\,, \quad \varepsilon_{i,j}^{(s)} \sim \mathrm{N}(1, 0)\,, \qquad 0 \le i + j \le I\,.$$

2. A pseudo-trapezoid of incremental claims is generated as:

$$X_{i,j}^{(s)} = a_i\, \widetilde{\gamma}_j\, \Theta_i^{(s)} + \sqrt{a_i\, \widetilde{\gamma}_j}\; \widetilde{\sigma}\, \varepsilon_{i,j}^{(s)}\,, \qquad 0 \le i + j \le I\,.$$

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

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

5. The squared errors are computed:

$$\Delta_i^{(s)} = \left(\widehat{R}_i^{(s)} - \widetilde{R}_i^{(s)}\right)^2, \quad i = 1, \ldots, I, \qquad \Delta^{(s)} = \left(\sum_{i=1}^{I} \left(\widehat{R}_i^{(s)} - \widetilde{R}_i^{(s)}\right)\right)^2.$$

The estimation errors are estimated as $\widehat{\mathrm{EE}}_i^{(\boldsymbol{\gamma})} = \frac{1}{S}\sum_{s=1}^{S} \Delta_i^{(s)}$ and $\widehat{\mathrm{EE}}^{(\boldsymbol{\gamma})} = \frac{1}{S}\sum_{s=1}^{S} \Delta^{(s)}$.

MSEP in the BSCR model including γ estimation error

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

i	$\text{msep}^{1/2}_{R_i}\big(\widehat{R}_i^{inh}\big)$	(%)	$\text{msep}^{1/2}_{R_i}\big(\widehat{R}_i^{hom}\big)$	(%)
1	19,072	122.3	18,783	123.7
2	23,140	86.2	22,829	87.4
3	26,067	72.8	25,754	74.1
4	38,856	43.7	38,516	44.6
5	51,524	31.5	51,166	32.2
6	69,385	23.0	69,025	23.6
7	88,730	18.4	88,386	18.9
8	135,231	12.1	134,973	12.5
9	291,912	6.9	292,987	7.2
Total	395,536	6.1	395,910	6.4
Total no EE	326,035	5.1	329,048	5.3

The convergence issue

In all the numerical examples we considered, the iterative estimation algorithm ever converged to a unique point $(\widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{\gamma}})$, independently of the initial d.p. $\boldsymbol{\gamma}^{(0)}$.

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 $\boldsymbol{\Theta}(\boldsymbol{\gamma})$ and $\boldsymbol{\gamma}(\boldsymbol{\Theta})$ 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:

Bühlmann, H. and Moriconi, F. (2015), *Credibility Claims Reserving with Stochastic Diagonal Effects.* ASTIN Bulletin **45(2)**, 309-353.

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 $0 \le j \le J$:

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

and:

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

A2. All η, ζ variables are independent, with:

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

We interpret:

· η_i: *random effect of accident year* i,

· ζ_{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 $\boldsymbol{\gamma}$ 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 $\boldsymbol{\Theta}$ which takes the form:

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

  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:

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

  with $\mathbb{I}_{i,k} = 1$ if $i = k$ and 0 elsewhere.
  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.

! 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 $X_{ij} \in \mathcal{D}_I^c$, we now consider the predictor $\widehat{X}_{ij} = a_i \, \gamma_j \, \widehat{\eta}_i$.

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

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

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 $X_{i,j}, X_{k,l} \in \mathcal{D}_I$, we consider the covariance between the incremental loss ratios:

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

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

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

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

Under the model assumptions one has:

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

Structure of the $\boldsymbol{\Omega}$ matrix with $I = J = 2$

λ \ ν		1 (0,0)	2 (0,1)	3 (0,2)	4 (1,0)	5 (1,1)	6 (2,0)
1	(0,0)	$\tau^2+\chi^2+\frac{\sigma^2}{a_0\gamma_0}$	τ^2	τ^2	0	0	0
2	(0,1)	τ^2	$\tau^2+\chi^2+\frac{\sigma^2}{a_0\gamma_1}$	τ^2	χ^2	0	0
3	(0,2)	τ^2	τ^2	$\tau^2+\chi^2+\frac{\sigma^2}{a_0\gamma_2}$	0	χ^2	χ^2
4	(1,0)	0	χ^2	0	$\tau^2+\chi^2+\frac{\sigma^2}{a_1\gamma_0}$	τ^2	0
5	(1,1)	0	0	χ^2	τ^2	$\tau^2+\chi^2+\frac{\sigma^2}{a_1\gamma_1}$	χ^2
6	(2,0)	0	0	χ^2	0	χ^2	$\tau^2+\chi^2+\frac{\sigma^2}{a_2\gamma_0}$

- If $\chi^2 = 0$ (BSCR model) the covariance matrix $\boldsymbol{\Omega}$ is *block diagonal*, where each block corresponds to an accident year.
  If $\boldsymbol{\Omega}$ is a block diagonal matrix with blocks B_i, the inverse matrix $\boldsymbol{\Omega}^{-1}$ 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 α_i, $\overline{Z}_i$ and $\widehat{\eta}_i$) are based only on the observations of accident year i.

- If $\chi^2 > 0$ (ADR model) $\boldsymbol{\Omega}$ 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 η_i, $0 \le i \le I$, is given by:*

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

where:

$$\alpha_i := \tau^2 \sum_{k=0}^{I} \sum_{l=0}^{I-k} \sum_{j=0}^{I-i} \omega^{(-1)}_{(i,j),(k,l)} \ , \qquad \bar{\bar{Z}}_i := \frac{\tau^2}{\alpha_i} \sum_{k=0}^{I} \sum_{l=0}^{I-k} \sum_{j=0}^{I-i} \omega^{(-1)}_{(i,j),(k,l)} \, Z_{k,l} \, .$$

The $\mathbf{\Omega}^{-1}$ matrix with $I = J = 2$

λ \ ν		1 (0,0)	2 (0,1)	3 (0,2)	4 (1,0)	5 (1,1)	6 (2,0)
1	(0,0)				$\omega^{-1}_{(0,0),(1,0)}$	$\omega^{-1}_{(0,0),(1,1)}$	$\omega^{-1}_{(0,0),(2,0)}$
2	(0,1)		B_0^{-1}		$\omega^{-1}_{(0,1),(1,0)}$	$\omega^{-1}_{(0,1),(1,1)}$	$\omega^{-1}_{(0,1),(2,0)}$
3	(0,2)				$\omega^{-1}_{(0,2),(1,0)}$	$\omega^{-1}_{(0,2),(1,1)}$	$\omega^{-1}_{(0,2),(2,0)}$
4	(1,0)	$\omega^{-1}_{(1,0),(0,0)}$	$\omega^{-1}_{(1,0),(0,1)}$	$\omega^{-1}_{(1,0),(0,2)}$	B_1^{-1}		$\omega^{-1}_{(1,0),(2,0)}$
5	(1,1)	$\omega^{-1}_{(1,1),(0,0)}$	$\omega^{-1}_{(1,1),(0,1)}$	$\omega^{-1}_{(1,1),(0,2)}$			$\omega^{-1}_{(1,1),(2,0)}$
6	(2,0)	$\omega^{-1}_{(2,0),(0,0)}$	$\omega^{-1}_{(2,0),(0,1)}$	$\omega^{-1}_{(2,0),(0,2)}$	$\omega^{-1}_{(2,0),(1,0)}$	$\omega^{-1}_{(2,0),(1,1)}$	B_2^{-1}

α_i is obtained by $\mathbf{\Omega}^{-1}$ by computing the sum of all the elements on the band corresponding to accident year i (and then multiplying by τ^2).

Theorem 4.4. *The best linear homogeneous estimator for η_i, $0 \le i \le I$, is given by:*

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

Credibility decomposition of reserve estimates

$$\widehat{R}_i^{hom} := \sum_{j=I-i+1}^{J} \widehat{X}_{i,j}^{hom} = \alpha_i \, \underbrace{a_i\,\bar{\bar{Z}}_i\,(1-\beta_{I-i})}_{Projective\ Reserve} + (1-\alpha_i)\, \underbrace{a_i\,\widehat{\mu}_0\,(1-\beta_{I-i})}_{Allocative\ Reserve}$$

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 τ^2, χ^2, σ^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 $\mathbf{E}(\mathrm{SS})$ of each of these sums is expressed as a function $f(\tau^2, \chi^2, \sigma^2)$. If the expectation $\mathbf{E}(\mathrm{SS})$ is replaced by the corresponding observed value SS^*, the equations $\mathrm{SS}^* = f(\tau^2, \chi^2, \sigma^2)$ 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 model assumptions, the f functions are linear, and we are led to a system of three independent linear equations. We obtain estimates for τ^2, χ^2, σ^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 $\boldsymbol{\gamma} = \boldsymbol{\gamma}^{CL}$)

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

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

i	α_i	$\widehat{R}_i^{inh}$	$\text{msep}_{R_i}^{1/2}(\widehat{R}_i^{inh})$	(%)	$\widehat{R}_i^{hom}$	$\text{msep}_{R_i}^{1/2}(\widehat{R}_i^{hom})$	(%)
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
Total in BSCR		6,413,077	326,040	5.1	6,167,268	329,031	5.3

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 $\widehat{\eta}_i^{inh/hom}$ has been obtained by the previous formulae, we adopt for γ_j the normalized estimator:

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

Using these estimators, we can introduce the definition of compatible pair $(\widetilde{\boldsymbol{\eta}}, \widetilde{\boldsymbol{\gamma}})$ 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 γ estimation error

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

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

The convergence issue

As concerning the convergence of the estimation algorithm, however, the situation is diffent 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.
  ? convergence fails for very large values of the ratio χ^2/σ^2?