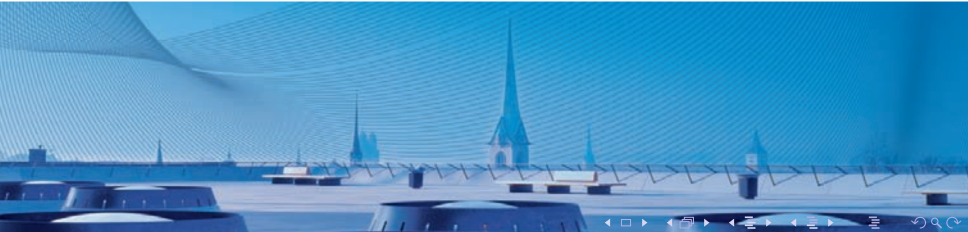
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# Development Pattern and Prediction Error for the Stochastic Bornhuetter-Ferguson Claims Reserving Method

Annina Saluz, Alois Gisler, Mario V. Wüthrich

ETH Zurich

ASTIN Colloquium Madrid, June 2011



# Overview


1. Notation
2. Bornhuetter-Ferguson Method (BF)
3. Normal Model
4. Estimation of the Parameters and their Correlations
5. Prediction Uncertainty
6. Conclusions and Remarks

# Notation

- Accident years $i$, $0 \le i \le I$

Table: Claims development triangle

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

# Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$

Table: Claims development triangle

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

# Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$

Table: Claims development triangle

| accident year $i$ | development year $j$ <br> $0 \quad \ldots \quad j \quad \ldots \quad J$ |
|---|---|
| 0 | |
| $\vdots$ | observations $\mathcal{D}_I$ |
| $i$ | |
| $\vdots$ | $\mathcal{D}_I^c$ |
| $I$ | |

# Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$

Table: Claims development triangle

| accident year $i$ | development year $j$: $0$ … $j$ … $J$ |
|---|---|
| 0 | |
| ⋮ | observations $\mathcal{D}_I$ |
| $i$ | |
| ⋮ | $\mathcal{D}_I^c$ |
| $I$ | |

# Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$
- Observed claims at time $I$: $\mathcal{D}_I = \{X_{i,j};\ i+j \leq I\}$

Table: Claims development triangle

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

# Notation

- Accident years $i$, $0 \leq i \leq I$
- Development years $j$, $0 \leq j \leq J \leq I$
- Incremental claims $X_{i,j}$
- Cumulative claims $C_{i,j}$
- Observed claims at time $I$: $\mathcal{D}_I = \{X_{i,j};\ i + j \leq I\}$
- Goal: Predict $\mathcal{D}_I^c = \{X_{i,j};\ i + j > I, i \leq I, j \leq J\}$

Table: Claims development triangle

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

# Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

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

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$

# Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

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

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
- $1 - \hat{\beta}_{I-i}$ estimated still to come factor at the end of development year $I - i$, $\hat{\beta}_J = 1$, $((\hat{\beta}_j)_{j=0,\dots,J}$ estimated cumulative development pattern)

# Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

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

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
- $1 - \hat{\beta}_{I-i}$ estimated still to come factor at the end of development year $I - i$, $\hat{\beta}_J = 1$, ($(\hat{\beta}_j)_{j=0,\ldots,J}$ estimated cumulative development pattern)
- $\hat{\gamma}_0 = \hat{\beta}_0$, $\hat{\gamma}_j = \hat{\beta}_j - \hat{\beta}_{j-1}$, $1 \leq j \leq J$, (estimated incremental development pattern)

# Bornhuetter-Ferguson Method (BF)

BF predictor for the outstanding loss liabilities (IBNR and IBNeR in case of incurred claims data) $R_i = C_{i,J} - C_{i,I-i}$ at time $I$

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

where

- $\hat{\mu}_i$ a prior estimate for $\mu_i = E[C_{i,J}]$
- $1 - \hat{\beta}_{I-i}$ estimated still to come factor at the end of development year $I - i$, $\hat{\beta}_J = 1$, $((\hat{\beta}_j)_{j=0,\ldots,J}$ estimated cumulative development pattern)
- $\hat{\gamma}_0 = \hat{\beta}_0$, $\hat{\gamma}_j = \hat{\beta}_j - \hat{\beta}_{j-1}$, $1 \leq j \leq J$, (estimated incremental development pattern)
- Assumption of a cross-classified model $E[X_{i,j}] = \mu_i \gamma_j$

# Motivation for the Estimation of the Pattern

- Often the chain ladder (CL) development pattern is used for $\hat{\gamma}_j$

$$\hat{\gamma}_j^{CL} = \prod_{k=j}^{J-1} \hat{f}_k^{-1} - \prod_{k=j-1}^{J-1} \hat{f}_k^{-1}, \quad \hat{f}_k = \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}}.$$

Issue: In the BF method we are given a priori estimates of the $\mu_i$. The CL pattern ignores any prior information.
Goal: Incorporate the a priori estimates in the estimation of the development pattern.

# Motivation for the Estimation of the Pattern

- Often the chain ladder (CL) development pattern is used for $\hat{\gamma}_j$

$$\hat{\gamma}_j^{CL} = \prod_{k=j}^{J-1} \hat{f}_k^{-1} - \prod_{k=j-1}^{J-1} \hat{f}_k^{-1}, \quad \hat{f}_k = \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}}.$$

  Issue: In the BF method we are given a priori estimates of the $\mu_i$. The CL pattern ignores any prior information.
  Goal: Incorporate the a priori estimates in the estimation of the development pattern.

- If the $\mu_i$ were known then the best linear unbiased estimate of $\gamma_j$ would be

$$\gamma_j^{(0)} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \mu_i}, \quad 0 \leq j \leq J.$$

  $\Rightarrow$ A first candidate is the raw estimate

$$\hat{\gamma}_j^{(0)} = \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i},$$

  However, they do not sum up to 1.

# Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

$$\hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \leq j \leq J.$$

# Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

$$\hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \le j \le J.$$

- If only the upper trapezoid and the $\mu_i$ are known, replace the unknown $X_{i,j}$ by predictors $\mu_i \hat{\gamma}_j$

$$\hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} X_{i,j} + \sum_{i=I-j+1}^{I} \mu_i \hat{\gamma}_j}{\sum_{i=0}^{I-J} C_{i,J} + \sum_{i=I-J+1}^{I} \left( C_{i,I-i} + \sum_{l=I-i+1}^{J} \mu_i \hat{\gamma}_l \right)}, \quad 0 \le j \le J.$$

# Intuitive Estimation of the Pattern

- If the full rectangle was known an obvious estimate is given by

$$\hat{\gamma}_j = \frac{\sum_{i=0}^{I} X_{i,j}}{\sum_{i=0}^{I} C_{i,J}}, \quad 0 \leq j \leq J.$$

- If only the upper trapezoid and the $\mu_i$ are known, replace the unknown $X_{i,j}$ by predictors $\mu_i \hat{\gamma}_j$

$$\hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} X_{i,j} + \sum_{i=I-j+1}^{I} \mu_i \hat{\gamma}_j}{\sum_{i=0}^{I-J} C_{i,J} + \sum_{i=I-J+1}^{I} \left( C_{i,I-i} + \sum_{l=I-i+1}^{J} \mu_i \hat{\gamma}_l \right)}, \quad 0 \leq j \leq J.$$

- In an over-dispersed Poisson model the maximum likelihood estimates (MLE's) satisfy these equations.

# Normal Model

**Model Assumptions (Normal Model)**

N1 The $X_{i,j}$ are independent and normally distributed and there exist parameters $\mu_0$, $\mu_1$, ..., $\mu_I$ and $\gamma_0, \gamma_1, \ldots, \gamma_J$ with $\sum_{j=0}^{J} \gamma_j = 1$ and $\sigma_0^2, \ldots, \sigma_J^2$ such that

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

where $\sigma_j^2 > 0$, $0 \leq j \leq J$.

N2 The a priori estimates $\hat{\mu}_i$ for $\mu_i = E[C_{i,J}]$ are unbiased and independent of all $X_{l,j}$.

# Maximum Likelihood Estimation (MLE) of the $\gamma_j$'s

- We calculate the MLE's assuming that the true $\mu_i$'s and $\sigma_j^2$'s are known and then replace the $\mu_i$'s by the a priori estimates $\hat{\mu}_i$ and the $\sigma_j^2$'s by estimates $\hat{\sigma}_j^2$.

# Maximum Likelihood Estimation (MLE) of the $\gamma_j$'s

- We calculate the MLE's assuming that the true $\mu_i$'s and $\sigma_j^2$'s are known and then replace the $\mu_i$'s by the a priori estimates $\hat{\mu}_i$ and the $\sigma_j^2$'s by estimates $\hat{\sigma}_j^2$.
- In the Normal Model we obtain

$$\hat{\gamma}_j = \underbrace{\frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i}}_{\text{raw estimate}} + \underbrace{\frac{\hat{\sigma}_j^2 / \sum_{i=0}^{I-j} \hat{\mu}_i}{\sum_{l=0}^{J} \left( \hat{\sigma}_l^2 / \sum_{i=0}^{I-l} \hat{\mu}_i \right)} \left( 1 - \sum_{l=0}^{J} \frac{\sum_{i=0}^{I-l} X_{i,l}}{\sum_{i=0}^{I-l} \hat{\mu}_i} \right)}_{\text{correction term}},$$

where

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j} \hat{\mu}_i \left( \frac{X_{i,j}}{\hat{\mu}_i} - \frac{\sum_{i=0}^{I-j} X_{i,j}}{\sum_{i=0}^{I-j} \hat{\mu}_i} \right)^2, \quad 0 \le j \le J, \ j \ne I.$$

# Covariance matrix of the $\hat{\gamma}_j$'s

We use the asymptotic properties of MLE's (Fisher Information matrix) to estimate the covariance matrix of the $\hat{\gamma}_j$'s.:

$$\widehat{\mathrm{Cov}}(\hat{\gamma}_j, \hat{\gamma}_k) = \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{I-j} \hat{\mu}_i} \left( \mathbb{1}_{\{j=k\}} - \frac{\left(\hat{\sigma}_k^2 / \sum_{i=0}^{I-k} \hat{\mu}_i\right)}{\sum_{l=0}^{J} \left(\hat{\sigma}_l^2 / \sum_{i=0}^{I-l} \hat{\mu}_i\right)} \right).$$

Remarks:

- For the best linear unbiased estimate $\gamma_j^{(0)}$ (in the case of known $\mu_i$) we have

$$\mathrm{Cov}(\gamma_j^{(0)}, \gamma_k^{(0)}) = \mathbb{1}_{\{j=k\}} \frac{\sigma_j^2}{\sum_{i=0}^{I-j} \mu_i},$$

  (compare with first summand).
- Because of the side constraint $\sum_{j=0}^{J} \gamma_j = 1$ the off-diagonal correlations must be negative (compare with second summand).

# Prediction Uncertainty

Given all information $\mathcal{I}_I$ (i.e. $\mathcal{D}_I$ and all $\hat{\mu}_i$), the conditional MSEP of the predictor $\widehat{R}_i = \hat{\mu}_i(1-\hat{\beta}_{I-i})$ is given by

$$\begin{aligned}
\mathrm{msep}_{R_i|\mathcal{I}_I}(\widehat{R}_i) &= E\left[\left(R_i - \widehat{R}_i\right)^2 \middle| \mathcal{I}_I\right] \\
&= E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \hat{\mu}_i(1-\hat{\beta}_{I-i})\right)^2 \middle| \mathcal{I}_I\right] \\
&= \underbrace{\sum_{j=I-i+1}^{J} \mathrm{Var}(X_{i,j})}_{\text{Process Variance } (\mathrm{PV}_i)} + \underbrace{\left(\hat{\mu}_i(1-\hat{\beta}_{I-i}) - \mu_i(1-\beta_{I-i})\right)^2}_{\text{Estimation Error } (\mathrm{EE}_i)}.
\end{aligned}$$

# MSEP

We estimate the conditional MSEP given $\mathcal{I}_I$ as follows

$$\widehat{\mathrm{msep}}_{R_i|\mathcal{I}_I}(\widehat{R}_i) = \sum_{j=I-i+1}^{J} \widehat{\mathrm{Var}}(X_{i,j}) + \widehat{\mathrm{Var}}(\hat{\mu}_i)(1-\hat{\beta}_{I-i})^2$$
$$+ \hat{\mu}_i^2 \left( \sum_{j=0}^{I-i} \widehat{\mathrm{Var}}(\hat{\gamma}_j) + 2 \sum_{0 \le j < k \le I-i} \widehat{\mathrm{Cov}}(\hat{\gamma}_j, \hat{\gamma}_k) \right),$$

$$\widehat{\mathrm{msep}}_{\sum_{i=I-J+1}^{I} R_i|\mathcal{I}_I} \left( \sum_{i=I-J+1}^{I} \widehat{R}_i \right) = \sum_{i=I-J+1}^{I} \widehat{\mathrm{msep}}_{R_i|\mathcal{I}_I} \left( \widehat{R}_i \right)$$
$$+ 2 \sum_{I-J+1 \le i < k \le I} \Bigg( (1-\hat{\beta}_{I-i})(1-\hat{\beta}_{I-k}) \widehat{\mathrm{Cov}}(\hat{\mu}_i, \hat{\mu}_k)$$
$$+ \hat{\mu}_i \hat{\mu}_k \sum_{j=0}^{I-i} \sum_{l=0}^{I-k} \widehat{\mathrm{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l) \Bigg).$$

# Conclusions and Remarks

Under distributional assumptions we have derived

- estimates for the development pattern taking all relevant information into account
- formulas for the smoothing from the raw estimates $\hat{\gamma}_j^{(0)}$ to the final estimates $\hat{\gamma}_j$
- estimates for the correlations of these estimates.

We recommend to use these formulas also in the distribution-free case (currently there are no estimators available from which we know that they perform better).