Annals of Actuarial Science, Vol. 5, part 1, pp. 7–17.
doi:10.1017/S1748499510000023
First published online 21 October 2010

Prediction Uncertainty in the Bornhuetter-Ferguson Claims Reserving Method: Revisited

By D. H. Alai, M. Merz and M. V. Wüthrich

Abstract

We revisit the stochastic model of Alai *et al.* (2009) for the Bornhuetter-Ferguson claims reserving method, Bornhuetter & Ferguson (1972). We derive an estimator of its conditional mean square error of prediction (MSEP) using an approach that is based on generalized linear models and maximum likelihood estimators for the model parameters. This approach leads to simple formulas, which can easily be implemented in a spreadsheet.

Keywords

Claims Reserving; Bornhuetter-Ferguson; Overdispersed Poisson Distribution; Chain Ladder Method; Generalized Linear Models; Fisher Information Matrix; Conditional Mean Square Error of Prediction

Contact address

M. V. Wüthrich, RiskLab, Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland.
E-mail: mario.wuethrich@math.ethz.ch

1. Introduction

The prediction uncertainty in the Bornhuetter-Ferguson (BF) claims reserving method, Bornhuetter & Ferguson (1972), has recently been studied by several authors; see e.g. Mack (2008), Verrall (2004) and Alai *et al.* (2009). We revisit the model studied in Alai *et al.* (2009). In the present paper we provide a different method of approximating the mean square error of prediction (MSEP), which substantially simplifies the formulas while preserving the accuracy established in the previous paper.

Alai *et al.* (2009) maintain that in practice the chain ladder (CL) development pattern is used for calculating the BF reserves, and hence incorporate this into their model assumptions. This is done by assuming the data to be overdispersed Poisson distributed. This allows one to recreate the CL estimate of the development pattern; a result dating back to Hachemeister & Stanard (1975) and Mack (1991). This is different from the approach taken in Mack (2008), but closer to the implementation of practitioners. We furthermore maintain the necessary assumption that the initial estimates of the expected ultimate claims are independent of the data, an assumption that is the basis of the BF methodology. This independence assumption is certainly challenged in practice, both Mack (2008) and Schmidt & Zocher (2008) suggest that estimates of the expected ultimate claims come from pricing, and it remains unclear how independent such information is from the data. We do not dwell on this issue at present and leave it for future consideration.

In this paper our main objective is to provide a far simpler method of implementing the results derived in Alai *et al.* (2009). We direct the reader to the previous paper, as well as the works of Mack (2008), Neuhaus (1992) and Schmidt & Zocher (2008) for elaboration on the motivation of studying the BF method and when it should be applied in practice. Furthermore, we do not consider the claims inflation problem at present, a topic with recent developments made by Kuang *et al.* (2008a,b).

Organization of the paper. In Section 2 we provide the notation and data structure as well as the model considerations. In Section 3 we give a short review of the BF method. In Section 4 we give a simplified estimation procedure for the conditional MSEP in the BF method. Finally, in Section 5 we revisit the case study presented in Alai *et al.* (2009) and compare our results with Mack (2008) and Verrall (2004).

2. Data and Model

2.1. Setup

Let $X_{i,j}$ denote the incremental claims of accident year $i \in \{0, 1, \dots, I\}$ and development year $j \in \{0, 1, \dots, J\}$. We assume the data is given by a claims development triangle, i.e. $I = J$, and that after J development periods all claims are settled. At time I, we have observations $\mathcal{D}_I = \{X_{i,j},\ i + j \leq I\}$. We are interested in predicting the corresponding lower triangle $\{X_{i,j},\ i + j > I,\ i \leq I\}$. Furthermore, define $C_{i,j}$ to be the cumulative claims of accident year i up to development year j. Hence,

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

2.2. Model Considerations

We adopt the overdispersed Poisson model presented in Alai *et al.* (2009). Please refer to Section 4.1 of that paper for the density function, from which it is shown that the overdispersed Poisson belongs to the exponential dispersion family. The reader is also advised to see Kuang *et al.* (2009) for further elaboration on the connection between the overdispersed Poisson model and the CL method using maximum likelihood estimators (MLEs), as well as Section 2.3 below.

Model Assumptions 2.1 (Overdispersed Poisson Model)

- The incremental claims $X_{i,j}$ are independent overdispersed Poisson distributed and there exist positive parameters $\gamma_0, \dots, \gamma_I$, $\mu_0, \dots, \mu_I$ and $\phi > 0$ with

$$E[X_{i,j}] = m_{i,j} = \mu_i \gamma_j,$$
$$\mathrm{Var}(X_{i,j}) = \phi\, m_{i,j},$$

  and $\sum_{j=0}^{I} \gamma_j = 1$.
- $\widehat{\nu}_k$ are random variables that are unbiased estimators of the expected ultimate claim $\mu_k = E[C_{k,I}]$ for all $k \in \{0, \dots, I\}$.
- $X_{i,j}$ and $\widehat{\nu}_k$ are independent for all i, j, k.

Remarks 2.2:

- The exogenous estimator $\widehat{\nu}_k$ is a prior estimate of the expected ultimate claims $E[C_{k,I}]$, which is used for the BF method; see also Section 2 in Mack (2008). In the work of Alai *et al.* (2009) these estimators were assumed to be independent, we now adopt a more general assumption in which their dependence structure can be modelled.

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

- For MSEP considerations, an estimate of the uncertainty of the $\widehat{\nu}_k$ is required. Below, we assume that a prior variance estimate $\widehat{\mathrm{Var}}(\widehat{\nu}_k)$ is given exogenously.
- For additional model interpretations we refer to Alai *et al.* (2009).

2.3. Maximum Likelihood Estimators

Under Model Assumptions 2.1 the log-likelihood function for $\mathcal{D}_I$ is given by

$$l_{\mathcal{D}_I}(\mu_i, \gamma_j, \phi) = \sum_{\substack{i+j\leq I \\ j<I}} \left(\frac{1}{\phi}(X_{i,j}\log(\mu_i\gamma_j) - \mu_i\gamma_j) + \log c(X_{i,j};\phi) \right)$$
$$+ \left(\frac{1}{\phi}\left(X_{0,I}\log\left[\mu_0\left(1 - \sum_{n=0}^{I-1}\gamma_n \right) \right] - \mu_0\left(1 - \sum_{n=0}^{I-1}\gamma_n \right) + \log c(X_{0,I};\phi) \right) \right),$$

where $c(\cdot,\phi)$ is the suitable normalizing function. Notice that the substitution, $\gamma_I = (1 - \sum_{n=0}^{I-1}\gamma_n)$ has been made in accordance with the constraint provided in Model Assumptions 2.1. The MLEs $\widehat{\mu}_i, \widehat{\gamma}_j$ are found by taking the derivates with respect to μ_i, γ_j and setting the resulting equations equal to zero. They are given by,

$$\begin{aligned}
\widehat{\mu}_0 &= \sum_{j=0}^{I} X_{0,j}, \\
\widehat{\mu}_i \sum_{j=0}^{I-i}\widehat{\gamma}_j &= \sum_{j=0}^{I-i} X_{i,j}, \qquad i \in \{1,\ldots,I\}, \\
\widehat{\gamma}_j \left(\sum_{i=1}^{I-j}\widehat{\mu}_i + X_{0,I}\frac{1}{1-\sum_{n=0}^{I-1}\widehat{\gamma}_n} \right) &= \sum_{i=0}^{I-j} X_{i,j}, \qquad j \in \{0,\ldots,I-1\}.
\end{aligned} \tag{1}$$

Furthermore, we define $\widehat{\gamma}_I = 1 - \sum_{n=0}^{I-1}\widehat{\gamma}_n$. The $\widehat{\mu}_i, \widehat{\gamma}_j$ can also be calculated with help from the well-known CL factors,

$$\widehat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}};$$

see e.g. Corollary 2.18 and Remarks 2.19 in Wüthrich & Merz (2008), i.e.

$$\widehat{\gamma}_j = \prod_{k=j}^{I-1}\frac{1}{\widehat{f}_k}\left(1 - \frac{1}{\widehat{f}_{j-1}} \right), \qquad \widehat{\mu}_i = C_{i,I-i}\widehat{f}_{I-i}\cdots\widehat{f}_{I-1}. \tag{2}$$

Although, as is clear in (1), ϕ has no influence on the parameter estimation of μ_i, γ_j, an estimate of ϕ is required to estimate the prediction uncertainty. As done in Alai *et al.* (2009), we use Pearson residuals to estimate ϕ:

$$\widehat{\phi} = \frac{1}{d}\sum_{i+j\leq I}\frac{(X_{i,j} - \widehat{m}_{i,j})^2}{\widehat{m}_{i,j}}, \tag{3}$$

where $d = \frac{(I+1)(I+2)}{2} - 2I - 1$ is the degrees of freedom of the model and $\widehat{m}_{i,j} = \widehat{\mu}_i\widehat{\gamma}_j$.

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

2.4. Asymptotic Properties of the MLE

As previously stated, the overdispersed Poisson model is a member of the exponential dispersion family. We use the proposition directly below which yields the asymptotic behaviour of the MLEs to quantify the parameter estimation uncertainty $\widehat{\gamma}_j - \gamma_j$; see e.g. Lehmann (1983), Theorem 6.2.3.

Proposition 2.3 *Assume $X_1, \ldots, X_n$ are i.i.d. with density $f_{\zeta}(\cdot)$ from the exponential dispersion family with parameters $\boldsymbol{\zeta}=(\zeta_1, \ldots, \zeta_m)^T$. Furthermore, $\widehat{\boldsymbol{\zeta}} = (\widehat{\zeta}_1, \ldots, \widehat{\zeta}_m)^T$ is the MLE of $\boldsymbol{\zeta}$, then,*

$$\sqrt{n}(\widehat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}) \xrightarrow{(d)} \mathcal{N}(0, H(\boldsymbol{\zeta})^{-1}), \quad \text{as } n \to \infty,$$

where we define the Fisher information matrix by $H(\boldsymbol{\zeta}) = (h_{r,s})_{r,s=1,\ldots,m}$ with

$$h_{r,s} = H(\boldsymbol{\zeta})_{r,s} = -E_{\boldsymbol{\zeta}}\left[\frac{\partial^2}{\partial\zeta_r\partial\zeta_s}\log f_{\boldsymbol{\zeta}}(X)\right].$$

Remark 2.4 One must be careful when considering asymptotic behaviour with respect to studying claims reserving triangles. We inherently limit ourselves to a finite dataset and hence introduce some error when using results from asymptotic theory. This issue was studied numerically, using the bootstrap method, in Section 7.3 of Wüthrich & Merz (2008). There it was observed, using the same dataset we study in Section 5, that the bias and the estimation error can be estimated accurately under the asymptotic normal approximation.

We use the notation $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_{2I+1}) = (\mu_0, \ldots, \mu_I, \gamma_0, \ldots, \gamma_{I-1})$ and $\widehat{\boldsymbol{\zeta}}$ for the corresponding MLE. Under Model Assumptions 2.1, we obtain for the components of the Fisher information matrix:

$$h_{i+1,i+1} = \frac{\mu_i^{-1}}{\phi}\sum_{j=0}^{I-i}\gamma_j, \qquad i \in \{0,\ldots,I\},$$

$$h_{I+2+j,I+2+j} = \frac{\gamma_j^{-1}}{\phi}\sum_{i=0}^{I-j}\mu_i + \frac{\mu_0}{\phi\left(1-\sum_{n=0}^{I-1}\gamma_n\right)}, \qquad j \in \{0,\ldots,I-1\},$$

$$h_{I+2+j,I+2+l} = \frac{\mu_0}{\phi\left(1-\sum_{n=0}^{I-1}\gamma_n\right)}, \qquad j,l \in \{0,\ldots,I-1\},\ j \neq l,$$

$$h_{i+1,I+2+j} = \frac{1}{\phi}, \qquad i \in \{1,\ldots,I\},\ j \in \{0,\ldots,I-i\},$$

$$h_{I+2+j,i+1} = \frac{1}{\phi}, \qquad j \in \{0,\ldots,I-1\},\ i \in \{1,\ldots,I-j\}.$$

The remaining entries of the $(2I+1) \times (2I+1)$ matrix $H(\boldsymbol{\zeta})$ are zero. By replacing the parameters $\boldsymbol{\zeta}$ and ϕ by their estimates given in (1) and (3), respectively, we obtain the estimated Fisher information matrix $H(\widehat{\boldsymbol{\zeta}}, \widehat{\phi})$. The inverse of the estimated Fisher information matrix, $H(\widehat{\boldsymbol{\zeta}}, \widehat{\phi})^{-1}$, contains, for our purposes, unnecessary information regarding the parameters μ_i. Therefore, we define the $(I+1) \times (I \times 1)$ matrix

$$\mathcal{G} = (g_{j,l})_{j,l=0,\ldots,I},$$

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

with

$$g_{j,l} = \widehat{\mathrm{Cov}}(\widehat{\gamma}_j, \widehat{\gamma}_l) = H(\widehat{\zeta}, \widehat{\phi})^{-1}_{I+2+j,I+2+l}, \qquad j,l \in \{0,\ldots,I-1\},$$

$$g_{j,I} = g_{I,j} = \widehat{\mathrm{Cov}}(\widehat{\gamma}_j, \widehat{\gamma}_I) = -\sum_{m=0}^{I-1} H(\widehat{\zeta}, \widehat{\phi})^{-1}_{I+2+j,I+2+m}, \quad j \in \{0,\ldots,I-1\}, \tag{4}$$

$$g_{I,I} = \widehat{\mathrm{Var}}(\widehat{\gamma}_I) = \sum_{\substack{0\le m\le I-1\\ 0\le n\le I-1}} H(\widehat{\zeta}, \widehat{\phi})^{-1}_{I+2+m,I+2+n}.$$

The first equation of (4) gives an estimator for the covariances between the MLEs $\widehat{\gamma}_j$ and $\widehat{\gamma}_l$, whereas the last two equations of (4) incorporate the MLE $\widehat{\gamma}_I = 1-\sum_{n=0}^{I-1}\widehat{\gamma}_n$.

3. The Bornhuetter-Ferguson Method

In practice, the BF predictor, which dates back to Bornhuetter & Ferguson (1972), relies on the data for the development pattern γ_j and on external data or expert opinion for the expected ultimate claims $E[C_{i,I}]$. The ultimate claim $C_{i,I}$ of accident year i under Model Assumptions 2.1 using the BF method, given $\mathcal{D}_I$, is predicted by

$$\widehat{C}_{i,I}^{BF} = C_{i,I-i} + \widehat{\nu}_i \sum_{j>I-i} \widehat{\gamma}_j, \tag{5}$$

where $\widehat{\gamma}_j$ are the MLEs produced in Section 2.3 and $\widehat{\nu}_i$ is an exogenous prior estimator for the expected ultimate claim $E[C_{i,I}]$ introduced in Model Assumptions 2.1.

Note that we define the BF predictor with the CL development pattern $\widehat{\gamma}_j$, which is the approach used in practice; see equation (2). A different approach for the estimation of the development pattern γ_j is given in Mack (2008), we further discuss this in the case study in Section 5.

4. The MSEP of the Bornhuetter-Ferguson Method

We begin by considering the (conditional) MSEP of the BF predictor $\widehat{C}_{i,I}^{BF}$ for single accident years $i\in\{1, \ldots, I\}$. From (5.5) in Alai *et al.* (2009) we have

$$\begin{aligned}\mathrm{msep}_{C_{i,I}|\mathcal{D}_I}(\widehat{C}_{i,I}^{BF}) &= E\left[(\widehat{C}_{i,I}^{BF} - C_{i,I})^2 | \mathcal{D}_I\right] \\ &= \sum_{j>I-i} \mathrm{Var}(X_{i,j}) + \left(\sum_{j>I-i}\widehat{\gamma}_j\right)^2 \mathrm{Var}(\widehat{\nu}_i) + \mu_i^2\left(\sum_{j>I-i}\widehat{\gamma}_j - \sum_{j>I-i}\gamma_j\right)^2.\end{aligned} \tag{6}$$

The first term on the right-hand side of equation (6) is the (conditional) process variance, it represents the stochastic movement of the $X_{i,j}$, the inherent uncertainty from our model assumptions.
The latter two terms form the (conditional) estimation error; these terms constitute the uncertainty in the prediction of the prior estimate $\widehat{\nu}_i$ and the MLEs $\widehat{\gamma}_j$. The first two terms on the right-hand side of equation (6) can be estimated by replacing unknowns with their estimates; see e.g. Sections 5.1.1 and 5.1.2 in Alai *et al.* (2009). The last term, however, if tackled this way would equal zero.
The standard approach, see England & Verrall (2002), is to estimate

$$\left(\sum_{j>I-i}(\widehat{\gamma}_j - \gamma_j)\right)^2$$

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

by the unconditional expectation

$$E\left[\left(\sum_{j>I-i}(\widehat{\gamma}_j-\gamma_j)\right)^2\right]=\sum_{\substack{j>I-i\\l>I-i}}E[(\widehat{\gamma}_j-\gamma_j)(\widehat{\gamma}_l-\gamma_l)].$$

Neglecting that MLEs have a possible bias term, see Remark 2.4, we make the following approximation:

$$\sum_{\substack{j>I-i\\l>I-i}}E[(\widehat{\gamma}_j-\gamma_j)(\widehat{\gamma}_l-\gamma_l)]\approx\sum_{\substack{j>I-i\\l>I-i}}\mathrm{Cov}(\widehat{\gamma}_j-\widehat{\gamma}_l).$$

We now deviate from Alai *et al.* (2009) and directly use $\mathcal{G}$, given by equations (4) to estimate the covariance terms. Hence, an estimate of the MSEP in the BF method for single accident year i is given by:

Estimator 4.1 (MSEP for the BF method, single accident year) *Under Model Assumptions 2.1 an estimator for the (conditional) MSEP for a single accident year $i\in\{1,\ldots,I\}$ is given by*

$$\widehat{\mathrm{msep}}_{C_{i,I}|\mathcal{D}_I}(\widehat{C}_{i,I}^{BF})=\sum_{j>I-i}\widehat{\phi}\widehat{\nu}_i\widehat{\gamma}_j+\left(\sum_{j>I-i}\widehat{\gamma}_j\right)^2\widehat{\mathrm{Var}}(\widehat{\nu}_i)+\widehat{\nu}_i^2\sum_{\substack{j>I-i\\l>I-i}}g_{j,l}.$$

Remark 4.2 If we compare the above estimator to equation (5.30) in Alai *et al.* (2009) we observe that the first two terms on the right-hand side are identical. However, the last term, i.e. the uncertainty in $\widehat{\gamma}_j$, has substantially simplified and can be easily calculated in a spreadsheet environment.

For multiple accident years the (conditional) MSEP is defined as follows:

$$\begin{aligned}\mathrm{msep}_{\sum_{i=1}^{I}C_{i,I}|\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{C}_{i,I}^{BF}\right)&=E\left[\left(\sum_{i=1}^{I}\widehat{C}_{i,I}^{BF}-\sum_{i=1}^{I}C_{i,I}\right)^2\middle|\mathcal{D}_I\right]\\&=\sum_{i=1}^{I}\mathrm{msep}_{C_{i,I}|\mathcal{D}_I}(\widehat{C}_{i,I}^{BF})+2\sum_{i<k}\mu_i\mu_k\sum_{\substack{j>I-i\\l>I-k}}(\widehat{\gamma}_j-\gamma_j)(\widehat{\gamma}_l-\gamma_l)\\&\quad+2\sum_{i<k}\left(\sum_{j>I-i}\widehat{\gamma}_j\right)\left(\sum_{l>I-k}\widehat{\gamma}_l\right)\mathrm{Cov}(\widehat{\nu}_i,\widehat{\nu}_k).\end{aligned}$$

Remark 4.3 Here we obtain an additional term compared to Alai *et al.* (2009) since we allow the exogenous estimators $\widehat{\nu}_i$ to depend on one another, which is a natural assumption if one estimates $\widehat{\nu}_i$ from the time series $\widehat{\nu}_l, l\leq i$. One can now make use of a variety of methods to estimate the covariances, methods that imply decaying positive correlation are recommended; see e.g. Mack (2008).

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

Table 1. Observed incremental claims $X_{i,j}$.

$i\backslash j$	0	1	2	3	4	5	6	7	8	9
0	5,946,975	3,721,237	895,717	207,760	206,704	62,124	65,813	14,850	11,130	15,813
1	6,346,756	3,246,406	723,222	151,797	67,824	36,603	52,752	11,186	11,646	
2	6,269,090	2,976,233	847,053	262,768	152,703	65,444	53,545	8,924		
3	5,863,015	2,683,224	722,532	190,653	132,976	88,340	43,329			
4	5,778,885	2,745,229	653,894	273,395	230,288	105,224				
5	6,184,793	2,828,338	572,765	244,899	104,957					
6	5,600,184	2,893,207	563,114	225,517						
7	5,288,066	2,440,103	528,043							
8	5,290,793	2,357,936								
9	5,675,568									

Estimator 4.4 (MSEP for the BF method, aggregated accident years) *Under Model Assumptions 2.1 an estimator for the (conditional) MSEP for aggregated accident years is given by*

$$\widehat{\mathrm{msep}}_{\sum_{i=1}^{I} C_{i,I}|\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{C}_{i,I}^{BF}\right) = \sum_{i=1}^{I}\widehat{\mathrm{msep}}_{C_{i,I}|\mathcal{D}_I}(\widehat{C}_{i,I}^{BF}) + 2\sum_{i<k}\widehat{\nu}_i\widehat{\nu}_k\sum_{\substack{j>I-i\\ l>I-k}} g_{j,l}$$
$$+2\sum_{i<k}\left(\sum_{j>I-i}\widehat{\gamma}_j\right)\left(\sum_{l>I-k}\widehat{\gamma}_l\right)\widehat{\mathrm{Cov}}(\widehat{\nu}_i,\widehat{\nu}_k).$$

Remark 4.5 The above estimator should be compared with equation (5.35) in Alai *et al.* (2009). Again, the additional term present due to our more general assumptions on the exogenous estimators of the expected ultimate claims. On the other hand, the second term on the right-hand side of Estimator 4.4 has a much simpler form compared with Alai *et al.* (2009).

5. Case Study

We utilize the dataset $\{X_{i,j} : i + j \leq I\}$ provided in Alai *et al.* (2009), which is shown in Table 1. We assume given external estimates $\widehat{\nu}_i$ of the ultimate claims, presented in Table 2. Furthermore, we assume the uncertainty of these estimates to be given by a coefficient of variation of 5% and assume

Table 2. Prior estimates for the expected ultimate claims.

i	$\widehat{\nu}_i$
1	11,364,606
2	10,962,965
3	10,616,762
4	11,044,881
5	11,480,700
6	11,413,572
7	11,126,527
8	10,986,548
9	11,618,437

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

Table 3. Reserve and uncertainty results for single and aggregated accident years using the method in Alai *et al.* (2009).

accident year i	BF reserves	Process std. dev.	Prior std. dev.	Parameter std. dev.	Prior and parameter std. dev.	msep$^{1/2}$	Vco
1	16,120	15,401	806	15,539	15,560	21,893	135.8%
2	26,998	19,931	1,350	17,573	17,624	26,606	98.5%
3	37,575	23,514	1,879	18,545	18,639	30,005	79.9%
4	95,434	37,473	4,772	24,168	24,635	44,845	47.0%
5	178,023	51,181	8,901	29,600	30,910	59,790	33.6%
6	341,305	70,866	17,065	35,750	39,614	81,187	23.8%
7	574,089	91,909	28,704	41,221	50,231	104,739	18.2%
8	1,318,645	139,294	65,932	53,175	84,703	163,025	12.4%
9	4,768,385	264,882	238,419	75,853	250,195	364,362	7.6%
covariance				195,409	195,409	195,409	
total	7,356,575	329,007	249,828	228,249	338,396	471,971	6.4%

they are uncorrelated. Hence,

$$\widehat{\mathrm{Var}}(\widehat{v}_i) = \widehat{v}_i^2 (0.05)^2.$$

Using equation (3), we obtain for the dispersion parameter ϕ, the estimate $\widehat{\phi} = 14{,}714$.

We demonstrate the numerical results in Table 3. Note that they are the same as the results presented in Alai *et al.* (2009), but the implementation is much simpler now. We compare the results in Table 3 to those from Mack (2008). We start by calculating the development pattern using equation (3) in Mack (2008). We normalize these results such that the pattern sums to one. Note that the normalization is necessary due to the fact that the prior estimates $\widehat{v}_i$ are rather conservative (as mentioned in Wüthrich & Merz (2008), Example 2.11).

In Table 4 we compare the cumulative development pattern (referred to as $\widehat{z}_j^*$ in Mack (2008)) with the cumulative development pattern obtained using the method of Alai *et al.* (2009)

Table 4. Cumulative development pattern, a comparison.

j	Alai *et al.* (2009)		Mack (2008)		Simulation	
	$\widehat{\beta}_j$	s.e. $(\widehat{\beta}_j)$	$\widehat{z}_j^*$	s.e. $(\widehat{z}_j^*)$	$\widehat{\beta}_j$	s.e. $(\widehat{\beta}_j)$
0	58.96%	0.653%	58.60%	1.717%	58.96%	0.654%
1	88.00%	0.484%	87.66%	0.616%	88.00%	0.486%
2	94.84%	0.370%	94.60%	0.326%	94.84%	0.373%
3	97.01%	0.313%	96.84%	0.271%	97.01%	0.317%
4	98.45%	0.258%	98.35%	0.131%	98.45%	0.260%
5	99.14%	0.219%	99.07%	0.054%	99.14%	0.220%
6	99.65%	0.175%	99.62%	0.025%	99.65%	0.177%
7	99.75%	0.160%	99.73%	0.018%	99.75%	0.162%
8	99.86%	0.137%	99.85%	0.012%	99.86%	0.138%
9	100.00%		100.00%		100.00%	

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

Table 5. $\widehat{s}_j^2$ calculated from equation (3) in Mack (2008).

j	0	1	2	3	4	5	6	7	8	9
$\widehat{s}_j^2$	69,990	25,848	2,450	260	376	79	11	.918	.407	.180

(referred to as $\widehat{\beta}_j$). Also shown in Table 4 are the standard errors calculated for the cumulative development patterns using the respective methods. In Alai *et al.* (2009) the approximated standard errors of the development pattern were compared with empirical standard errors obtained from simulation. The simulation study consisted of 10,000 simulated run-off triangles under the assumption of overdispersed Poisson data. The results from this simulation study shows the accuracy of the approximation, which is maintained here. Therefore, although a different, much simpler, method of approximation is utilized in this paper, it preserves the accuracy established in the previous paper.

Remark 5.1 The distinction is made between estimates of the development pattern $\widehat{\gamma}_j$ and of the cumulative development pattern $\widehat{\beta}_j$; the latter being defined as follows:

$$\widehat{\beta}_j = \sum_{k=0}^{j} \widehat{\gamma}_k, \qquad \text{for } j \in \{0, \ldots, I\}.$$

Table 4 indicates a slower decrease of the uncertainty in our approach.

In Table 5 we provide the $\widehat{s}_j^2$ calculated using equation (4) in Mack (2008). The role of the $\widehat{s}_j^2$ are comparable to that of $\widehat{\phi}$. The difference originates from the fact that the $\widehat{s}_j^2$ depend on the development year j, whereas $\widehat{\phi}$ does not.

Finally, we apply the same coefficient of variation to determine the standard error of the ultimates using the method in Mack (2008), namely 5%. Table 6 provides the MSEP results under the method described in Mack (2008). It should be compared to Table 3, which provides the results under the method described in Alai *et al.* (2009) and in this paper.

Table 6. Reserve and uncertainty results for single and aggregated accident years using the method in Mack (2008).

accident year i	BF reserves	process std. dev.	prior std. dev.	parameter std. dev.	prior and parameter std. dev.	msep$^{1/2}$	Vco
1	17,420	1,431	871	1,415	1,661	2,193	12.6%
2	29,059	2,536	1,453	1,998	2,470	3,540	12.2%
3	40,480	3,996	2,024	2,607	3,300	5,183	12.8%
4	102,383	11,541	5,119	6,025	7,906	13,989	13.7%
5	189,802	32,332	9,490	15,060	17,801	36,908	19.4%
6	360,691	73,010	18,035	30,914	35,790	81,310	22.5%
7	600,764	89,940	30,038	36,319	47,131	101,541	16.9%
8	1,355,361	186,841	67,768	67,773	95,842	209,988	15.5%
9	4,809,547	580,712	240,477	199,723	312,600	659,504	13.7%
covariance				173,602	173,602	173,602	
total	7,505,506	621,899	252,532	277,796	375,424	726,431	9.7%

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

Table 7. Aggregate reserve and uncertainty results for the CL method, the BF approach of Alai *et al.* (2009), the BF approach of Mack (2008), and the BF approach of Verrall (2004).

	reserves	process error	estimation error	$msep^{1/2}$	Vco
CL method	6,047,061	424,379	185,026	462,960	7.7%
BF Alai *et al.* (2009)	7,356,575	329,007	338,396	471,971	6.4%
BF Mack (2008)	7,505,506	621,899	375,424	726,431	9.7%
BF Verrall (2004)	7,356,560	-	-	427,278	5.8%

As becomes clear from comparing Tables 3 and 6, one main difference between the two methods lies in the estimated process variance. It is evident that this difference originates in the model assumptions with respect to the structure of the variance of the incremental claims. Alai *et al.* (2009) assume

$$\mathrm{Var}(X_{i,j}) = \phi\, m_{i,j},$$

whereas Mack (2008) assumes

$$\mathrm{Var}(X_{i,j}) = s_j^2\, m_{i,j}.$$

Table 5 shows the volatility of the s_j^2, which heavily impacts the process variance. A similar picture is obtained for the parameter standard deviation, in contrast to the prior standard deviation, which almost perfectly coincide.

Finally, in Table 7, we present the MSEP results for the distribution-free CL method, see Mack (1993), as well as the MSEP results from Verrall (2004). To obtain the (conditional) MSEP in the distribution-free CL method, we use the approach described in Buchwalder *et al.* (2006). The calculations for the Bayesian negative binomial approach presented in Verrall (2004) were performed using WinBugs, we ran 20,000 iterations, discarding the first 10,000. Although in no way conclusive, the overall approach of Alai *et al.* (2009) is more in line with the CL MSEP figures.

References

Alai, D. H., Merz, M. & Wüthrich, M. V. (2009). Mean square error of prediction in the Bornhuetter-Ferguson claims reserving method. *Annals of Actuarial Science*, **4**, 7–31.

Bornhuetter, R. L. & Ferguson, R. E. (1972). The actuary and IBNR. *Proceedings CAS*, **LIX**, 181–195.

Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. V. (2006). The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin*, **36**(2), 543–571.

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

Hachemeister, C. & Stanard, J. (1975). IBNR claims count estimation with static lag functions. Presented at ASTIN Colloquium, Portimao, Portugal.

Kuang, D., Nielsen, B. & Nielsen, J. P. (2008a). Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 987–991.

https://doi.org/10.1017/S1748499510000023 Published online by Cambridge University Press

Kuang, D., Nielsen, B. & Nielsen, J. P. (2008b). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 979–986.

Kuang, D., Nielsen, B. & Nielsen, J. P. (2009). Chain-Ladder as maximum likelihood revisited. *Annals of Actuarial Science*, **4**, 105–121.

Lehmann, E. L. (1983). *Theory of Point Estimation*. Wiley & Sons, New York.

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

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

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

Neuhaus, W. (1992). Another pragmatic loss reserving method or Bornhuetter/Ferguson revisited. *Scandinavian Actuarial Journal*, **2**, 151–162.

Schmidt, K. D. & Zocher, M. (2008). The Bornhuetter-Ferguson principle. *Variance*, **2**(1), 85–110.

Verrall, R. J. (2004). A Bayesian generalized linear model for the Bornhuetter-Ferguson method of claims reserving. *North American Actuarial Journal*, **8**(3), 67–89.

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