A.A.S. **4**, I, 7-31 (2009)

MEAN SQUARE ERROR OF PREDICTION IN THE BORNHUETTER–FERGUSON CLAIMS RESERVING METHOD

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

ABSTRACT

The prediction of adequate claims reserves is a major subject in actuarial practice and science. Due to their simplicity, the chain ladder (CL) and Bornhuetter–Ferguson (BF) methods are the most commonly used claims reserving methods in practice. However, in contrast to the CL method, no estimator for the conditional mean square error of prediction (MSEP) of the ultimate claim has been derived in the BF method until now, and as such, this paper aims to fill that gap. This will be done in the framework of generalized linear models (GLM) using the (overdispersed) Poisson model motivation for the use of CL factor estimates in the estimation of the claims development pattern.

KEYWORDS

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

CONTACT ADDRESS

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

1. Introduction

Often in non-life insurance, claims reserves are the largest item on the liability side of the balance sheet. Therefore, given the available information about the past, the prediction of an adequate amount to face the responsibilities assumed by the non-life insurance company as well as the quantification of the uncertainties in these reserves are major issues in actuarial practice and science; see e.g. Casualty Actuarial Society (1990) and Teugels & Sundt (2004).

Due to their simplicity, the chain ladder (CL) and Bornhuetter–Ferguson (BF) methods are among the easiest claims reserving methods and, therefore, the most commonly used techniques in practice. Mack (1993) published a fundamental article on claims reserving regarding the estimation of the conditional mean square error of prediction (MSEP) in the CL method. Unfortunately, until now, no estimator for the conditional MSEP of the ultimate claim has been derived in the BF method.

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

The BF method goes back to Bornhuetter & Ferguson (1972). Apart from its simplicity, the BF method is a very popular claims reserving method since it is rather robust against outliers in the observations and allows for incorporating prior knowledge from experts, premium calculations or strategic business plans. Furthermore, in contrast to the CL method, the BF method has proven to be a very robust method, in particular, against instability in the proportion of ultimate claims paid in early development years. The BF method, as it was stated in the original work of Bornhuetter & Ferguson (1972), was not formulated in a probabilistic way. The work of Mack (2000) and Verrall (2004) puts the BF method into a probabilistic framework; we discuss this further below. Before describing the method in detail we would like to mention that there are also rather sceptic and critical opinions of the use of the BF method. The purpose of this paper is not to improve the method in view of this criticism but rather to explain from a probabilistic point of view, what is done when actuaries use the BF method in its current state and to derive analytic estimators for the prediction uncertainty.

The BF method is based on the simple idea of stabilizing the BF estimate $\widehat{C_{i,J}}^{BF}$ using an initial estimate $\widehat{\mu}_i$ of the ultimate claim $C_{i,J}$ based on external knowledge. Then it is standard practice to use the prior estimate $\widehat{\mu}_i$ with the CL factor estimates $\widehat{f}_j$ to predict the ultimate claim. In this case, the CL method and the BF method only differ in the choice of the estimate for the ultimate claim (CL estimate versus prior estimate). Hence, in this regard the BF method is a variant of the CL method that uses external information to obtain an initial estimate for the ultimate claim. Mack (2000) studied this from a probabilistic point of view. In his work he analysed the stochastic model for given (deterministic) claims development patterns, which are the analogon to the CL factors, and random initial estimates $\widehat{\mu}_i$. Mack (2000) then derived optimal credibility weighted averages between the CL and the BF method. However, in most practical applications the claims development pattern and the CL factors are unknown and need to be estimated from the data. This adds an additional source of uncertainty to the problem. Verrall (2004) has studied these uncertainties using a Bayesian approach to the BF method. If one uses an appropriate Bayesian approach with improper priors and an appropriate two stage procedure, one then arrives at the BF method. We use a similar procedure within generalized linear models (GLM) using maximum likelihood estimators (MLE). This framework allows for an analytic estimate for the mean square error of prediction using asymptotic properties of MLE. Note that in a Bayesian framework one can, in general, only give numerical answers using simulation techniques such as the Markov chain Monte Carlo method.

A criticism of the BF method as it is currently used is that the use of the CL estimates $\widehat{f}_j$ contradicts the basic idea of independence between the last observed cumulative claims $C_{i,I-i}$ and the estimated outstanding claims liabilities $\widehat{C_{i,J}}^{BF} - C_{i,I-i}$, which was fundamental to the BF method; see e.g.

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Mack (2006). Therefore, Mack (2006) proposed different estimators for the claims development pattern. In this paper however, we do not follow this route. We rather use the well-known fact that the (overdispersed) Poisson model leads to the same claims reserves and payout pattern as the CL model. This means that we use the (overdispersed) Poisson model motivation for the use of the CL factor estimates $\widehat{f}_j$. It is then straightforward to use GLM methods for parameter estimation and to derive an estimator for the conditional MSEP of the ultimate claim in the BF method.

Organization of the paper. In Section 2 we provide the notation and data structure. In Section 3 we give a short review of the CL and BF methods and compare these two techniques. Section 4 is dedicated to the overdispersed Poisson model and its representation as a GLM. In Section 5 we give an estimation procedure for the conditional MSEP in the BF method. Finally, in Section 6 we discuss an example.

2. Notation and Data Structure

Throughout, we assume the loss data for the run-off portfolio is given by a claims development triangle of observations. However, all claims reserving methods discussed in this paper can also be applied to other shapes of loss data (e.g. claims development trapezoids). In this claims development triangle the indices $i \in \{0, 1, \ldots, I\}$ and $j \in \{0, 1, \ldots, J\}$ with $I \geq J$ refer to accident years and development years, respectively. The incremental claims (i.e. incremental payments, change of reported claim amount or number of newly reported claims) for accident year i and development year j are denoted by $X_{i,j}$ and cumulative claims (i.e. cumulative payments, claims incurred or total number of reported claims) of accident year i up to development year j are given by

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k}. \tag{2.1}$$

We assume that the last development year is given by J, i.e. $X_{i,j} \equiv 0$ for all $j > J$, and the last accident year is given by I. Moreover, our assumption that we consider claims development triangles implies $I = J$.

Usually, at time I (i.e. calendar year I), we have observations $\mathcal{D}_I$ in the upper claims development triangle, defined as follows,

$$\mathcal{D}_I = \{X_{i,j};\, i + j \leq I\}. \tag{2.2}$$

We need to predict the random variables in its complement

$$\mathcal{D}_I^c = \{X_{i,j};\, i + j > I, i \leq I\}. \tag{2.3}$$

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

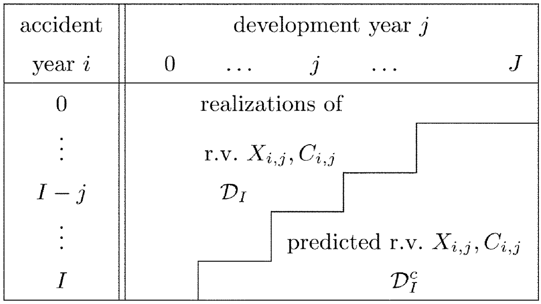

Figure 1. Claims development triangle

Figure 1 shows the claims data structure for the claims development triangle described above.

Furthermore, let R_i and R denote the outstanding claims liabilities for accident year i at time I,

$$R_i = \sum_{j=I-i+1}^{J} X_{i,j} = C_{i,J} - C_{i,I-i} \qquad \text{for } 1 \leq i \leq I, \tag{2.4}$$

and the total outstanding claims liabilities for aggregated accident years,

$$R = \sum_{i=1}^{I} R_i, \tag{2.5}$$

respectively. The prediction of the outstanding claims liabilities R_i and R by the so-called claims reserves or best estimates, as well as quantifying the uncertainty in this prediction, is the classical actuarial claims reserving problem studied at every non-life insurance company.

3. Bornhuetter–Ferguson and Chain Ladder Methods

In this section we give a short review of the CL and BF methods, which are the most commonly used claims reserving methods in practice on account of their simplicity. Our review is similar to the one given in Mack (2000).

3.1 *Chain Ladder Method*

The classical actuarial literature often explains the CL method as a pure computational algorithm to estimate claims reserves. The first distribution-free stochastic model was proposed by Mack (1993).

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Model Assumptions 3.1 (CL Model):

— There exist deterministic development factors $f_0, \ldots, f_{J-1} > 0$ such that for all $0 \leq i \leq I$ and all $1 \leq j \leq J$ we have

$$E[C_{i,j} \mid C_{i,0}, \ldots, C_{i,j-1}] = E[C_{i,j} \mid C_{i,j-1}] = f_{j-1} C_{i,j-1}. \tag{3.1}$$

— Claims $C_{i,j}$ of different accident years i are independent.

An easy exercise in calculating conditional expectation leads to

$$E[C_{i,J} \mid \mathcal{D}_I] = f_{J-1} E[C_{i,J-1} \mid C_{i,I-i}] = \ldots = C_{i,I-i} \prod_{j=I-i}^{J-1} f_j, \tag{3.2}$$

for all $1 \leq i \leq I$, where the factors f_j are called CL factors or development factors. Given the observations $\mathcal{D}_I$ and CL factors f_j, (3.2) gives a recursive algorithm for predicting the ultimate claim $C_{i,J}$. However, in most practical applications the CL factors f_j are not known and have to be estimated from the data $\mathcal{D}_I$. It is well known that the $\mathcal{D}_I$-measurable estimators for the CL factors f_j, defined by

$$\widehat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}, \tag{3.3}$$

for all $0 \leq j \leq J-1$, are unbiased and uncorrelated; see e.g. Mack (1993). However, they are not independent since the squares of two successive estimators $\widehat{f}_j$ and $\widehat{f}_{j+1}$ are negatively correlated; see e.g. Mack (2006) and Wüthrich *et al*. (2008).

The properties of the CL factor estimates $\widehat{f}_j$ imply that, given $C_{i,I-i}$, the CL estimator of the ultimate claim $C_{i,J}$, defined by

$$\widehat{C_{i,J}}^{CL} = C_{i,I-i} \prod_{j=I-i}^{J-1} \widehat{f}_j \qquad \text{for all } 1 \leq i \leq I, \tag{3.4}$$

is an unbiased estimator for $E[C_{i,J} \mid \mathcal{D}_I]$.

3.2 *Bornhuetter–Ferguson Method*

The BF method goes back to Bornhuetter & Ferguson (1972). Analogously to the CL method, the classical actuarial literature often explains the BF method as a pure computational algorithm to estimate claims reserves although there are several stochastic models that motivate the BF method.

The following stochastic model is consistent with the BF method.

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Model Assumptions 3.2 (BF Model):

— There exist parameters $\mu_0, \ldots, \mu_I > 0$ and a pattern $\beta_0, \ldots, \beta_J > 0$ with $\beta_J = 1$ such that for all $0 \le i \le I$, $0 \le j \le J-1$ and $1 \le k \le J-j$

$$E[C_{i,0}] = \beta_0 \mu_i, \tag{3.5}$$

$$E[C_{i,j+k} \mid C_{i,0}, \ldots, C_{i,j}] = C_{i,j} + \left(\beta_{j+k} - \beta_j\right)\mu_i. \tag{3.6}$$

— Claims $C_{i,j}$ of different accident years i are independent.

These assumptions imply

$$E[C_{i,j}] = \beta_j \mu_i \quad \text{and} \quad E[C_{i,J}] = \mu_i, \tag{3.7}$$

which is often used to explain the BF method; see e.g. Radtke & Schmidt (2004). The sequence $\left(\beta_j\right)_j$ denotes the claims development pattern and, if $C_{i,j}$ are cumulative payments, β_j is the expected cumulative cashflow pattern (also called payout pattern). Such a pattern is often used when one needs to build market-consistent/discounted reserves, where money values differ over time.

Assumption (3.6) motivates the BF estimator for the ultimate claim $C_{i,J}$ given by

$$\widehat{C_{i,J}}^{BF} = C_{i,I-i} + \left(1 - \widehat{\beta}_{I-i}\right)\widehat{\mu}_i, \tag{3.8}$$

for all $1 \le i \le I$, where $\widehat{\beta}_{I-i}$ is an appropriate estimate for β_{I-i} and $\widehat{\mu}_i$ is a prior estimate for the expected ultimate claim $E[C_{i,J}]$. In practice, $\widehat{\mu}_i$ is an exogenously determined estimate (i.e. without the observations $\mathcal{D}_I$) such as a plan value from a strategic business plan or the value used for premium calculations.

3.3 *Comparison of BF and CL Methods*

From the CL Assumptions 3.1 we obtain

$$E[C_{i,J}] = E[C_{i,j}] \prod_{k=j}^{J-1} f_k, \tag{3.9}$$

which implies that

$$E[C_{i,j}] = \prod_{k=j}^{J-1} f_k^{-1} E[C_{i,J}], \qquad \text{for all } 0 \le j \le J. \tag{3.10}$$

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

If we compare this to the BF method, see e.g. (3.7), we find that $\prod_{k=j}^{J-1} f_k^{-1}$ plays the role of β_j. Therefore, these parameters are often viewed equally and if one knows the CL factors f_k one can construct a development pattern $(\beta_j)_j$ and vice versa. That is, in practice the β_j are usually estimated by

$$\widehat{\beta}_j^{(CL)} = \widehat{\beta}_j = \prod_{k=j}^{J-1} \frac{1}{\widehat{f}_k}, \tag{3.11}$$

where $\widehat{f}_k$ are the CL factor estimates given in (3.3). Moreover, using the estimator $\widehat{\beta}_j^{(CL)}$ for β_j in the BF method, we see that

$$\widehat{C_{i,J}}^{BF} = C_{i,I-i} + \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i, \tag{3.12}$$

$$\widehat{C_{i,J}}^{CL} = C_{i,I-i} + \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{C_{i,J}}^{CL}, \tag{3.13}$$

which means that the CL method and BF method only differ in the choice of the estimator for the ultimate claim $C_{i,J}$ (prior estimate $\widehat{\mu}_i$ versus CL estimate $\widehat{C_{i,J}}^{CL}$). In other words, if we identify $\widehat{\beta}_j^{(CL)}$ and $\prod_{k=j}^{J-1} \widehat{f}_k^{-1}$, the BF method is a variant of the CL method that uses external information to obtain an initial estimate for the ultimate claim. The main criticism of this approach is that the use of the CL factor estimates $\widehat{f}_k$ contradicts the basic idea of independence between last observed cumulative claims $C_{i,I-i}$ and estimated outstanding claims liabilities $\widehat{C_{i,J}}^{BF} - C_{i,I-i}$, which was fundamental to the origin of the BF method; see e.g. Mack (2006). Therefore, Mack (2006) constructed different estimators for the claims development pattern $(\beta_j)_j$. However, we do not follow this route here. We rather concentrate on the overdispersed Poisson model motivation for the use of the CL factor estimates $\widehat{f}_k$ and utilize the fact that the overdispersed Poisson model is a GLM. It is then straightforward to use GLM methods for parameter estimation and to derive an estimator for the conditional MSEP of the ultimate claim in the BF method.

4. Overdispersed Poisson Model and Generalized Linear Models

In this section we give a brief review of the overdispersed Poisson model and its formulation in a GLM context.

4.1 *Overdispersed Poisson Model*

We define the overdispersed Poisson model by first considering the exponential dispersion family. Random variable Y belongs to the exponential dispersion family if its density or probability distribution function can be

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

written as

$$f_Y(y;\theta,\phi) = \exp\{[y\theta - b(\theta)]/a(\phi) + c(y,\phi)\}. \tag{4.1}$$

The overdispersed Poisson model is a member of this family with $a(\phi) = \phi$, $b(\theta) = e^\theta$ and $c(y,\phi) = -\ln y$. It differs from the Poisson model in that the variance is not equal to the mean. This model was introduced for claims reserving in a Bayesian context by Verrall (1990, 2000, 2004) and Renshaw & Verrall (1998) and it is also used in the GLM framework; see e.g. McCullagh & Nelder (1989) and England & Verrall (2002, 2007). It is well-known in actuarial literature that the (overdispersed) Poisson model leads to the same claims reserves as the CL model. This result goes back to Hachemeister & Stanard (1975) and can be found, for example, in Mack (1991) and Verrall & England (2000). This means that although the CL model and the overdispersed Poisson model are very different, they lead to the same reserve estimates, the difference in the two models is relevant only if we estimate higher moments. In the following, we will utilize this correspondence as we do not motivate the use of the estimate $\widehat{\beta}_{I-i}^{(CL)}$ by CL factor estimates $\widehat{f}_k$ but rather by the MLEs in the overdispersed Poisson model. Note that CL factor estimates are used to calculate the MLEs.

Model Assumptions 4.1 (Overdispersed Poisson Model):

— The increments $X_{i,j}$ are independent overdispersed Poisson distributed and there exist positive parameters $\gamma_0, \ldots, \gamma_J$, $\mu_0, \ldots, \mu_I$ and $\phi > 0$ such that

$$E[X_{i,j}] = m_{i,j} = \mu_i \gamma_j, \tag{4.2}$$

$$\mathrm{Var}(X_{i,j}) = \phi m_{i,j}, \tag{4.3}$$

with $\sum_{j=0}^{J} \gamma_j = 1$.

— $\widehat{\mu}_i$ are independent random variables that are unbiased estimators of $\mu_i = E[C_{i,J}]$ for all i.

— $X_{i,j}$ and $\widehat{\mu}_k$ are independent for all i, j, k.

From Model Assumptions 4.1 we obtain

$$E[C_{i,j+k} \mid C_{i,0}, \ldots, C_{i,j}] = C_{i,j} + \sum_{l=1}^{k} E[X_{i,j+l}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i, \tag{4.4}$$

where $\beta_j = \sum_{k=0}^{j} \gamma_k$. This means that the overdispersed Poisson model satisfies Model Assumptions 3.2 and can also be used to explain the BF method.

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Remarks 4.2:

— The so-called dispersion parameter ϕ does not depend on accident year i and development year j. The restriction in the overdispersed Poisson model is that we require $X_{i,j}$ to be non-negative.

— The parameters γ_k define an expected incremental reporting/cashflow pattern over the development years j.

— The exogenous estimator $\widehat{\mu}_k$ is a prior estimate for the expected ultimate claim $E[C_{k,J}]$, it is solely based on external data and expert opinion. Therefore, we assume that it is independent of the data $X_{i,j}$ (this is in the BF spirit as explained by Mack (2006)). Moreover, in order to obtain a meaningful model, we assume that it is unbiased for the expected ultimate claim. In this sense, we follow a pure BF method.

There are different methods for estimating the parameters μ_i and γ_j. In the following we use MLEs. The MLEs $\widehat{\mu}_i^{(MLE)}$ and $\widehat{\gamma}_j^{(MLE)}$ in the overdispersed Poisson model are found by solving

$$\widehat{\mu}_i^{(MLE)} \sum_{j=0}^{I-i} \widehat{\gamma}_j^{(MLE)} = \sum_{j=0}^{I-i} X_{i,j}, \tag{4.5}$$

$$\widehat{\gamma}_j^{(MLE)} \sum_{i=0}^{I-j} \widehat{\mu}_i^{(MLE)} = \sum_{i=0}^{I-j} X_{i,j}, \tag{4.6}$$

for all $0 \leq i \leq I$ and $0 \leq j \leq J$ under the constraint that $\sum_{j=0}^{J} \widehat{\gamma}_j^{(MLE)} = 1$.

Remarks 4.3:

— Because of the multiplicative structure of the overdispersed Poisson model, see e.g. (4.2), the parameters μ_i and γ_j can only be determined up to a constant factor, i.e. $\widetilde{\mu}_i = c\mu_i$ and $\widetilde{\gamma}_j = \gamma_j / c$ would give the same estimate for $m_{i,j}$. Therefore, we need to impose a side constraint. In our situation this becomes that the MLEs $\widehat{\gamma}_j^{(MLE)}$ form a development pattern, i.e. that $\sum_{j=0}^{J} \widehat{\gamma}_j^{(MLE)} = 1$.

— If we now use these MLEs $\widehat{\gamma}_j^{(MLE)}$ for the estimation of the expected incremental cashflow pattern γ_j we obtain the following BF estimator,

$$\widehat{\widehat{C}_{i,J}}^{BF} = C_{i,I-i} + \left(1 - \sum_{j=0}^{I-i} \widehat{\gamma}_j^{(MLE)}\right) \widehat{\mu}_i; \tag{4.7}$$

see e.g. (3.8). Moreover, under Model Assumptions 4.1, it has been proved that

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

$$\sum_{k=0}^{j} \widehat{\gamma}_k^{(MLE)} = \widehat{\beta}_j^{(CL)} = \prod_{k=j}^{J-1} \frac{1}{\widehat{f}_k}; \tag{4.8}$$

see e.g. Mack (1991) or Taylor (2000), which implies that the BF estimator,

$$\widehat{\widehat{C_{i,J}}}^{BF} = C_{i,I-i} + \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i, \tag{4.9}$$

is perfectly motivated by the overdispersed Poisson model and the use of the MLE $\widehat{\gamma}_j^{(MLE)}$ for γ_j. This is exactly the BF estimator as it is commonly used in practice; see e.g. (3.12). Henceforth, in this understanding we do not motivate the use of the estimate $\widehat{\beta}_{I-i}^{(CL)}$ by CL factor estimates $\widehat{f}_k$ but rather by the MLEs $\widehat{\gamma}_k^{(MLE)}$ in the overdispersed Poisson model. This means that (4.8) provides the essential steps for the use of $\widehat{\beta}_j^{(CL)}$ in the BF method.

— Note that we use a similar two stage procedure as described by Verrall (2004). First we estimate the claims development pattern γ_j and the exposures μ_i using MLE methods resulting in $\widehat{\gamma}_j^{(MLE)}$ and $\widehat{\mu}_i^{(MLE)}$. Only the simultaneous MLE of γ_j and μ_i give the CL pattern given in (4.8). In the second step we then replace the MLE $\widehat{\mu}_i^{(MLE)}$ by an external estimate $\widehat{\mu}_i$. It could be argued that this is, in some sense, inconsistent, but it describes what practitioners do in the BF method to smoothen the claims reserves estimates.

4.2 *Overdispersed Poisson Model as a Generalized Linear Model*

Renshaw (1995) and Renshaw & Verrall (1998) were the first to implement the standard GLM techniques for the derivation of estimates for incremental data in a claims reserving context. In this subsection we give a brief description of the overdispersed Poisson model in the GLM framework. For more details on the overdispersed Poisson model in the GLM context and on GLMs and their statistical background we refer to England & Verrall (2002) and McCullagh & Nelder (1989) or Fahrmeir & Tutz (2001), respectively.

A specific GLM model (in a parametrization suitable for claims reserving) is fully characterized by the following three components

(a) the type of the Exponential Dispersion Family for the *random component* $X_{i,j}$;

(b) the *link function* g relating the expectation of the random component $X_{i,j}$ to the linear predictor $\eta_{i,j} = \Gamma_{i,j}\boldsymbol{b}$, i.e.

$$g\left(E[X_{i,j}]\right) = \eta_{i,j}$$

for all $0 \le i \le I$ and $0 \le j \le J$;

(c) the *design matrices* $\Gamma_{i,j}$ for all $0 \le i \le I$ and $0 \le j \le J$.

In the overdispersed Poisson model the distribution of the random component is given by the overdispersed Poisson distribution, and for the multiplicative structure of Model 4.1 it is straightforward to choose the log-link $g(\cdot) = \log(\cdot)$ as link function. Then we have

$$\eta_{i,j} = g\big(m_{i,j}\big) = \log(\mu_i) + \log(\gamma_j). \tag{4.10}$$

For GLMs, it is easy to obtain MLEs of the parameters μ_i and γ_j using standard GLM software. However, since the multiplicative structure of Model 4.1 is overparametrized it becomes necessary to set constraints which could take a number of different forms. In the last subsection we derived the MLEs $\widehat{\gamma}_j^{(MLE)}$ under the normalization assumption $\sum_{j=0}^{J} \widehat{\gamma}_j^{(MLE)} = 1$ in order to obtain a claims development pattern; see e.g. Remarks 4.3. However, in the framework of GLMs it is more natural and indeed convenient to choose the constraint $\mu_0 = 1$, hence $\log(\mu_0) = 0$ and

$$\eta_{0,j} = \log(\gamma_j) \qquad \text{for all } 0 \le j \le J; \tag{4.11}$$

see e.g. (4.10). This parametrization leads to the following vector of unknown parameters

$$\boldsymbol{b} = \big(\log(\mu_1), \ldots, \log(\mu_I), \log(\gamma_0), \ldots, \log(\gamma_J)\big)', \tag{4.12}$$

and the $1 \times (I + J + 1)$ design matrices

$$\Gamma_{0,j} = (0, \ldots, 0, 0, 0, \ldots, 0, e_{I+j+1}, 0, \ldots, 0), \tag{4.13}$$

$$\Gamma_{i,j} = (0, \ldots, 0, e_i, 0, \ldots, 0, e_{I+j+1}, 0, \ldots, 0), \tag{4.14}$$

for $1 \le i \le I$ and $0 \le j \le J$, where the entries $e_i = 1$ and $e_{I+j+1} = 1$ are on the i-th and the $(I + j + 1)$-th position, respectively. We obtain the linear predictor

$$\eta_{i,j} = \Gamma_{i,j}\boldsymbol{b}. \tag{4.15}$$

Hence, we have now reduced the dimension from $(I + 1) \times (J + 1)$ unknown parameters $m_{i,j}$ to $p = I + J + 1$ unknown parameters $\log(\mu_i)$ and $\log(\gamma_j)$.

Using standard GLM software based on the Fisher scoring method, these parameters are then estimated with the MLE method. We obtain the MLEs

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

$$\widehat{\boldsymbol{b}} = \left(\widehat{\log(\mu_1)}^{GLM}, \ldots, \widehat{\log(\mu_I)}^{GLM}, \widehat{\log(\gamma_0)}^{GLM}, \ldots, \widehat{\log(\gamma_J)}^{GLM}\right)', \tag{4.16}$$

which implies a second "payout" pattern

$$\widehat{\gamma}_0^{(GLM)}, \ldots, \widehat{\gamma}_J^{(GLM)}, \tag{4.17}$$

where

$$\widehat{\gamma}_j^{(GLM)} = \exp\left(\widehat{\log \gamma_j}^{GLM}\right) \tag{4.18}$$

for all $0 \leq j \leq J$. The following relationships hold,

$$\widehat{\gamma}_j^{(MLE)} = \frac{\widehat{\gamma}_j^{(GLM)}}{\sum_{l=0}^{J} \widehat{\gamma}_l^{(GLM)}} \quad \text{for all } 0 \leq j \leq J. \tag{4.19}$$

Remarks 4.4:

— Note the superscripts MLE and GLM, which are used to differentiate between the two normalizations, one natural to maximum likelihood for claims reserving and the other more practical for GLM modelling purposes.

— In multiplicative models like the overdispersed Poisson model it is natural to use the log-link $g(\cdot) = \log(\cdot)$ as the link function since the systematic effects are additive on the scale given by the log-link function. Moreover, the log-link is the so-called canonical link function for the (overdispersed) Poisson distribution that has convenient mathematical and statistical properties; see e.g. McCullagh & Nelder (1989) or Fahrmeir & Tutz (2001).

— In the next section, relationship (4.19) will be crucial to incorporate our results from GLM theory in the derivation of an estimate of the conditional MSEP in the BF method.

— From GLM theory it is well-known that the MLE (4.16) is asymptotically multivariate normally distributed with covariance matrix $\text{Cov}\left(\widehat{\boldsymbol{b}}, \widehat{\boldsymbol{b}}\right)$ estimated by the inverse of the Fisher information matrix (denoted by $H^{-1}\left(\widehat{\boldsymbol{b}}\right)$), which is a standard output in all GLM software packages; see e.g. Panjer (2006) or Fahrmeir & Tutz (2001).

5. MSEP in the BF Method using GLM

In this section we quantify the uncertainty in the estimation of the ultimate claims $C_{i,J}$ and $\sum_{i=1}^{I} C_{i,J}$ by the estimators $\widehat{\widehat{C_{i,J}}}^{BF}$ and $\sum_{i=1}^{I} \widehat{\widehat{C_{i,J}}}^{BF}$,

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

respectively, given the observations $\mathcal{D}_I$. More precisely, our goal is to derive an estimate of the conditional MSEP for single accident years $1 \le i \le I$,

$$\mathrm{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF}\right) = E\left[\left(C_{i,J} - \widehat{C_{i,J}}^{BF}\right)^2 \middle| \mathcal{D}_I\right] \tag{5.1}$$

$$= E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i\right)^2 \middle| \mathcal{D}_I\right], \tag{5.2}$$

as well as an estimate of the conditional MSEP for aggregated accident years

$$\mathrm{msep}_{\sum_{i=1}^{I} C_{i,J}|\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{C_{i,J}}^{BF}\right) = E\left[\left(\sum_{i=1}^{I} C_{i,J} - \sum_{i=1}^{I}\widehat{C_{i,J}}^{BF}\right)^2 \middle| \mathcal{D}_I\right]. \tag{5.3}$$

This is described in the next subsections.

5.1 *MSEP in the BF Method, Single Accident Year*

We choose $1 \le i \le I$. Since the incremental claims $X_{i,j}$ are independent the conditional MSEP (5.1) can be decoupled in the following way:

$$\begin{aligned} &E\left[\left(\sum_{j=I-i+1}^{J} X_{i,j} - \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i\right)^2 \middle| \mathcal{D}_I\right] \\ &= \sum_{j=I-i+1}^{J} \mathrm{Var}(X_{i,j}) + E\left[\left(\sum_{j=I-i+1}^{J} E[X_{i,j}] - \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i\right)^2 \middle| \mathcal{D}_I\right] \\ &\quad + 2E\left[\left(\sum_{j=I-i+1}^{J} (X_{i,j} - E[X_{i,j}])\right)\left(\sum_{j=I-i+1}^{J} E[X_{i,j}] - \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i\right) \middle| \mathcal{D}_I\right]. \end{aligned} \tag{5.4}$$

Note that $\widehat{\mu}_i$ is independent of $X_{k,j}$ for all k, j, that $\widehat{\beta}_{I-i}^{(CL)}$ is $\mathcal{D}_I$-measurable and that $E\left[\widehat{\mu}_i\right] = \mu_i$; see e.g. (3.11) and Model Assumptions 4.1. Therefore, the last term in the above equality disappears and we get

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

$$\mathrm{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF}\right) = \sum_{j=I-i+1}^{J} \mathrm{Var}\left(X_{i,j}\right) + \left(1-\widehat{\beta}_{I-i}^{(CL)}\right)^2 \mathrm{Var}(\widehat{\mu}_i)$$
$$+ \mu_i^2 \left(\sum_{j=I-i+1}^{J} \gamma_j - \sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)} \right)^2. \tag{5.5}$$

Hence, the three terms on the right-hand side of (5.5) need to be estimated in order to get an appropriate estimate for the conditional MSEP in the BF method. The first term as a (conditional) process variance originates from the stochastic movement of $X_{i,j}$. The second and third term on the right-hand side of (5.5) constitute the (conditional) estimation error which reflects the uncertainty in the prior estimate $\widehat{\mu}_i$ and the MLEs $\widehat{\gamma}_j^{(MLE)}$, respectively.

5.1.1 *Process variance*

For the estimation of the (conditional) process variance, Model Assumptions 4.1 motivates the following estimator:

$$\widehat{\mathrm{Var}}(X_{i,j}) = \widehat{\phi}\,\widehat{\mu}_i \sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}$$
$$= \widehat{\phi}\,\widehat{\mu}_i\left(1-\widehat{\beta}_{I-i}^{(CL)}\right), \tag{5.6}$$

where $\widehat{\phi}$ is an estimate of the dispersion parameter ϕ. Within the framework of GLM we use different types of residuals (Pearson, deviance, Anscombe, etc.) to estimate ϕ; see e.g. McCullagh & Nelder (1989) or Fahrmeir & Tutz (2001). In the following we will use the Pearson residuals defined by

$$\widehat{R}_{i,j}^{(p)} = \frac{X_{i,j} - \widehat{m}_{i,j}}{\sqrt{\widehat{m}_{i,j}}}, \tag{5.7}$$

where $\widehat{m}_{i,j}$ is the GLM estimate of $m_{i,j}$ given by

$$\widehat{m}_{i,j} = \widehat{\mu}_i^{(GLM)} \widehat{\gamma}_j^{(GLM)} \tag{5.8}$$

for all $0 \le i+j \le I$. The estimate of the dispersion parameter is then given by

$$\widehat{\phi} = \frac{\sum_{0\le i+j\le I}\left(\widehat{R}_{i,j}^{(p)}\right)^2}{N-p}, \tag{5.9}$$

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

where

$$N = \text{number of observations } X_{i,j} \text{ in } \mathcal{D}_I, \text{ i.e. } N = |\mathcal{D}_I| \tag{5.10}$$

$$p = \text{number of estimated parameters, i.e. } p = I + J + 1. \tag{5.11}$$

5.1.2 *Estimation error*

The (conditional) estimation error is given by the second and third term on the right-hand side of (5.5). This means that we need to quantify the volatility of the prior estimates $\widehat{\mu}_i$ and the MLEs $\widehat{\gamma}_j^{(MLE)}$ around the true parameters μ_i and γ_j, respectively.

Prior estimate $\widehat{\mu}_i$: The second term

$$\left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)^2 \mathrm{Var}(\widehat{\mu}_i) \tag{5.12}$$

quantifies the uncertainty in the prior estimate $\widehat{\mu}_i$ of the expected ultimate claim $E[C_{i,J}]$. Since $\widehat{\mu}_i$ is determined exogenously, see e.g. Remarks 4.2, this can generally only be done using external data like market statistics and expert opinion. The regulator, for example, provides an estimate for the coefficient of variation of $\widehat{\mu}_i$, denoted by $\widehat{\mathrm{Vco}}(\widehat{\mu}_i)$, that quantifies how good the exogenous estimator $\widehat{\mu}_i$ is. Statistical estimates based on impact studies for the determination of estimates $\widehat{\mathrm{Vco}}(\widehat{\mu}_i)$ exist, for example, in the context of the Swiss Solvency Test (2006). These studies suggest that 5% to 10% is a reasonable range for $\widehat{\mathrm{Vco}}(\widehat{\mu}_i)$. Hence the term (5.12) is estimated by

$$\left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)^2 \widehat{\mathrm{Var}}(\widehat{\mu}_i) = \left(1 - \widehat{\beta}_{I-i}^{(CL)}\right)^2 \widehat{\mu}_i^2 \widehat{\mathrm{Vco}}(\widehat{\mu}_i)^2. \tag{5.13}$$

Note that an appropriate choice for $\widehat{\mathrm{Vco}}(\widehat{\mu}_i)$ is crucial for a meaningful analysis. This choice is closely related to a Bayesian setup where one chooses an appropriate prior distribution for $\widehat{\mu}_i$; see e.g. Mack (2000). Of course, the choice of this prior distribution and/or its coefficient of variation depends on the internal processes of the company. Ideally, this is determined using market statistics as described above (and similarly as used, for example, in the context of modelling operational risk, see Lambrigger *et al.* (2007)). Unfortunately, in many cases there are no market statistics available and one tries to adjust the priors using internal data. However, this approach contradicts the BF method if we interpret it in the strict sense described by Mack (2006), since the choice of the prior $\widehat{\mu}_i$ should be independent from the observations $X_{k,j}$. We would like to motivate further research into this direction, i.e. (1) finding appropriate priors and (2) describe the internal

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

processes as they are used in practice. This could lead to a new theory and method using Kalman filters, see e.g. Chapter 9 in Bühlmann & Gisler (2005) and Chapter 10 in Taylor (2000), to describe loss ratio prediction based on observations of past accident years. One then immediately loses the independence assumptions and the conditional MSEP no longer decouples in a nice way.

MLEs $\widehat{\gamma}_j^{(MLE)}$: The estimation of the third term on the right-hand side of (5.5) requires more work. We have to study the fluctuations of the MLEs $\widehat{\gamma}_j^{(MLE)}$ around the true parameters γ_j

$$\left(\sum_{j=I-i+1}^{J} \gamma_j - \sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}\right)^2. \tag{5.14}$$

Neglecting that MLEs have a possible bias term we estimate (5.14) by

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}\right) &= \sum_{j,k=I-i+1}^{J} \mathrm{Cov}\left(\widehat{\gamma}_j^{(MLE)}, \widehat{\gamma}_k^{(MLE)}\right) \\
&= \sum_{j,k=I-i+1}^{J} \mathrm{Cov}\left(\frac{\widehat{\gamma}_j^{(GLM)}}{\sum_{l=0}^{J} \widehat{\gamma}_l^{(GLM)}}, \frac{\widehat{\gamma}_k^{(GLM)}}{\sum_{l=0}^{J} \widehat{\gamma}_l^{(GLM)}}\right) \\
&= \sum_{j,k=I-i+1}^{J} \mathrm{Cov}\left(\frac{1}{1+\sum_{l\neq j} \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}}, \frac{1}{1+\sum_{l\neq k} \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}}\right);
\end{aligned} \tag{5.15}$$

see e.g. (4.19). Here, we restricted our probability space such that a solution to the above equations exist. We define

$$\Delta_j = \sum_{\substack{l=0 \\ l\neq j}}^{J} \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}} \qquad \text{and} \qquad \delta_j = \mathrm{E}[\Delta_j] \tag{5.16}$$

for all $I-i+1 \leq j \leq J$. Hence we need to calculate

$$\mathrm{Cov}\left(\frac{1}{1+\Delta_j}, \frac{1}{1+\Delta_k}\right) \tag{5.17}$$

for all $I-i+1 \leq j, k \leq J$. We first do a Taylor approximation around δ_j. To this end we define the function

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

$$f(x) = \frac{1}{1+x} \qquad \text{with} \qquad f'(x) = -\frac{1}{(1+x)^2} \tag{5.18}$$

and obtain for the first order Taylor approximation around δ_j

$$f(x) \approx f(\delta_j) + f'(\delta_j)(x-\delta_j) = \frac{1}{1+\delta_j} - \frac{1}{(1+\delta_j)^2}(x-\delta_j). \tag{5.19}$$

This implies that

$$\begin{aligned} \mathrm{Cov}\left(\frac{1}{1+\Delta_j}, \frac{1}{1+\Delta_k}\right) &\approx \frac{1}{(1+\delta_j)^2}\frac{1}{(1+\delta_k)^2}\mathrm{Cov}(\Delta_j, \Delta_k) \\ &= \frac{1}{(1+\delta_j)^2}\frac{1}{(1+\delta_k)^2}\sum_{l\neq j}\sum_{m\neq k}\mathrm{Cov}\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}, \frac{\widehat{\gamma}_m^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}\right). \end{aligned} \tag{5.20}$$

It only remains to calculate the covariance terms on the right-hand side of (5.20). We use the following linearization (Taylor approximation for the exponential function)

$$\begin{aligned} \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}} &= \exp\left(\log\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\right)\right) \\ &= \frac{\gamma_l}{\gamma_j}\exp\left(\log\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\right) - \log\left(\frac{\gamma_l}{\gamma_j}\right)\right) \\ &\approx \frac{\gamma_l}{\gamma_j}\left(1 + \log\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\right) - \log\left(\frac{\gamma_l}{\gamma_j}\right)\right). \end{aligned} \tag{5.21}$$

Using (5.21) we obtain for the covariance terms on the right-hand side of (5.20)

$$\mathrm{Cov}\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}, \frac{\widehat{\gamma}_m^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}\right) \approx \frac{\gamma_l}{\gamma_j}\frac{\gamma_m}{\gamma_k}\mathrm{Cov}\left(\log\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\right), \log\left(\frac{\widehat{\gamma}_m^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}\right)\right). \tag{5.22}$$

Now, we define the slightly modified design matrices

$$\widetilde{\Gamma}_j = (0, \ldots, 0, e_{I+j+1}, 0, \ldots, 0)' \qquad \text{for all } 0 \le j \le J, \tag{5.23}$$

which implies that, see (4.16) and (4.18),

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

$$\log\left(\widehat{\gamma}_j^{(GLM)}\right) = \widetilde{\Gamma}_j \widehat{\boldsymbol{b}} \qquad \text{for all } 0 \le j \le J. \tag{5.24}$$

Hence

$$\mathrm{Cov}\left(\log\left(\frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\right), \log\left(\frac{\widehat{\gamma}_m^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}\right)\right) = \left(\widetilde{\Gamma}_l - \widetilde{\Gamma}_j\right)\mathrm{Cov}\left(\widehat{\boldsymbol{b}}, \widehat{\boldsymbol{b}}\right)\left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_k\right)'. \tag{5.25}$$

Using the inverse of the Fisher information matrix $H(\widehat{\boldsymbol{b}})$ for the estimation of the covariance term $\mathrm{Cov}(\widehat{\boldsymbol{b}}, \widehat{\boldsymbol{b}})$ we obtain for (5.15)

$$\mathrm{Var}\left(\sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}\right) \approx \sum_{j,k=I-i+1}^{J} \frac{1}{(1+\delta_j)^2}\frac{1}{(1+\delta_k)^2} \times \sum_l \sum_m \frac{\gamma_l}{\gamma_j}\frac{\gamma_m}{\gamma_k}\left(\widetilde{\Gamma}_l - \widetilde{\Gamma}_j\right) H(\widehat{\boldsymbol{b}})^{-1}\left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_k\right)'; \tag{5.26}$$

see e.g. Remarks 4.4. Hence we define the estimator

$$\widehat{\mathrm{Var}}\left(\sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}\right) = \sum_{j,k=I-i+1}^{J} \frac{1}{(1+\widehat{\delta}_j)^2}\frac{1}{(1+\widehat{\delta}_k)^2} \times \sum_l \sum_m \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}\frac{\widehat{\gamma}_m^{(GLM)}}{\widehat{\gamma}_k^{(GLM)}}\left(\widetilde{\Gamma}_l - \widetilde{\Gamma}_j\right) H(\widehat{\boldsymbol{b}})^{-1}\left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_k\right)', \tag{5.27}$$

where we set $\widehat{\delta}_j = \sum_{\substack{l=0 \\ l\neq j}}^{J} \frac{\widehat{\gamma}_l^{(GLM)}}{\widehat{\gamma}_j^{(GLM)}}$; see e.g. (5.16). This can be rewritten in matrix notation, we define the parameter $c_{j,k} = \widehat{\gamma}_j^{(GLM)}\widehat{\gamma}_k^{(GLM)}\left(\widehat{\mu}_0^{(MLE)}\right)^{-4}$ and $\widehat{\boldsymbol{\gamma}} = (0, \ldots, 0, \widehat{\gamma}_0^{(GLM)}, \ldots, \widehat{\gamma}_J^{(GLM)})'$. Furthermore, we define

$$\Psi_{j,k} = c_{j,k}\left[\widehat{\boldsymbol{\gamma}}' H(\widehat{\boldsymbol{b}})^{-1}\widehat{\boldsymbol{\gamma}} - \widehat{\mu}_0^{(MLE)}\widehat{\boldsymbol{\gamma}}' H(\widehat{\boldsymbol{b}})^{-1}\left(\widetilde{\Gamma}_j + \widetilde{\Gamma}_k\right) + \left(\widehat{\mu}_0^{(MLE)}\right)^2 H(\widehat{\boldsymbol{b}})^{-1}_{j,k}\right]. \tag{5.28}$$

Then we obtain

$$\widehat{\mathrm{Var}}\left(\sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}\right) = \sum_{j,k=I-i+1}^{J} \Psi_{j,k}. \tag{5.29}$$

Putting the three estimates (5.6), (5.13) and (5.29) together we obtain the following estimator for the (conditional) MSEP for a single accident year:

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Estimator 5.1 (MSEP for the BF method, single accident year)

Under Model Assumptions 4.1 an estimator for the (conditional) MSEP for a single accident year $I-J+1 \le i \le I$ is given by

$$\widehat{\mathrm{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF}\right) = \widehat{\phi}\left(1-\widehat{\beta}_{I-i}^{(CL)}\right)\widehat{\mu}_i + \left(1-\widehat{\beta}_{I-i}^{(CL)}\right)^2 \widehat{\mu}_i^2 \widehat{\mathrm{Vco}}(\widehat{\mu}_i)^2 + \widehat{\mu}_i^2 \sum_{j,k=I-i+1}^{J} \Psi_{j,k}, \tag{5.30}$$

see (5.6), (5.13) *and* (5.29).

5.2 MSEP in the BF Method, Aggregated Accident Years

In this subsection we derive an estimate for the conditional MSEP for aggregated accident years (5.3). We start by considering two different accident years $i < l$,

$$\mathrm{msep}_{C_{i,J}+C_{l,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF} + \widehat{C_{l,J}}^{BF}\right) = E\left[\left(C_{i,J} + C_{l,J} - \widehat{C_{i,J}}^{BF} - \widehat{C_{l,J}}^{BF}\right)^2 \middle| \mathcal{D}_I\right]. \tag{5.31}$$

By the usual decomposition we find

$$\begin{aligned}
&\mathrm{msep}_{C_{i,J}+C_{l,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF} + \widehat{C_{l,J}}^{BF}\right) \\
&= \mathrm{msep}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C_{i,J}}^{BF}\right) + \mathrm{msep}_{C_{l,J}|\mathcal{D}_I}\left(\widehat{C_{l,J}}^{BF}\right) \\
&\quad + 2\mu_i\mu_l\left(\sum_{j=I-i+1}^{J}\gamma_j - \sum_{j=I-i+1}^{J}\widehat{\gamma}_j^{(MLE)}\right)\left(\sum_{k=I-l+1}^{J}\gamma_k - \sum_{k=I-l+1}^{J}\widehat{\gamma}_k^{(MLE)}\right).
\end{aligned} \tag{5.32}$$

That is, we need to give an estimate for the term on the right-hand side of (5.32). Analogously to (5.14), we have to study the fluctuations of the MLEs $\widehat{\gamma}_j^{(MLE)}$ around the true parameters γ_j. Again, neglecting the possible bias of the MLEs, we estimate this term by

$$\begin{aligned}
&\mathrm{Cov}\left(\sum_{j=I-i+1}^{J}\widehat{\gamma}_j^{(MLE)}, \sum_{k=I-l+1}^{J}\widehat{\gamma}_k^{(MLE)}\right) \\
&= \sum_{j=I-i+1}^{J}\sum_{k=I-l+1}^{J}\mathrm{Cov}\left(\widehat{\gamma}_j^{(MLE)}, \widehat{\gamma}_k^{(MLE)}\right) \\
&\approx \sum_{j=I-i+1}^{J}\sum_{k=I-l+1}^{J}\frac{1}{(1+\delta_j)^2}\frac{1}{(1+\delta_k)^2} \\
&\quad \times \sum_n\sum_m \frac{\gamma_n}{\gamma_j}\frac{\gamma_m}{\gamma_k}\left(\widetilde{\Gamma}_n - \widetilde{\Gamma}_j\right)H(\widehat{\boldsymbol{b}})^{-1}\left(\widetilde{\Gamma}_m - \widetilde{\Gamma}_k\right)';
\end{aligned} \tag{5.33}$$

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

see e.g. (5.26). Again, as with the single accident year case, we restrict our probability space such that the above covariance exists. This motivates the following estimator for the covariance term (5.33)

$$\Upsilon_{i,l} = \widehat{\mathrm{Cov}}\left(\sum_{j=I-i+1}^{J} \widehat{\gamma}_j^{(MLE)}, \sum_{k=I-l+1}^{J} \widehat{\gamma}_k^{(MLE)} \right) = \sum_{j=I-i+1}^{J} \sum_{k=I-l+1}^{J} \Psi_{j,k}. \quad (5.34)$$

This leads to the following estimator for the (conditional) MSEP for aggregated accident years:

Estimator 5.2 (MSEP for the BF method, aggregated accident years)
Under Model Assumptions 4.1 an estimator for the (conditional) MSEP for aggregated accident years is given by

$$\widehat{\mathrm{msep}}_{\sum_{i=1}^{I} C_{i,J}|\mathcal{D}_I}\left(\sum_{i=1}^{I} \widehat{\widehat{C}_{i,J}}^{BF} \right) = \sum_{i=1}^{I} \widehat{\mathrm{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{\widehat{C}_{i,J}}^{BF} \right) + 2 \sum_{1 \le i < l \le I} \widehat{\mu}_i \widehat{\mu}_l \Upsilon_{i,l}. \quad (5.35)$$

The natural extension of this work would be to obtain some properties of Estimators 5.1 and 5.2, for example, asymptotic behaviour. However, we omit this presently, because it would go beyond the context of this work.

6. Example and Simulation

In this section we state an example and a simulation.

6.1 *Example of MSEP in the BF Method*

Using the incremental claims data provided in Table 1, which are scaled incremental payments from property business, we calculate the BF reserves and the estimators for the conditional MSEP as derived for Estimator 5.1 (single accident year) and Estimator 5.2 (aggregated accident years). Furthermore, we compare these results with the estimators for the conditional MSEP of the overdispersed Poisson method/Poisson GLM method, with reserves matching those of the CL method. However, before we can estimate the conditional MSEP for the BF method, we must first specify the assumptions regarding ultimate claim estimates $\widehat{\mu}_i$.

We assume the prior estimates $\widehat{\mu}_i$ for the expected ultimate claims $E[C_{i,J}]$ are given by Table 2. Furthermore, we have to specify the uncertainty in these estimates. To this end we assume for the coefficient of variation a flat rate of 5% for all accident years, i.e.

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

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

	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									

Table 2. Prior estimates for the expected ultimate claims

i	$\widehat{\mu}_i$
0	11,653,101
1	11,367,306
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

$$\widehat{\text{Vco}}(\widehat{\mu}_i) = 0.05 \qquad \text{for } 1 \le i \le I. \tag{6.1}$$

This assumption, as previously stated, has been shown to be reasonable in the Swiss Solvency Test (2006). To view the effect of changing this input see Table 5 where we have presented the BF MSEP as a function of the coefficient of variation for the expected ultimate estimates. For the dispersion parameter ϕ we obtain the estimate

$$\widehat{\phi} = 14{,}714; \tag{6.2}$$

see e.g. (5.9).

Now we can calculate the estimators for the conditional MSEPs as given by Estimator 5.1 and 5.2 and obtain the values given in Table 3. If we compare these results to the results for the overdispersed Poisson GLM method in Table 4 we observe the following: (a) The claims reserve estimates are rather large for the BF method, reflecting the conservative prior estimates $\widehat{\mu}_i$ for the expected ultimate claims given by Table 2. (b) As a

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

Table 3. Results for the stochastic BF method

i	BF reserves	Process std. dev.	Prior std. dev.	Parameter β 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%
total	7,356,575	329,007	249,828	228,249	338,396	471,971	6.4%

Table 4. Results for the overdispersed Poisson method

i	Reserves	Process std. dev.	Parameter std. dev.	msep$^{1/2}$	Vco
1	15,125	14,918	14,611	20,882	138.1%
2	26,257	19,656	17,160	26,093	99.4%
3	34,538	22,543	17,159	28,331	82.0%
4	85,301	35,428	22,040	41,724	48.9%
5	156,493	47,986	27,108	55,113	35.2%
6	286,120	64,885	32,927	72,761	25.4%
7	449,166	81,296	38,935	90,139	20.1%
8	1,043,242	123,897	66,175	140,462	13.5%
9	3,950,816	241,107	227,661	331,605	8.4%
total	6,047,059	298,290	309,563	429,891	7.1%

Table 5. The BF msep$^{1/2}$ as a function of $\widehat{\text{Vco}}(\hat{\mu}_i)$

$\widehat{\text{Vco}}(\hat{\mu}_i)$	msep$^{1/2}$
0.00	400,428
0.01	403,534
0.02	412,710
0.03	427,565
0.04	447,535
0.05	471,971
0.06	500,219
0.07	531,671
0.08	565,794
0.09	602,133
0.10	640,311

https://doi.org/10.1017/S1748499500000580 Published online by Cambridge University Press

consequence of (a) the process standard deviations of the BF method are larger than the ones of the overdispersed Poisson GLM method. (c) The totals of prior and parameter uncertainty in the estimators $\widehat{\mu}_i$ and $\widehat{\beta}_j^{CL}$ are slightly higher than the estimation errors in the overdispersed Poisson GLM method. As a consequence, the conditional MSEPs of the BF method are larger than the ones of the overdispersed Poisson GLM method. However, due to the conservative estimation of the claims reserves, the coefficients of variation are smaller than the ones of the overdispersed Poisson GLM model.

6.2 *Simulation*

In the derivation of Estimators 5.1 and 5.2 we used various approximations. We test the strength of our results by simulation. Assuming the incremental claims $X_{i,j}$ are overdispersed Poisson distributed with parameters $\widehat{\gamma}_j^{(MLE)}$ and $\widehat{\mu}_j^{(MLE)}$ and dispersion parameter (6.2), see e.g. Model Assumptions 4.1, we generate 10,000 run-off triangles $\mathcal{D}_I$ from which we calculate (for each simulation/run-off triangle) the estimated payout pattern $\widehat{\beta}_j^{(CL)}$. From the generated samples we obtain the empirical distribution of the payout pattern, and hence calculate its empirical standard deviation.

Table 6 provides the resulting empirical payout pattern, the empirical variance and the estimated variance for each development year *j*. From these results it is clear that the approximation of the variance of the cumulative payout pattern is very close to the empirical value. This means that the approximation given in (5.26) performs very well for a typical payout pattern in practice.

Table 6. Simulation results comparing empirical and estimated standard deviations of the cumulative payout pattern

j	$\widehat{\beta}_j^{(CL)}$	Empirical s.d.$(\widehat{\beta}_j^{(CL)})$	Estimated s.d.$(\widehat{\beta}_j^{(CL)})$ using (5.27)
0	58.96%	0.654%	0.653%
1	88.00%	0.486%	0.484%
2	94.84%	0.373%	0.370%
3	97.01%	0.317%	0.313%
4	98.45%	0.260%	0.258%
5	99.14%	0.220%	0.219%
6	99.65%	0.177%	0.175%
7	99.75%	0.162%	0.160%
8	99.86%	0.138%	0.137%
9	100.00%	0.000%	0.000%

Acknowledgements

The authors would like to thank Prof. Paul Embrechts for helpful discussion related to the paper. Furthermore, the first author is grateful to ACE Limited for providing the financial support that made this research a reality.

References

Bornhuetter, R.L. & Ferguson, R.E. (1972). The actuary and IBNR. *Proceedings of the Casualty Actuarial Society*, **LIX**, 181-195.

Bühlmann, H. & Gisler, A. (2005). *A course in credibility theory and its applications*. Springer, Berlin.

Casualty Actuarial Society (1990). *Foundations of casualty actuarial science.* 4th ed. Casualty Actuarial Society, New York.

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

England, P.D. & Verrall, R.J. (2007). Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, **1**, 221-270.

Fahrmeir, L. & Tutz, G. (2001). *Multivariate statistical modelling based on generalized linear models.* 2nd ed. Springer, Berlin.

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

Lambrigger, D.D., Shevchenko, P.V. & Wüthrich, M.V. (2007). The quantification of operational risk using internal data, relevant external data and expert opinion. *Journal of Operational Risk*, **2**, 3-27.

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

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

Mack, T. (2000). Credible claims reserves: The Benktander method. *ASTIN Bulletin*, **30**, 333-347.

Mack, T. (2006). Parameter estimation for Bornhuetter/Ferguson. *CAS Forum*, **Fall 2006**, 141-157.

Mack, T., Quarg, G. & Braun, C. (2006). The mean square error of prediction in the chain ladder reserving method — a comment. *ASTIN Bulletin*, **36**, 543-552.

McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models.* 2nd ed. Chapman and Hall, London.

Panjer, H.H. (2006). *Operational risk: modeling analytics*. Wiley, New York.

Radtke, M. & Schmidt, K.D. (2004). *Handbuch zur Schadenreservierung*. Verlag Versicherungswirtschaft, Karlsruhe.

Renshaw, A.E. (1995). Claims reserving by joint modelling. Presented at ASTIN Colloquium, Cannes.

Renshaw, A.E. & Verrall, R.J. (1998). A stochastic model underlying the chain ladder technique. *British Actuarial Journal*, **4**, 903-923.

Swiss Solvency Test (2006). BPV SST Technisches Dokument, Version October 2, 2006.

Taylor, G. (2000). *Loss reserving: an actuarial perspective*. Kluwer Academic Publishers, Boston.

Teugels, J.L. & Sundt, B. (2004). *Encyclopedia of actuarial science.* Volume 1. Wiley, Chichester.

Verrall, R.J. (1990). Bayesian and empirical Bayes estimation for the chain ladder model. *ASTIN Bulletin*, **20**, 217-238.

VERRALL, R.J. (2000). An investigation into stochastic claims reserving models and the chain-ladder technique. *Insurance: Mathematics & Economics*, **26**, 91-99.

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

VERRALL, R.J. & ENGLAND, P.D. (2000). Comments on: "A comparison of stochastic models that reproduce chain ladder reserve estimates", by Mack and Venter. *Insurance: Mathematics & Economics*, **26**, 109-111.

WÜTHRICH, M.V., MERZ, M. & BÜHLMANN, H. (2008). Bounds on the estimation error in the chain ladder method. *Scandinavian Actuarial Journal*, **4**, 283-300.