THE PREDICTION ERROR OF BORNHUETTER/FERGUSON

BY

THOMAS MACK

ABSTRACT

Together with the Chain Ladder (CL) method, the Bornhuetter/Ferguson (BF) method is one of the most popular claims reserving methods. Whereas a formula for the prediction error of the CL method has been published already in 1993, there is still nothing equivalent available for the BF method. On the basis of the BF reserve formula, this paper develops a stochastic model for the BF method. From this model, a formula for the prediction error of the BF reserve estimate is derived.

Moreover, the model gives important advice on how to estimate the parameters for the BF reserve formula. E.g. it turns out that the appropriate BF development pattern is different from the CL pattern. This is a nice add-on as it makes BF to a standalone reserving method which is fully independent from CL. The other parameter required for the BF reserve is the well-known initial estimate for the ultimate claims amount. Here the stochastic model clearly shows what has to be meant with 'initial'.

In order to apply the formula for the prediction error, the actuary must assess his uncertainty about both sets of parameters, about the development pattern and about the initial ultimate claims estimates. But for both, much guidance can be drawn from the estimates itself and from the run-off data given. Finally, a numerical example shows how the resulting prediction error compares to the one of the CL method.

1. INTRODUCTION

For most insurance companies and their auditors, the use of the Chain Ladder method (CL) and of the Bornhuetter/Ferguson method (BF) has become a certain standard or benchmark in claims reserving. This means that these methods are applied in almost every case, and only if they seem to fail, one looks for other methods. Originally, these methods gave only a point estimate for the claims reserve. But this was not satisfactory because then one could not decide whether the estimates differ significantly or not. Moreover, for the calculation of risk based capital and of premium loadings one needs to assess the prediction error of the estimate (i.e. the standard deviation of the true claims reserve from the point estimate).

Astin Bulletin 38(1), 87-103. doi: 10.2143/AST.38.1.2030404

In 1993, a formula for the prediction error of the CL reserve estimate was published (Mack (1993)) which in the mean time is very popular. This formula gives an answer to the question of significant differences to other methods and measures the variability of the true reserves for business segments where CL is acceptable. But for BF, such a formula is still missing. This may seem strange because BF is even simpler than CL. But this simplicity is just the problem. The prediction error consists of two components, the process error and the estimation error. Whereas the estimation error basically always can be calculated via the laws of error propagation, for the process error a stochastic model of the claims process is required. The latter was feasible in the CL case because the way in which the CL age-to-age factors are estimated contains implicit information on the underlying stochastics. In the BF case, no clear procedure on how to estimate the parameters has been established. In such a situation, many models may seem admissible.

The stochastic model for BF which is introduced in this paper is very similar in its structure to the CL model of Mack (1993) but adequately reflects the two fundamental differences between CL and BF. The first difference is the fact that the CL reserve is directly proportional to the claims amount known so far whereas the BF reserve does not depend at all from the known claims amount. This is reflected in an additional independence assumption of the BF model. The second difference is the fact that the BF reserve estimate includes the full tail of the claims development whereas the standard CL reserve (i.e. without additional tail factor) only considers the development until a given last development year. The latter fact implies that the parameter estimation for the BF model has also to consider the tail of the development where there is no data and some judgement is required. Therefore, we do not give a unique estimation formula for the tail parameters but discuss two alternative ways to cope with this problem. In any case, the development pattern suggested by the BF model turns out to be different from the well-known CL pattern. This makes BF to a really standalone reserving method. But still, the actuary may make his own selections regarding the development pattern, especially for the tail.

In addition to the development pattern, the BF reserve formula requires another element, an initial estimate for the ultimate claims amount. Of course, the uncertainty of this estimate must have a high impact on the prediction error. As this estimate usually comes from outside (e.g. from pricing) or is simply set by the actuary on the basis of his knowledge of the business, its uncertainty must be assessed from outside of the run-off triangle, too. And an actuary who is able to set (or accept) a point estimate should also be able to quantify (or ask for quantification of) the uncertainty of this estimate. Moreover, from the stochastic model important advice can be derived for the assessment of these estimates and their uncertainty. Altogether, this means that the prediction error of the BF reserve estimate depends largely on the (more or less subjective) assessment of the actuary as it is already the case with the BF reserve estimate itself.

Section 2 gives a short review of the BF method and of its connections and differences to the CL method. Section 3 describes the appropriate stochastic BF model. Section 4 shows two ways to estimate or select the model parameters. The estimation of the standard error of the parameters is discussed in Section 5 where also the formula for the prediction error and its components is derived. Section 6 gives a numerical example.

2. The BF method

Let $C_{i,k}$ denote the cumulative claims amount (either paid or incurred) of accident year i after k years of development, $1 \le i,\ k \le n$, and v_i be the premium volume of accident year i where n denotes the most recent accident year. Then $C_{i,n+1-i}$ denotes the currently known claims amount of accident year i. Let further $S_{i,k} = C_{i,k} - C_{i,k-1}$ denote the incremental claims amount (with $C_{i,0} = 0$) and U_i the (unknown) ultimate claims amount of accident year i. Then $R_i = U_i - C_{i,n+1-i}$ is the (unknown true) claims reserve for accident year i. Let finally $S_{i,n+1} = U_i - C_{i,n}$ be the incremental claims amount after development year n (tail development).

Bornhuetter/Ferguson (1972) introduced their method to estimate R_i in order to cope with a major weakness of the CL method. Therefore we first consider this weakness. CL uses link ratios (age-to-age factors) $\hat{f}_k$ and a tail factor $\hat{f}_\infty$ in order to project the current claims amount $C_{i,n+1-i}$ to ultimate, i.e. it estimates $\hat{U}_i^{CL} = C_{i,n+1-i} \cdot \hat{f}_{n+2-i} \cdot \ldots \cdot \hat{f}_n \cdot \hat{f}_\infty$, and therefore the CL reserve is

$$\hat{R}_i^{CL} = \hat{U}_i^{CL} - C_{i,n+1-i} = C_{i,n+1-i}\left(\hat{f}_{n+2-i} \cdot \ldots \cdot \hat{f}_\infty - 1\right).$$

This means that the reserve strongly depends on the current amount $C_{i,n+1-i}$ which can e.g. lead to a nonsense reserve $\hat{R}_i^{CL} = 0$ for accident years where currently no claims are paid or reported which is not unusual in excess-of-loss reinsurance for the most recent accident year(s).

The BF reserve estimate avoids this dependency from the current claims amount $C_{i,n+1-i}$. It is

$$\hat{R}_i^{BF} = \hat{U}_i(1 - \hat{z}_{n+1-i})$$

where $\hat{U}_i = v_i \hat{q}_i$ with a prior estimate $\hat{q}_i$ for the ultimate claims ratio $q_i = U_i / v_i$ of accident year i, $\hat{z}_k \in [0,1]$ is the estimated percentage of the ultimate claims amount which is expected to be known after development year k.

$\hat{q}_i$ is called 'prior' (or 'initial') as opposed to the posterior estimate $(C_{i,n+1-i} + \hat{R}_i^{BF}) / v_i$ for the ultimate claims ratio which is based on the prior $\hat{q}_i$ and is different iff $C_{i,n+1-i} \neq \hat{z}_{n+1-i} v_i \hat{q}_i$, i.e. if the current claims amount deviates from its estimated expectation. The percentages $z_1, z_2, \ldots$ constitute the expected cumulative development pattern and $1 - \hat{z}_{n+1-i}$ is therefore an estimate for the percentage of the expected outstanding claims of accident year i.

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

Having already an estimate $\hat{U}_i$, the question may arise why BF does not simply use $\hat{R}_i = \hat{U}_i - C_{i,n+1-i}$ as reserve estimate. In that case, the reserve estimate would become the higher, the smaller the current amount $C_{i,n+1-i}$ is and would again strongly depend on $C_{i,n+1-i}$. With CL, the reserve estimate behaves just in the opposite way, i.e. is the smaller, the smaller $C_{i,n+1-i}$ is. Here BF takes a neutral position: It does not care about the size of $C_{i,n+1-i}$ at all, i.e. it considers the deviation between the observed amount $C_{i,n+1-i}$ and the expected amount $\hat{z}_{n+1-i}\hat{U}_i$ as purely random and by no means indicative for the future development. Altogether, the essential feature of the BF method is to avoid any dependency between $C_{i,n+1-i}$ and $\hat{R}_i^{BF}$.

In order to apply the BF method, the actuary has to estimate the parameters q_i and z_k for all i and k. In practice, the ultimate claims ratios q_i are estimated in various ways, mainly based on additional pricing and market information in such a way that any expected differences between the accident years are reasonably reflected. The z_k are usually derived from the (selected) CL link ratios $\hat{f}_2, \ldots, \hat{f}_n$ together with a selected tail factor $\hat{f}_\infty$ in the following way:

$$\hat{z}_n = \hat{f}_\infty^{-1},\ \hat{z}_{n-1} = (\hat{f}_n \cdot \hat{f}_\infty)^{-1}, \ldots, \hat{z}_1 = (\hat{f}_2 \cdot \ldots \cdot \hat{f}_n \cdot \hat{f}_\infty)^{-1}.$$

The systematic use of the CL link ratios assumes that the outstanding claims part is a direct multiple of the already known part at each point of the development. This contradicts to the basic BF idea of the independence between $C_{i,n+1-i}$ and $\hat{R}_i^{BF}$, i.e. between past and future claims, which was fundamental for the origin of the BF method. At least, with the use of the CL pattern, the BF method cannot really claim to be a standalone reserving method. Moreover, in the following we will see that the stochastic BF model suggests a different way to estimate the BF development pattern.

3. A stochastic model underlying the BF method

From the BF reserve formula it is clear that the appropriate model for BF has to be cross-classified of the type

$$E(C_{i,k}) = x_i z_k \text{ or equivalently } E(S_{i,k}) = x_i y_k \text{ for } 1 \le i \le n \text{ and } 1 \le k \le n+1.$$

Because of $x_i y_k = (x_i a)(y_k/a)$ for any $a > 0$, x_i and y_k are only unique up to a constant factor. Thus we can – without loss of generality – impose the restriction $y_1 + \ldots + y_n + y_{n+1} = 1$. This yields $E(U_i) = E(S_{i,1} + \ldots + S_{i,n+1}) = x_i$ and shows that x_i can be considered to be a measure of volume for accident year i. We therefore will assume in addition that $Var(U_i)$ is proportional to x_i or $Var(U_i/x_i)$ proportional to $1/x_i$. This is the usual assumption for the influence of the volume on the variance. Furthermore, the fundamental BF property of independence between past and future claims suggests to assume that all

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

increments $S_{i,k}$ of the same accident year are independent – the independence of the accident years themselves being a standard assumption anyway. Note that the independence within the accident years does not hold in the CL model of Mack (1993).

Thus we work with the following model for the increments $S_{i,k}$, $1 \le i \le n$, $1 \le k \le n+1$:

(BF1) All increments $S_{i,k}$ are independent.

(BF2) There are unknown parameters x_i, y_k with $E(S_{i,k}) = x_i y_k$ and $y_1 + \ldots + y_{n+1} = 1$.

(BF3) There are unknown proportionality constants s_k^2 with $Var(S_{i,k}) = x_i s_k^2$.

From these assumptions, we deduce

$$E(R_i) = x_i(y_{n+2-i} + \ldots + y_{n+1}) = x_i(1 - z_{n+1-i}) \text{ with } z_k := y_1 + \ldots + y_k$$

which shows that the expected claims reserve has the same form as the BF reserve estimate.

This model is thought to be the most general model fitting to the philosophy of the BF method. Like with the CL model and as suggested by having only one single column parameter y_k for the expectation, it here, too, makes sense to assume that the variability constant s_k^2 is the same for all $S_{i,k}$ within each column k but differs from column to column. The simpler assumption $Var(S_{i,k}) = c x_i y_k$ for all i,k seems to contradict to reality as has already been mentioned by Taylor (2002) because then 'the coefficient of variation of the claim size is inversely related to the mean claim size' which is 'opposite of what one observes'. Moreover, this last variance assumption is just a special case of (BF3) and thus less general. Finally, this variance assumption would imply that all y_k be > 0 which is not the case with (BF3) and which would prevent from using the model for incurred claims amounts where negative incremental claims are not uncommon.

Like with the CL model of Mack (1993), this model is heavily parametrized, especially for the late development years. But of course, the actuary may – depending on the data – apply additional regression assumptions in order to reduce the number of parameters and to stabilize the estimates. This is shown in the numerical example below.

From the above model, we deduce further

$$Var(R_i) = x_i \left(s_{n+2-i}^2 + \ldots + s_{n+1}^2 \right).$$

As background for the next section, we note that with $x_1, \ldots, x_n$ known,

$$\hat{y}_k = \sum_{i=1}^{n+1-k} S_{i,k} \Big/ \sum_{i=1}^{n+1-k} x_i, \tag{1}$$

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

is a best linear unbiased estimate of y_k, $1 \le k \le n$, and

$$\hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} \left(S_{i,k} - x_i \hat{y}_k\right)^2 \Big/ x_i \tag{2}$$

is an unbiased estimate of s_k^2, $1 \le k \le n-1$.

4. Parameter estimation for the BF model

From the model above we clearly see what is meant with calling $\hat{U}_i$ a 'prior' estimate: It has to be an estimate $\hat{x}_i$ for $x_i = E(U_i)$ and not for the 'posterior' expectation $E(U_i \mid C_{i,n+1-i})$, given $C_{i,n+1-i}$. This shows that the claims amount $C_{i,n+1-i} = S_{i,1} + \ldots + S_{i,n+1-i}$ known so far should not be the main basis for the estimate $\hat{x}_i$. For example, it would be wrong to use the posterior estimate $C_{i,n-i} + \hat{R}_i^{BF(n-1)}$ of last year's reserving as $\hat{x}_i$ because it is an estimate rather for $E(U_i \mid C_{i,n+1-i})$ than for $E(U_i)$. Even a very large random claim which happened in accident year i and is already known must not change the estimate $\hat{x}_i$. As an extreme example we might have an accident year where $\hat{x}_i < C_{i,n+1-i}$. But this does not mean that the prior estimate $\hat{x}_i$ cannot change during the claims development.

To fix ideas, let us assume that $\hat{x}_i$ originally stems from pricing (which has taken place before the end of development year 1). Usually, the pricing is based on the (trended) claims experience of the preceding accident years (i.e. on the years $i-1$, $i-2$, …) and on assumptions on the future claims cost inflation. This basic information develops from year to year because the claims experience of the preceding years develops as well as the relevant inflation index. Thus, we can reprice the business of accident year i every later year and thus arrive at updated estimates for $x_i = E(U_i)$. We may even include the claims experience of the accident years i, $i+1$, … into this repricing of accident year i as long as they can be translated to the portfolio of accident year i. In any case, the own claims experience $C_{i,n+1-i}$ should only have a marginal influence on $\hat{x}_i$ otherwise we would rather estimate $E(U_i \mid C_{i,n+1-i})$. Thus, the estimate $\hat{x}_i$ may change over the years but normally not to a large extent, at least if the first estimate for x_i came from a sound pricing.

When the actuary does not have the result of a complete repricing available, he has at least the data $\{v_i, C_{ik}\}$ of the run-off triangle. On basis of this data and some rather general information on rate level changes, he may follow the procedure outlined in Mack (2006) which is not a full repricing but brings all accident years on about the same claims ratio level as basis for the calculation of the initial ultimate claims ratio $\hat{q}_i$.

After these clarifying remarks, we assume that the initial estimate $\hat{U}_i$ of Section 2 fulfills the requirements for being an estimate of $x_i = E(U_i)$. Thus we write $\hat{U}_i$ instead of $\hat{x}_i$ in the following. Having now an estimate $\hat{U}_i$ for $E(U_i)$, we are only left with the task to estimate y_k and s_k^2. The main problem here is the fact that we have only very few observations for the late development years.

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

As we do not have any observations beyond development year n, we cannot estimate the tail ratio y_{n+1} without further assumptions. An outside estimate may be gained from similar portfolios with more accident years where the claims experience of later development years than year n is available. Without such information, the actuary may arrive at an estimate $\hat{y}_{n+1}$ by extrapolation from $\hat{y}_1, \ldots, \hat{y}_n$ (which are not available yet). Similarly, an estimate for s_n^2 cannot be obtained from the only available observation of column n alone but may be obtained by extrapolation, too. Therefore, in order to fix ideas for an iterative procedure, we first consider the situation where we have already reasonable estimates $\hat{y}_{n+1}$, $\hat{s}_1^2, \ldots, \hat{s}_n^2$. Then we can get a weighted least squares estimate (i.e. with the weights inversely proportional to the variances) for $y_1, \ldots, y_n$ by minimizing

$$Q = \sum_{i=1}^{n} \sum_{k=1}^{n+1-i} \frac{\left(S_{i,k} - \hat{U}_i \hat{y}_k\right)^2}{\hat{U}_i \hat{s}_k^2}$$

under the constraint $\hat{y}_1 + \ldots + \hat{y}_n = 1 - \hat{y}_{n+1}$. Note that $\hat{y}_n = S_{1,n}/\hat{U}_1$ may not be optimal due to this constraint. As starting values for the minimization we can use

$$\hat{\hat{y}}_k = \sum_{i=1}^{n+1-k} S_{i,k} \Big/ \sum_{i=1}^{n+1-k} \hat{U}_i, \tag{3}$$

(see (1)) but these will usually not fulfill the constraint.

In most cases the data will not be so stable that the resulting least squares estimates $\hat{y}_1, \ldots, \hat{y}_n$ seem reliable enough to leave them as they are (especially for k large). Therefore, the actuary will apply a smoothing procedure to select his own final $\hat{y}_1^*, \ldots, \hat{y}_n^*, \hat{y}_{n+1}^*$ (i.e. including a possible revision of the tail ratio in view of the other $\hat{y}_k^*$) with $\hat{y}_1^* + \ldots + \hat{y}_n^* + \hat{y}_{n+1}^* = 1$.

On the basis of the fact that the actuary will in any case make some own selections due to the few data, he can dispense with the above exact minimization and just proceed as follows: He starts with the raw estimates $\hat{\hat{y}}_k$, $1 \leq k \leq n$, as given in (3) and applies some manual smoothing and extrapolating in order to arrive at his final selection for $\hat{y}_1^*, \ldots, \hat{y}_n^*, \hat{y}_{n+1}^*$ fulfilling $\hat{y}_1^* + \ldots + \hat{y}_n^* + \hat{y}_{n+1}^* = 1$. In view of (2), he then estimates s_k^2 by

$$\hat{\hat{s}}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} \left(S_{i,k} - \hat{U}_i \hat{y}_k^*\right)^2 \Big/ \hat{U}_i, \quad 1 \leq k \leq n-1, \tag{4}$$

and again applies some smoothing in order to select his final $\hat{s}_1^{2*}, \ldots, \hat{s}_{n-1}^{2*}$ and an extrapolation to obtain $\hat{s}_n^{2*}$. Note that $\hat{s}_{n+1}^{2*}$ cannot be obtained in this way because it usually has to cover several development years as is the case for $\hat{y}_{n+1}$, too. Therefore, $\hat{s}_{n+1}^{2*}$ may be arrived at by interpolating a regression of $\hat{s}_k^{2*}$ against $|\hat{y}_k^*|$ at the point $|\hat{y}_{n+1}^*|$. (Note that some $\hat{y}_k$ may be negative.) The whole estimation procedure is shown in the numerical example.

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

A more formal way to estimate the parameters y_k, s_k^2 (in case of rather stable data) would be as follows: On the basis of $\hat{\hat{y}}_k$, $1 \le k \le n$ we decide on the formula for a smoothing regression, e.g. $ln(\hat{y}_k) = \alpha - \beta \cdot k$ for k above some $k_1 < n$ (assuming $y_k > 0$ there), which then is extrapolated until some final development year $k_2 > n$. Then we calculate $\hat{\hat{s}}_k^2$ (according to (4) but using the smoothened $\hat{y}_k$ for $k > k_1$). The resulting values $\hat{\hat{s}}_1^2, \ldots, \hat{\hat{s}}_{n-1}^2$ are now kept fixed and used in the above constrained minimization of Q to obtain better values for $\hat{y}_1, \ldots, \hat{y}_{k_1}, \alpha, \beta$ under the constraint

$$\hat{y}_1 + \ldots + \hat{y}_{k_1} + exp\big(\alpha - \beta(k_1 + 1)\big) + \ldots + exp(\alpha - \beta k_2) = 1.$$

Note that in Q we have to leave out the term for $(i, k) = (1, n)$ because now we do not yet have a value for $\hat{s}_n$. This minimization yields our selections for all $\hat{y}_k^*$: The values for $k = 1, \ldots, k_1$ are obtained directly, those for $k = k_1 + 1, \ldots, n$ are taken from the smoothing regression and $\hat{y}_{n+1}^*$ is obtained by adding up the extrapolated values of the regression up to development year k_2. Using these $\hat{y}_k^*$, we calculate new values $\hat{\hat{s}}_k^2$ according to (4) and plot $ln(\hat{\hat{s}}_k^2)$ for $k > k_1$ against $|\hat{y}_k^*|$ or $ln\big(|\hat{y}_k^*|\big)$ in order to select appropriate values for $\hat{s}_k^{2*}$, especially for $k = n$ (over $|\hat{y}_n^*|$) and $k = n + 1$ (over $|\hat{y}_{n+1}^*|$). Of course, we could now apply another constraint minimization with these new values of $\hat{s}_k^{2*}$, but usually this will not change much. Note that the values of $\hat{s}_k^{2*}$ for $k > k_1$ will be overestimated a little as we did not change the degrees of freedom in formula (4) for $\hat{\hat{s}}_k^2$ which would have been possible as the regression employs fewer parameters.

As result of each of these two estimation procedures we have selected $\hat{y}_1^*, \ldots, \hat{y}_n^*, \hat{y}_{n+1}^*$ and $\hat{s}_1^{2*}, \ldots, \hat{s}_n^{2*}, \hat{s}_{n+1}^{2*}$ from which we estimate the BF claims reserve by

$$\hat{R}_i^{BF} = \hat{U}_i(\hat{y}_{n+2-i}^* + \ldots + \hat{y}_{n+1}^*) = \hat{U}_i(1 - \hat{z}_{n+1-i}^*) \text{ with } \hat{z}_k^* = \hat{y}_1^* + \ldots + \hat{y}_k^*.$$

$\hat{s}_1^{2*}, \ldots, \hat{s}_n^{2*}, \hat{s}_{n+1}^{2*}$ will be needed for the prediction error.

The properties of the above estimators can be sketched as follows:

(a) $\hat{y}_1^*, \ldots, \hat{y}_n^*, \hat{y}_{n+1}^*$ are pairwise (slightly) negatively correlated as they have to add up to unity.

(b) $\hat{y}_1^*, \ldots, \hat{y}_n^*, \hat{y}_{n+1}^*$ and therefore also $\hat{z}_1^*, \ldots, \hat{z}_{n+1}^*$ are practically independent from $\hat{U}_1, \ldots, \hat{U}_n$ as the latter do not really influence the size of any $\hat{y}_k^*$ because these have to add up to unity in any case and because of selections and regressions used.

(c) $\hat{R}_i^{BF}$ and R_i are independent (due to BF1).

(d) $E(\hat{U}_i) = E(U_i) = x_i$, $1 \le i \le n$.

(e) $E(\hat{y}_k^*) = y_k$, $1 \le k \le n + 1$, and therefore $E(\hat{z}_k^*) = z_k$, $1 \le k \le n + 1$.

(f) $E(\hat{s}_k^{2*}) = s_k^2$, $1 \le k \le n + 1$.

In (d)-(f) we have simply assumed that the actuary's selections are unbiased.

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

The unbiasedness of the reserve estimate follows directly from these properties:

$$E(\hat{R}_i^{BF}) = E(\hat{U}_i)\, E(1-\hat{z}^*_{n+1-i}) = x_i(1-z_{n+1-i}) = E(R_i).$$

Note that the raw estimates $\hat{\hat{y}}_k$ according to (3) are identical to the estimates $\hat{\beta}_k$ in Mack (2006) which were shown there as being suggested directly by the BF reserve formula itself. In any case and even without any smoothing of $\hat{\hat{y}}_k$, the resulting development pattern will turn out to be different from the CL pattern (see also the numerical example below).

Now we are prepared to derive the formula for the prediction error.

5. The prediction error of the BF method

As one is interested in the future variability only, given the data observed so far, the mean squared error of prediction of any reserve estimate $\hat{R}_i$ is defined to be

$$msep\left(\hat{R}_i\right) = E\left(\left(\hat{R}_i - R_i\right)^2 \middle| S_{i,1}, ..., S_{i,n+1-i}\right).$$

According to (BF1), $R_i = S_{i,n+2-i} + \ldots + S_{i,n+1}$ is independent from $S_{i,1}, \ldots, S_{i,n+1-i}$. Also, the BF reserve estimate $\hat{R}_i^{BF}$ can be taken as being independent from $S_{i,1}, \ldots, S_{i,n+1-i}$ (as these play at most a marginal role when selecting $\hat{U}_i$ and $\hat{z}^*_k$), more precisely, R_i and $\hat{R}_i^{BF}$ are taken to be commonly independent from $S_{i,1}, \ldots, S_{i,n+1-i}$. Thus we have

$$\begin{aligned} msep\left(\hat{R}_i^{BF}\right) &= E\left(\left(\hat{R}_i^{BF} - R_i\right)^2\right) \\ &= Var\left(\hat{R}_i^{BF} - R_i\right) + \left(E\left(\hat{R}_i^{BF}\right) - E(R_i)\right)^2 \\ &= Var\left(\hat{R}_i^{BF}\right) + Var(R_i), \end{aligned}$$

i.e. the mean squared error of prediction is the sum of the (squared) estimation error $Var(\hat{R}_i^{BF})$ and of the (squared) process error $Var(R_i)$.

For the process error we simply have

$$Var(R_i) = Var(S_{i,n+2-i}) + \ldots + Var(S_{i,n+1}) = x_i(s^2_{n+2-i} + \ldots + s^2_{n+1})$$

which will be estimated by

$$\hat{Var}(R_i) = \hat{U}_i(\hat{s}^{2*}_{n+2-i} + \ldots + \hat{s}^{2*}_{n+1}).$$

For the estimation error of $\hat{R}_i^{BF} = \hat{U}_i(1-\hat{z}^*_{n+1-i})$ we use the general formula

$$Var(XY) = (E(X))^2\, Var(Y) + Var(X)Var(Y) + Var(X)(E(Y))^2$$

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

for independent random variables X and Y and obtain

$$\begin{aligned} Var(\hat{R}_i^{BF}) &= (E(\hat{U}_i))^2 Var(\hat{z}^*_{n+1-i}) + Var(\hat{U}_i) Var(\hat{z}^*_{n+1-i}) + Var(\hat{U}_i)(1 - E(\hat{z}^*_{n+1-i}))^2 \\ &= (x_i^2 + Var(\hat{U}_i)) Var(\hat{z}^*_{n+1-i}) + Var(\hat{U}_i)(1 - z_{n+1-i})^2. \end{aligned}$$

Whereas we have already estimators $\hat{U}_i$ for x_i and $\hat{z}^*_{n+1-i}$ for z_{n+1-i}, we still need estimates for $Var(\hat{U}_i)$ and $Var(\hat{z}^*_{n+1-i})$, i.e. we have to quantify the precision of $\hat{U}_i$ and $\hat{z}^*_{n+1-i}$.

The standard error $s.e.(\hat{U}_i)$, i.e. an estimate for $\sqrt{Var(\hat{U}_i)}$, cannot be obtained from the estimation error $s.e.(\hat{R}_i^{BF(n-1)})$ of last year's reserving because this would ignore the variability of $C_{i,n+1-i}$ which has to be included into $s.e.(\hat{U}_i)$. Like $\hat{U}_i$ itself, $s.e.(\hat{U}_i)$ is best be obtained from a repricing of the business. But one has to be cautious there. For example, the variability of the posterior claims ratio estimates $\hat{U}_1^{post}/v_1, \ldots, \hat{U}_n^{post}/v_n$ would underestimate $s.e.(\hat{U}_i/v_i)$ because these estimates are positively correlated via the common estimates $\hat{z}^*_k$. Similarly, also the prior estimates $\hat{U}_1, \ldots, \hat{U}_n$ will usually be positively correlated. Thus the formula

$$\left(s.e.\left(\hat{U}_i\right)\right)^2 = \frac{v_i}{n-1} \sum_{j=1}^{n} v_j \left(\frac{\hat{U}_j}{v_j} - \hat{q}\right)^2 \text{ with } \hat{q} = \sum_{j=1}^{n} \hat{U}_j \Big/ \sum_{j=1}^{n} v_j \qquad (5)$$

(which is analogous to (1), (2) for BF3) is applicable only if the prior estimates $\hat{U}_j$ can assumed to be uncorrelated. But even then, using the real premiums v_j would include the market cycle of premium adequacy into $s.e.(\hat{U}_i)$ which would overestimate $s.e.(\hat{U}_i)$ in those situations where we can predict the market cycle rather well. Thus, we should remove the influence of the market cycle from (5) by using on-level premiums $\tilde{v}_j$. In addition, we should correct for any positive correlation between the $\hat{U}_i$'s by replacing the term $n-1$ of (5) with e.g. $n-\sqrt{n}$ for a constant correlation coefficient $\hat{\rho}_{ij}^U = 1/\sqrt{n}$ between $\hat{U}_i$ and $\hat{U}_j$ or with (approximately) $n-\sqrt{2n}$ for a decreasing correlation coefficient $\hat{\rho}_{ij}^U = 1/(1+|i-j|)$; the precise formula being $n - \sum_{i,j} \rho_{ij}^U \sqrt{\frac{v_i}{v_+}\frac{v_j}{v_+}}$ with $v_+ = \sum_{i=1}^{n} v_i$.

These standard errors $s.e.(\hat{U}_i)$ usually will not change much over the years. Of course, we will have slight changes as long as the $\hat{U}_i$ change. But even at the end of the development, we will not know $E(U_i)$ much more precisely than at the beginning. The actuary should plausibilize the resulting values of $s.e.(\hat{U}_i)$, for instance in the following way: If we assume a normal distribution, then the interval $\left(\hat{U}_i - 2 \cdot s.e.(\hat{U}_i), \hat{U}_i + 2 \cdot s.e.(\hat{U}_i)\right)$ will contain the true $E(U_i)$ with 95% probability. Thus, if the size of the interval is plausible, then $s.e.(\hat{U}_i)$ is plausible, too.

Next, we have to decide on how to estimate

$$Var(1 - \hat{z}^*_{n+1-i}) = Var(\hat{z}^*_{n+1-i}) = Var(\hat{y}^*_1 + \ldots + \hat{y}^*_{n+1-i}) = Var(\hat{y}^*_{n+2-i} + \ldots + \hat{y}^*_{n+1}).$$

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

From property (a) we see that we will be on the safe side when we replace $Var(\hat{y}_1^* + \ldots + \hat{y}_{n+1-i}^*)$ with $Var(\hat{y}_1^*) + \ldots + Var(\hat{y}_{n+1-i}^*)$. But whereas the latter sum increases with each additional term, this is not the case with $Var(\hat{y}_1^* + \ldots + \hat{y}_{n+1-i}^*)$ as finally $Var(\hat{y}_1^* + \ldots + \hat{y}_{n+1}^*) = Var(1) = 0$. Therefore we replace $Var(\hat{z}_k^*) = Var(1 - \hat{z}_k^*)$ for small k with $Var(\hat{y}_1^*) + \ldots + Var(\hat{y}_k^*)$ and for large k with $Var(\hat{y}_{k+1}^*) + \ldots + Var(\hat{y}_{n+1}^*)$. More precisely, we replace – still being on the safe side –

$$Var(\hat{z}_k^*) \text{ with } min\Big(Var(\hat{y}_1^*) + \ldots + Var(\hat{y}_k^*),\ Var(\hat{y}_{k+1}^*) + \ldots + Var(\hat{y}_{n+1}^*)\Big).$$

Due to $\hat{y}_k^* \approx \hat{\hat{y}}_k \approx \sum_{j=1}^{n+1-k} S_{j,k} \Big/ \sum_{j=1}^{n+1-k} x_j$ we can assume that

$$Var\Big(\hat{y}_k^*\Big) \approx Var\left(\sum_{j=1}^{n+1-k} S_{j,k} \Big/ \sum_{j=1}^{n+1-k} x_j\right) = \frac{s_k^2}{\sum_{j=1}^{n+1-k} x_j}, \quad 1 \le k \le n.$$

Therefore we estimate $Var\Big(\hat{y}_k^*\Big)$ by

$$\Big(s.e.\Big(\hat{y}_k^*\Big)\Big)^2 = \frac{\hat{s}_k^{2\,*}}{\sum_{j=1}^{n+1-k} \hat{U}_j}, \quad 1 \le k \le n. \tag{6}$$

But the value of $s.e.(\hat{y}_{n+1}^*)$ must come from outside. Without this, a plausible choice is often a coefficient of variation $c.v.(\hat{y}_{n+1}^*) = 50\%$ assuming a normal distribution with 95% probability within the interval $(0;\ 2\,\hat{y}_{n+1}^*)$.

Altogether, our estimate $\big(s.e.(\hat{z}_k^*)\big)^2$ for $Var(\hat{z}_k^*)$ is

$$\big(s.e.(\hat{z}_k^*)\big)^2 = min\Big(\big(s.e.(\hat{y}_1^*)\big)^2 + \ldots + \big(s.e.(\hat{y}_k^*)\big)^2,\ \big(s.e.(\hat{y}_{k+1}^*)\big)^2 + \ldots + \big(s.e.(\hat{y}_{n+1}^*)\big)^2\Big). \tag{7}$$

In any case, we have $s.e.(\hat{z}_{n+1}^*) = s.e.(1) = 0$. Of course, the actuary will plausibilize $s.e.(\hat{z}_k^*)$ similarly as $s.e.(\hat{U}_i)$ and, if necessary, manually adjust some of the resulting values.

Thus we finally obtain the following estimator for the mean squared error of prediction:

$$\begin{aligned} \hat{msep}(\hat{R}_i^{BF}) = {} & \hat{U}_i(\hat{s}_{n+2-i}^{2*} + \ldots + \hat{s}_{n+1}^{2*}) + \Big(\hat{U}_i^2 + (s.e.(\hat{U}_i))^2\Big)\Big(s.e.(\hat{z}_{n+1-i}^*)\Big)^2 \\ & + \big(s.e.(\hat{U}_i)\big)^2(1 - \hat{z}_{n+1-i}^*)^2. \end{aligned}$$

This is the formula one needs for risk based capital and premium loading calculations as well as for the construction of a confidence interval for R_i. In order to check the significance of differences between alternative reserve estimates or to construct a confidence interval for $E(U_i)$ one only needs the pure estimation error

$$\Big(s.e.(\hat{R}_i^{BF})\Big)^2 = \Big(\hat{U}_i^2 + (s.e.(\hat{U}_i))^2\Big)\Big(s.e.(\hat{z}_{n+1-i}^*)\Big)^2 + \big(s.e.(\hat{U}_i)\big)^2(1 - \hat{z}_{n+1-i}^*)^2.$$

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

A closer analysis of this formula shows that

$$s.e.(\hat{R}_i^{BF})/\hat{U}_i \approx s.e.(\hat{z}^*_{n+1-i}) \quad \text{for } \hat{z}^*_{n+1-i} \text{ close to 1,}$$
$$s.e.(\hat{R}_i^{BF})/\hat{U}_i \approx s.e.(\hat{U}_i)/\hat{U}_i \quad \text{for } \hat{z}^*_{n+1-i} \text{ close to 0,}$$

i.e. for the very green accident years, the uncertainty of the initial ultimate claims estimate is directly transferred to the reserve estimate.

For the overall reserve $R = R_1 + \ldots + R_n$ we have the unbiased estimate $\hat{R}^{BF} = \hat{R}_1^{BF} + \ldots + \hat{R}_n^{BF}$. Its mean squared error of prediction is $msep(\hat{R}^{BF}) = Var(\hat{R}^{BF}) + Var(R)$. For the process error we have $Var(R) = Var(R_1) + \ldots + Var(R_n)$ due to the independence of the accident years (BF1) and thus get the estimate

$$\hat{Var}(R) = \sum_{i=1}^{n} \hat{U}_i(\hat{s}^{2*}_{n+2-i} + \ldots + \hat{s}^{2*}_{n+1}).$$

The estimation error $Var(\hat{R}^{BF})$ is more involved because $\hat{R}_1^{BF}, \ldots, \hat{R}_n^{BF}$ are positively correlated via the common parameter estimates $\hat{y}_k^*$ (and in addition via the $\hat{U}_i$'s). We have

$$Var\left(\hat{R}^{BF}\right) = \sum_{i=1}^{n} Var\left(\hat{R}_i^{BF}\right) + 2\sum_{i<j} Cov\left(\hat{R}_i^{BF}, \hat{R}_j^{BF}\right).$$

For $Cov(\hat{R}_i^{BF}, \hat{R}_j^{BF}) = Cov\left(\hat{U}_i(1-\hat{z}^*_{n+1-i}), \hat{U}_j(1-\hat{z}^*_{n+1-j})\right)$ we use the general formula

$$\begin{aligned} Cov(XY, WZ) = {} & Cov(X,W)E(Y)E(Z) + Cov(X,W)Cov(Y,Z) \\ & + E(X)E(W)\ Cov(Y,Z) \end{aligned}$$

for random variables X, Y, W, Z where the sets $\{X, W\}$ and $\{Y, Z\}$ are independent. We omit the term in the middle which is of lower order and obtain

$$\begin{aligned} & Cov\left(\hat{U}_i\left(1-\hat{z}^*_{n+1-i}\right), \hat{U}_j\left(1-\hat{z}^*_{n+1-j}\right)\right) = \\ & \quad = \rho_{ij}^U \sqrt{Var\left(\hat{U}_i\right) Var\left(\hat{U}_j\right)}\ E\left(1-\hat{z}^*_{n+1-i}\right) E\left(1-\hat{z}^*_{n+1-j}\right) \\ & \quad + \rho_{ij}^z \sqrt{Var\left(\hat{z}^*_{n+1-i}\right) Var\left(\hat{z}^*_{n+1-j}\right)}\ E\left(\hat{U}_i\right) E\left(\hat{U}_j\right) \end{aligned}$$

with the correlation coefficients

$$\rho_{ij}^U = Cov\left(\hat{U}_i, \hat{U}_j\right) \Big/ \sqrt{Var\left(\hat{U}_i\right) Var\left(\hat{U}_j\right)},$$

$$\rho_{ij}^z = Cov\left(1-\hat{z}^*_{n+1-i}, 1-\hat{z}^*_{n+1-j}\right) \Big/ \sqrt{Var\left(\hat{z}^*_{n+1-i}\right) Var\left(\hat{z}^*_{n+1-j}\right)}.$$

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

Thus, we only have to estimate these correlation coefficients as we have estimates for all the other terms. If the actuary does not has the possibility to obtain data-based estimates for ρ_{ij}^{U} (e.g. from repricing) and ρ_{ij}^{z}, he may simply use one of the two estimates $\hat{\rho}_{ij}^{U}$ as given above (after (5)) and

$$\hat{\rho}_{ij}^{z} = \frac{\hat{z}_{n+1-j}^{*}\left(1-\hat{z}_{n+1-i}^{*}\right)}{\hat{z}_{n+1-i}^{*}\left(1-\hat{z}_{n+1-j}^{*}\right)} \text{ for } i<j \text{ and } \hat{z}_{1}^{*} \le \ldots \le \hat{z}_{n+1}^{*}.$$

The latter estimate stems from assuming a Dirichlet distribution (which is a generalization of the Beta distribution) for $\hat{y}_{1}^{*}, \ldots, \hat{y}_{n}^{*}$. Thus we finally get

$$\left(s.e.\left(\hat{R}^{BF}\right)\right)^{2} = \sum_{i=1}^{n}\left(s.e.\left(\hat{R}_{i}^{BF}\right)\right)^{2} + 2\sum_{i<j}\hat{C}ov\left(\hat{R}_{i}^{BF}, \hat{R}_{j}^{BF}\right)$$

with

$$\begin{aligned}\hat{C}ov(\hat{R}_{i}^{BF}, \hat{R}_{j}^{BF}) = {} & \hat{\rho}_{ij}^{U}\, s.e.(\hat{U}_{i})\, s.e.(\hat{U}_{j})\,(1-\hat{z}_{n+1-i}^{*})\,(1-\hat{z}_{n+1-j}^{*}) \\ & + \hat{\rho}_{ij}^{z}\, s.e.(\hat{z}_{n+1-i}^{*})\, s.e.(\hat{z}_{n+1-j}^{*})\,\hat{U}_{i}\hat{U}_{j}.\end{aligned}$$

6. Numerical example

The paid triangle of Exhibit A of Mack (2006) with $n = 13$ is used as example and we keep the initial ultimate claims estimates $\hat{U}_{i}$ (Exhibit C, column (I)) from there (see also Table 2 below, second column). In a first approach, we also keep the development pattern $\hat{z}_{k}^{*}$ (= b_{k} of Exhibit C, row (9), of Mack (2006)), see the row 'selected z' in the first block of Table 1 below. This pattern can also be obtained – except for rounding differences – from the raw estimates $\hat{\hat{y}}_{k}$ according to (3) by manually smoothing with the selections $\hat{y}_{8}^{*} = 8\%$, $\hat{y}_{9}^{*} = 5\%$, $\hat{y}_{10}^{*} = 3.7\%$, $\hat{y}_{11}^{*} = 2.1\%$, $\hat{y}_{12}^{*} = 1.5\%$, $\hat{y}_{13}^{*} = 1.4\%$ and a tail ratio $\hat{y}_{14}^{*} = 3.5\%$, see the second and third row of Table 1 below. In Mack (2006), this tail ratio was based on the calculation for the incurred data. From the pattern and the initial $\hat{U}_{i}$ the reserve estimates $\hat{R}_{i}^{BF1} = \hat{U}_{i}(\hat{y}_{n+2-i}^{*} + \ldots + \hat{y}_{n+1}^{*}) = \hat{U}_{i}(1-\hat{z}_{n+1-i}^{*})$ are calculated. These reserves, see the fourth column of Table 2, are thus the same as in Mack (2006) except for rounding differences.

For the prediction error, we first select $\hat{s}_{k}^{2*}$. For this purpose, we calculate the raw $\hat{\hat{s}}_{k}^{2}$ according to (4) and plot $ln(\hat{\hat{s}}_{k}^{2})$ against $|\hat{y}_{k}^{*}|$ for the decreasing part $k \ge 4$. We see that the plot looks reasonably smooth. Crucial cases are always $\hat{\hat{s}}_{n-1}^{2}$ and $\hat{\hat{s}}_{n-2}^{2}$ which rely on very few data. Here ($n = 13$), according to the plot, $\hat{\hat{s}}_{n-2}^{2} = 21.8$ and $\hat{\hat{s}}_{n-1}^{2} = 19.5$ seem to be rather small. Thus, we adjust these to $\hat{s}_{n-2}^{2*} = \hat{s}_{11}^{2*} = 30$, $\hat{s}_{n-1}^{2*} = \hat{s}_{12}^{2*} = 25$, leave $\hat{\hat{s}}_{k}^{2}$, $1 \le k \le 10$, as they are, i.e. $\hat{s}_{k}^{2*} = \hat{\hat{s}}_{k}^{2}$, and manually select from the plot the missing values $\hat{s}_{13}^{2*} = 20$ (over

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

Table 1. Parameter Estimates

devel. year k	1	2	3	4	5	6	7	8	9	10	11	12	13	14+
raw y	0.6%	4.4%	12.8%	19.0%	15.0%	10.6%	12.4%	6.5%	4.0%	1.2%	1.5%	1.9%	0.0%	
selected y	0.6%	4.5%	12.7%	19.0%	15.0%	10.6%	12.4%	**8.0%**	**5.0%**	**3.7%**	**2.1%**	**1.5%**	**1.4%**	**3.5%**
selected z	0.6%	5.1%	17.8%	36.8%	51.8%	62.4%	74.8%	82.8%	87.8%	91.5%	93.6%	95.1%	96.5%	100.0%
raw s^2	12.6	97.4	80.3	359.7	204.6	111.6	284.0	69.7	35.9	39.4	21.8	19.5		
selected s^2	12.6	97.4	80.3	359.7	204.6	111.6	284.0	69.7	35.9	39.4	**30.0**	**25.0**	**20.0**	**35.0**
selected s.e.(y)	0.27%	0.79%	0.75%	1.70%	1.39%	1.24%	2.50%	1.54%	1.30%	1.61%	1.63%	1.86%	2.49%	**1.50%**
s.e.(y) accumulated -->	0.27%	0.83%	1.12%	2.03%	2.46%	2.76%	3.72%	4.03%	4.23%	4.53%	4.81%	5.16%	5.73%	5.92%
s.e.(y) accumulated <--	5.91%	5.86%	5.81%	5.56%	5.38%	5.24%	4.60%	4.34%	4.14%	3.81%	3.45%	2.91%	1.50%	0.00%
s.e.(z)	0.27%	0.83%	1.12%	2.03%	2.46%	2.76%	3.72%	4.03%	4.14%	3.81%	3.45%	2.91%	1.50%	0.00%
Alternative Estimates														
selected y	0.7%	4.7%	13.0%	20.8%	15.6%	11.6%	8.7%	6.5%	4.8%	3.6%	2.7%	2.0%	1.5%	3.9%
selected z	0.7%	5.3%	18.3%	39.1%	54.7%	66.3%	75.0%	81.5%	86.3%	89.9%	92.6%	94.6%	96.1%	100.0%
selected s^2	12.6	98.1	80.9	406.2	209.2	122.4	388.9	57.4	34.3	37.2	28.0	**25.0**	**23.0**	**36.0**
selected s.e.(y)	0.27%	0.79%	0.75%	1.80%	1.41%	1.30%	2.92%	1.40%	1.27%	1.56%	1.57%	1.86%	2.67%	**1.93%**
s.e.(z)	0.27%	0.83%	1.12%	2.12%	2.55%	2.86%	4.09%	4.32%	4.38%	4.09%	3.78%	3.29%	1.93%	0.00%
Chain Ladder Estimates														
ATA factor f		7.830	3.440	2.052	1.400	1.204	1.203	1.088	1.049	**1.040**	**1.030**	**1.020**	**1.015**	**1.040**
cumulative pattern z	0.7%	5.3%	18.2%	37.5%	52.4%	63.1%	76.0%	82.6%	86.7%	90.2%	92.9%	94.7%	96.2%	100.0%
$sigma^2$		49623.1	17023.4	4633.7	893.8	456.2	766.1	87.0	45.2	37.6	32.0	22.2	**18.0**	**40.0**
s.e.(f)		2.315	0.479	0.143	0.048	0.035	0.052	0.020	0.016	0.017	0.018	0.019	0.025	**0.020**

Table 2. Reserves and Errors

Accident	Initial Ult.	Coeff. of Var.	Bornh/Ferg 1	Coeff. of Var.			Bornh/Ferg 2				Chain Ladder			
Year i	Claims U	c.v.(U)	Reserve R	Pred. Error	Pred. Error	Estim. Error	Proc.Error	Reserve R	Pred. Error	Estim. Error	Proc.Error	Reserve R	Pred. Error	Estim. Error	Proc.Error
1992	32299.9	**10.0%**	1130	1175	104%	500	1063	1247	1254	639	1078	1151	1218	576	1073
1993	40279.1	**10.0%**	1974	1907	97%	1193	1488	2162	2050	1351	1542	1992	1884	1182	1467
1994	40634.6	**10.0%**	2601	2302	89%	1432	1803	2999	2426	1573	1848	2699	2217	1374	1740
1995	39604.3	**10.0%**	3366	2602	77%	1554	2087	3992	2693	1678	2106	3653	2542	1499	2053
1996	58440.6	**10.0%**	7130	3892	55%	2532	2955	8002	3998	2695	2953	7655	3889	2508	2973
1997	81346.9	**10.0%**	13992	5280	38%	3578	3883	15075	5448	3842	3863	14132	5543	3766	4068
1998	163258.7	**10.0%**	41141	9789	24%	7362	6453	40833	10050	7854	6271	36855	10211	7632	6784
1999	268150.6	**10.0%**	100825	17364	17%	12528	12023	90335	17605	11877	12995	101544	25418	18233	17710
2000	331893.1	**10.0%**	159972	23225	15%	17985	14695	150375	23410	17274	15800	161368	34495	24501	24282
2001	193519.8	**10.0%**	122305	18186	15%	12854	12865	117793	18490	12483	13640	113303	27286	14459	23140
2002	169559.7	**10.0%**	139378	20097	14%	14068	14353	138541	20676	13986	15228	132795	41904	17226	38200
2003	157381.6	**10.0%**	149355	20703	14%	14993	14277	149014	21254	14960	15098	72180	54124	13339	52454
2004	156150.7	**10.0%**	155214	21413	14%	15527	14746	155128	21962	15519	15540	353492	265190	122286	235312

$\hat{y}^*_{13} = 1.4\%$) and $\hat{s}^{2*}_{14} = 35$ (over $\hat{y}^*_{14} = 3.5\%$). With these selections for $\hat{s}^{2*}_k$ we calculate s.e. $(\hat{y}^*_k)$ for $1 \le k \le n = 13$ according to (6) and find the resulting values and their coefficients of variation plausible. Then, we have to quantify our uncertainty in $\hat{y}^*_{14} = 3.5\%$ and select it to be $s.e.(\hat{y}^*_{14}) = 1.5\%$ assuming a 95%-range from 0.5% up to 6.5%. This fits well to the *s.e.* of $\hat{y}^*_{10}$ which is close to $\hat{y}^*_{14}$. Now we calculate $s.e.(\hat{z}^*_k)$ according to (7). All estimates and selections are shown in the first block of Table 1, where a bold number indicates a pure selection or a change from the raw estimate.

Finally, we have to select $s.e.(\hat{U}_i)$. In this example, we have an extreme premium cycle: The ultimate claims ratios $\hat{U}_i/v_i$ first decrease to 63%, then increase to 277%, then decrease again to 69% (see Mack (2006)). Thus, an application of equation (5) does not make sense. In Mack (2006), on-level premium factors r^*_i were estimated which bring all accident years on about the same claims ratio level. Then, the prior $\hat{U}_i$ were chosen to be

$$\hat{U}_i = v_i r^*_i \hat{m}^* (\hat{\tilde{y}}_1 + \ldots + \hat{\tilde{y}}_{n+1})$$

with $\hat{\tilde{y}}_k$ according to (3) and a certain constant factor $\hat{m}^*$. We can assume that the variability of $r^*_i \hat{m}^*$ is small compared to the one of $\hat{\tilde{y}}_1 + \ldots + \hat{\tilde{y}}_{n+1}$. Then we have

$$Var(\hat{U}_i) \approx (v_i r^*_i \hat{m}^*)^2 \, Var(\hat{\tilde{y}}_1 + \ldots + \hat{\tilde{y}}_{n+1}) = (v_i r^*_i \hat{m}^*)^2 \left(Var(\hat{\tilde{y}}_1) + \ldots + Var(\hat{\tilde{y}}_1) \right)$$

because the $\hat{\tilde{y}}_k$'s are fully independent due to BF1 as they do not have to add up to unity. As in the derivation of (6), we have

$$Var(\hat{\tilde{y}}_k) \approx s^2_k \Big/ \sum\nolimits_{j=1}^{n+1-k} \hat{U}_j,$$

$$\text{i.e. we take } \left(s.e.(\hat{\tilde{y}}_k) \right)^2 = \left(s.e.(\hat{y}^*_k) \right)^2 = \hat{s}^2_k \Big/ \sum\nolimits_{j=1}^{n+1-k} \hat{U}_j \, .$$

Finally, in order to get rid of the factor $v_i r^*_i \hat{m}^*$, we consider the coefficient of variation and obtain

$$c.v.\left(\hat{U}_i \right) = \frac{s.e.\left(\hat{U}_i \right)}{\hat{U}_i} \approx \frac{\sqrt{\left(s.e.\left(\hat{\tilde{y}}_1 \right) \right)^2 + \ldots + \left(s.e.\left(\hat{\tilde{y}}_{n+1} \right) \right)^2}}{\hat{\tilde{y}}_1 + \ldots + \hat{\tilde{y}}_{n+1}} = 6.7\%.$$

As we have ignored the variability of $r^*_i \hat{m}^*$ and have eliminated the full premium cycle (which probably would not have been achieved a priori), we deliberately increase this *c.v.* to c.v.$(\hat{U}_i) = 10\%$ for all accident years *i*. This is considered to be a rather high uncertainty for an estimate of $E(U_i)$ for classical insurance business because, e.g. for $\hat{U}_i / v_i = 90\%$, this corresponds to a wide 95% confidence range of (72%; 108%) – note that this is the range for $E(U_i)$ and not for U_i!

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

Note further that this approach only works for prior estimates $\hat{U}_i$ which were obtained in this specific way. It cannot be applied to estimates $\hat{U}_i$ obtained differently, e.g. via repricing, because each approach to $\hat{U}_i$ has its own uncertainties. Normally, $c.v.(\hat{U}_i)$ will not be the same for all accident years but will be lower for years with higher volume. In our example, we leave $c.v.(\hat{U}_i) = 10\%$ constant (see the third column of Table 2) assuming the varying volume has essentially been caused by writing varying shares of the same treaties. With these selections, we obtain the error estimates shown in the block 'Bornh/Ferg 1' of Table 2.

We also may apply the alternative estimation procedure described in Section 4: Then, we do not use the pattern of Mack (2006) but start with the original raw $\hat{\tilde{y}}_k$ according to (3) (see second row of Table 1) and select as last payment year $k_2 = 20$. Looking at the plot of $ln(|\hat{\tilde{y}}_k|)$ against k, we select $k_1 = 3$ and take an initial smoothing regression $ln(\hat{y}_k) = \alpha - \beta k$ with $\alpha = -0.03874$ and $\beta = 0.3632$ for $k > k_1$. With the resulting initial values for $\hat{y}_k$, initial values for $\hat{\tilde{s}}_1^2, \ldots, \hat{\tilde{s}}_{n-1}^2$ are calculated according to (4) which then are kept fixed during the following minimization of Q (without the term for $i = 1$ and $k = n = 13$). The minimum 79.98 is obtained at $\hat{y}_1^* = 0.65\%$, $\hat{y}_2^* = 4.7\%$, $\hat{y}_3^* = 13.0\%$, $\alpha = -0.4003$ and $\beta = 0.2920$ which leads to $\hat{y}_{14}^* = 3.9\%$ by adding up the extrapolated values for $\hat{y}_k$ from $k = 14$ to $k = 20$. For the other $\hat{y}_k^*$ (from the new regression) and the resulting $\hat{z}_k^*$ see the block 'Alternative Estimates' of Table 1. Then the corresponding new $\hat{\tilde{s}}_k^2$ are calculated according to (4) and the resulting values $ln(\hat{\tilde{s}}_k^2)$ are plotted against $|\hat{y}_k^*|$ for $k > k_1$. In view of this plot, we change $\hat{\tilde{s}}_{12}^2 = 18.7$ to $\hat{s}_{12}^{2*} = 25$ and select $\hat{s}_{13}^{2*} = 23$ and $\hat{s}_{14}^{2*} = 36$. Finally, we calculate $s.e.(\hat{y}_k^*)$ according to (6) and select $c.v.(\hat{y}_{14}^*) = 50\%$ which gives $s.e.(\hat{y}_{14}^*) = 1.93$. The resulting reserves $\hat{R}_i^{BF2}$, see Table 2, block 'Bornh/Ferg 2', are slightly higher than $\hat{R}_i^{BF1}$ for the old years and slightly lower for the new ones. The amounts (not the %ages) of the prediction error (using $c.v.(\hat{U}_i) = 10\%$ as before) are all a little bit higher. Using $\hat{\rho}_{ij}^U = 1/(1 + |i - j|)$, the overall reserve is $\hat{R}^{BF2} = 875.497$ with a prediction error of 72.940 consisting of an estimation error of 62.770 and a process error of 37.152.

As comparison we apply the Chain Ladder method, too. All parameters used are given in the last block of Table 1. We have replaced the last four raw age-to-age factors with 1.04, 1.03, 1.02, 1.015 and selected a tail factor of 1.04. The latter is in accordance with the tail ratio of 3.5%-3.9% used above. From the age-to-age factors we can derive the corresponding cumulative development pattern $\hat{z}_k$ as described in Section 2. The resulting values shown in Table 1 are close to the z-estimates of the two BF approaches but not identical. The implementation of the tail factor into the formulae for the prediction error has been done according to Mack (1999). The raw sigma-parameters (see Mack (1993) or Mack (1999)) have been kept and were supplemented with $\hat{\sigma}_n^2 = 18$ and $\hat{\sigma}_{n+1}^2 = 40$ on basis of a plot of $ln(\hat{\sigma}_k^2)$ against $ln(|\hat{f}_k - 1|)$. Finally, for the tail factor, $s.e.(\hat{f}_{n+1}) = 0.02$ was assumed, i.e. a 95%-range from 1.00 to 1.08. This

https://doi.org/10.2143/AST.38.1.2030404 Published online by Cambridge University Press

yields the results shown in the last block of Table 2. The CL reserves are close to the ones of BF except for the most recent years 2003 and 2004: In 2003, the CL reserve is about half of the BF reserve, whereas in 2004 the CL reserve is more than twice the BF reserve. This higher volatility is reflected in the markedly higher prediction errors for $i \geq 1999$, caused by a much higher process error. The CL and BF reserve estimates for 1992-2002 are not significantly different (i.e. not different by more than $2 \cdot s.e.(\hat{R}_i)$). But the reserves for 2003 are judged as being different by either method; the 2004 reserves are only different from the BF viewpoint whereas the CL estimation error is so large that the BF reserve is not judged to be different although it is less than 50% of the CL reserve. This is a good example for the fact that CL often cannot be reasonably applied in the standard way for new accident years in Excess business where almost nothing is paid in the first development year(s).

Acknowledgement

I am grateful to Dave Clark, Alois Gisler and Gerhard Quarg for helpful discussions in the conception phase of this paper.

References

Bornhuetter, R.L. and Ferguson, R.E. (1972) The Actuary and IBNR. *Proceedings of the CAS*, **LIX**, 181-195.

Mack, Th. (1993) Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, **23**, 213-225.

Mack, Th. (1999) The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor. *ASTIN Bulletin*, **29**, 361-366.

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

Taylor, G.C. (2002) Written discussion of the paper "Stochastic Claims Reserving in General Insurance" by P.D. England and R.J.Verrall, *British Actuarial Journal*, **8**, Part III (2002), 540-542.

Thomas Mack
Retired Chief Actuary Non-Life
Munich, Germany
E-Mail: tmack@gmx.net