# BOOTSTRAPPING INDIVIDUAL CLAIM HISTORIES

BY

STIG ROSENLUND


## ABSTRACT

The bootstrap method BICH is given for estimating mean square prediction errors and predictive distributions of non-life claim reserves under weak conditions. The dates of claim occurrence, reporting and finalization and the payment dates and amounts of individual finalized historic claims form a claim set from which samples with replacement are drawn. We assume that all claims are independent and that the historic claims are distributed as the object claims, possibly after inflation adjustment and segmentation on a background variable, whose distribution could have changed over time due to portfolio change. Also we introduce the new reserving function RDC, using all these dates and payments for reserve predictions. We study three reserving functions: chain ladder, the Schnieper (1991) method and RDC. Checks with simulated cases obeying the assumptions of Mack (1999) for chain ladder and Liu and Verrall (2009) for Schnieper's method, respectively, confirm the validity of our method. BICH is used to compare the three reserving functions, of which RDC is found overall best in simulated cases.



## KEYWORDS

BICH, Bootstrap, Claim reserve, RDC, Stochastic reserving.


## 1. INTRODUCTION

For claim reserve MSEP (*Mean Square Error* of *Prediction*) calculations, the analytic method of Mack (1999) is available for chain ladder under certain mean, variance and independence assumption. Liu and Verrall (2009) give assumptions and algorithms enabling analytic MSEP computations for the Schnieper (1991) reserves. See also Section 10.2 of Wüthrich and Merz (2008).

Other stochastic reserving methods use bootstrapping in the form of drawing random upper triangles, $\nabla_r$ of development history and lower triangles, $\Delta_r$ of future development (outcomes in the bootstrap world), by sampling with replacement from standardized residuals within a development triangle. On each such random triangle a reserving method is applied to obtain a reserve prediction. This yields MSEP estimates transferable to the real world. The model of

*Astin Bulletin* 42(1), 291-324. doi: 10.2143/AST.42.1.2160744 
Pdf by the author. Page breaks as in the printed article. See last page for errors of print in the latter.

this method is usually that the residuals are the deviations from estimated mean values in a GLM (Generalized Linear Model) that assumes that the increments $X$ obey $\text{Var}[X] = \phi \text{E}[X]^p$, with $p$ usually assumed to be either 1 (Poisson) or 2 (Gamma). See Björkwall et al. (2009) for an overview, contributions and further references.

Norberg (1993) and Norberg (1999) introduced the individual claim loss model as marked Poisson processes. The model is studied in Larsen (2007). Several Poisson and other parametric models are assumed and their parameters estimated. A bootstrap procedure is sketched, but not implemented in practice, as step 2 p. 131.

Zhao and Zhou (2010) used semi-competing risks copula and semi-survival copula models to fit the dependence structure of the claim occurrence times with reporting delays in the individual claim loss model of Larsen (2007) and Taylor et al. (2008), in order to study IBNR reserves. Also here a Poisson arrival process is assumed.

An overview of the area is given by Wüthrich and Merz (2008).

In this paper we describe how to use sampling with replacement from a set of detailed complete claim histories for claims that are finalized. Applying a reserving method to each such sample, using only payments up to relevant development periods, we get predictions in the bootstrap world that can be compared to the known outcomes of these finalized claims. The model is essentially that claims are IID and that the historic claims are distributed as the object claims, apart from inflation. Here 'object' refers to the set of partly non-finalized claims whose future we wish to predict. This allows variance estimates and also estimates of the full predictive distributions for the object claim reserves. The method is computer-intensive, since a large set of claim histories is read into memory and used repeatedly for sampling. In return it does not need much mathematical-statistical theory. We call it

**BICH** = **Bootstrapping Individual Claim Histories**

BICH can rank reserving methods by their MSEPs under weak conditions. A reserve model is one thing and a reserving method another thing. A method derived under a model can perform well even if the model is not true. BICH measures this.

We also give a new reserving method using all claim numbers of and payments on both open and settled individual claims without any distributional assumptions. It is a generalization of the PPCF (expected *Payments Per Claim Finalized*) method. See Fisher and Lange (1973) and Sawkins (1979). The parameters of the claim distribution are broken down in many small details while conditioning on observable variables with many combinations. It is described in Appendix A. We call it

**RDC** = **Reserve** by **Detailed Conditioning**

The reserving methods studied with BICH in the sequel are the chain ladder, the Schnieper (1991) method and RDC. In comparing reserving methods we can choose the best one. Also, variable parameters in RDC can be calibrated to give the smallest MSEP. Overall we find that RDC is best for the situations studied.

BICH and RDC are intended for cases where many claims are reported in the first development period, say at least a couple of hundreds. Then it is normally better to not use insurance exposures to generate claims in bootstrapping or to compute reserves by using e.g. a Poisson process, since claim frequencies mostly oscillate in a not completely predictable manner. Thus we can avoid stochastic process theory. Our study concerns IBNR (Incurred But Not Reported) and RBNS (Reported But Not Settled) claims. We do not study the UPR (Unearned Premium Reserve) of covered but not yet incurred claims, which would need insurance exposure.

Since the reserve of a business line is, in our model, a sum of the reserves of several independent claims, a CLT (central limit theorem) can be invoked unless a few large claims dominate the overall development. Hence the variance estimates are useful. They can e.g. be used to compute the variance for the sum of reserves from several business lines. Moreover, the full predictive distributions are given in the form of eleven important percentiles in the author's program.

As observed in Larsen (2007) and Zhao and Zhou (2010), distributions for payment delays and sizes may change over time, making the use of covariates desirable. We take account of such heterogeneity by segmentation on a background variable.

The word payment in the sequel can be replaced by change of incurred (payment sum plus claims-handler reserve). So payments can be both positive and negative.

Large claims, as judged at reporting, might have to be excluded from the analysis.

The organization of the paper is as follows. Section 2 introduces the claim info that is supposed to be available, assumptions and perspectives on these. The bootstrap procedure is described in Section 3. Section 4 deals with application guidelines. In Section 5 the segmentation mechanism is described. Section 6 gives estimates and tests. In Section 7 the necessary data are described and four numerical examples are given, where Section 7.7 gives results of benchmark tests to validate BICH and Section 7.8 compares reserving methods. After a conclusion in Section 8, the RDC method is described in Appendix A.

Free program for BICH, RDC, GLM etc.: `www.stigrosenlund.se/rapp.htm`

## 2. Model assumptions

We consider claim periods $i \in \{1, \ldots, n\}$, for which we want to predict future payments, and development periods $j \in \{1, \ldots, n\}$. The present time is at the end of period $n$ or in other words at the beginning of claim period $n+1$.

In Section 5 we describe a mechanism for segmentation by a background variable. For simplicity this mechanism will not appear in the notation of sections other than Section 5 and Section 7.6, where segmentation is applied in Example 4.

First we formulate historic and object claims, not including segmentation.

### 2.1. Historic claims for bootstrap

Let $Z = \{Z_1, Z_2, \ldots, Z_K\}$ be a set of $K$ historic finalized claims which occurred more than $n+s-1$ time periods ago, i.e. in periods $-s+1, -s, -s-1, \ldots$ . Their payments are thus known up to and including development period $n + s$. Here $s \in \{0, 1, \ldots\}$ is the length of a maximal tail time after $n$. A claim is here a set

$$Z_r = \{W(r), F(r), Y(r, 1), \ldots, Y(r, F(r))\}, \tag{2.1}$$

where $W(r)$ is the period index of the customer claim reporting date, $F(r)$ the period index of the claim finalization date, and $Y(r, j)$ the payment sum for development period index $j$. Period index is defined so that the claim occurrence date falls in the period with index 1. The notation $W(r)$ is chosen since this variable can be called waiting-for-report period or reporting delay, albeit with smallest value 1 instead of 0. Henceforth period will mean period index, for shortness.

A claim reported at occurrence thus has $W(r) = 1$. If it is reported one period later, e.g. in the following month if month is the time unit, then $W(r) = 2$, etc. The payment sum made in the claim occurrence period has $j = 1$. The claim occurrence dates of these claims do not appear in this notation – only the report, finalization and development periods. The claims in $Z$ will be used for bootstrapping.

### 2.2. Object claims

Let $T^i = \{T^i(1), T^i(2), \ldots\}$ be a set of finalized and non-finalized claims from the claim occurrence period $i \in \{1, \ldots, n\}$. Claims in $T^i$ are defined by the common occurrence period $i$, the report and finalization periods, and a sequence of payments per development period. As in (2.1), with $i$ added as superscript. Thus, using the form of (2.1),

$$T^i(k) = \{W^i(k), F^i(k), Y^i(k, 1), \ldots, Y^i(k, F^i(k))\} \tag{2.2}$$

is the $k$:th claim in $T^i$. The whole collection of sets is $T = \{T^1, \ldots, T^n\}$. The payments are known only up to and including development period $n - i + 1$ for claims in claim period $i$. Our objective is, for each $i$, to predict the sum of remaining payments and estimate the MSEP from the real remaining sum for these claims. The claims for $i$ that are finalized could be part of $Z$. If the business is special and so new that data are available only back to the first finalized period, then only the claims of $T^1$ can be used for bootstrap. Let $(\Omega, \mathcal{F}, \mathrm{P})$ denote the probability space. Define the known history of claim development in $T$ up to and including calendar period $n$ as a $\sigma$-algebra $\mathcal{G} \subset \sigma\{T\}$. Here $\sigma\{T\}$ denotes the $\sigma$-algebra induced by $T$. Let $F^{i\prime}(k) = F^i(k)$ if $F^i(k) \leq n{+}1{-}i$, otherwise 0. Then $F^{i\prime}(k)$ is known even if $F^i(k)$ is not, and thus set

$$\mathcal{G} = \sigma\{W^i(k), F^{i\prime}(k), Y^i(k,1), \ldots, Y^i(k, n{-}i{+}1);\ 1{\leq}i{\leq}n;\ W^i(k){\leq}n{-}i{+}1\}.$$

Consider a $\mathcal{G}$-measurable reserving function

$$\widehat{R}_i = \text{ reserve-ex-ante for claim period } i \in \{i = 1, \ldots, n\}, \tag{2.3}$$

which is a prediction computed before the actual remaining payment sum is known.

BICH is in the program currently equipped with these $\widehat{R}_i$ functions: chain ladder (possibly combined with an exponential tail predicting payments at $n{+}1, \ldots, n{+}s$), the Schnieper (1991) method and RDC.

For claim set $T^i$ we define these further random variables.

$$M_i = \text{ number of reported claims now, i.e. with } W^i(k) \leq n - i + 1 \tag{2.4}$$

$$N_i = \text{ total number of claims in } T^i, \text{ not known at the end of claim period } n \tag{2.5}$$

$$Y_{ij} = \sum_{k=1}^{N_i} Y^i(k,j) = \text{payment sum for } i \text{ and } j \text{ over all claims } T^i(k) \tag{2.6}$$

$$H_i = \sum_{j=1}^{n-i+1} Y_{ij} = \text{sum of known payments per claim period} \tag{2.7}$$

$$\begin{aligned} R_i &= \sum_{j=n-i+2}^{n+s} Y_{ij} = \text{reserve ex-post} = \text{unknown remaining payment sum} \\ &= \text{ outstanding loss liabilities, see (1.17) in Wüthrich and Merz (2008).} \end{aligned} \tag{2.8}$$

$$Q_i = R_i/\widehat{R}_i = \text{ratio of reserve-ex-post to reserve-ex-ante} \tag{2.9}$$

$$\tau_i^2 = \mathrm{E}[(R_i - \widehat{R}_i)^2 \mid \mathcal{G}] = \text{MSEP of period } i \tag{2.10}$$

The MSEP $\tau_i^2$ is the main object of our study.

### 2.3. BICH assumptions and a hypothesis

In Section 3 we describe bootstrap images $\mathcal{G}^{(\nu)}$ of $\mathcal{G}$ and $\tau_i^{(\nu)2}$ of $\tau_i^2$ using $Z$, where $\nu$ is the bootstrap repetition of sample index. Assumption A5 deals with these images.

**A1.** All claims are finalized after development period $n + s$, i.e. $Y^i(k,j) = 0$ for $j > n + s$. For methods Schnieper and RDC we require that $s = 0$.

**A2.** $Z_1, Z_2, \ldots, Z_K$ are IID random vectors with variable size of the form (2.1).

**A3.** $T^i(1), T^i(2), \ldots, T^i(N_i)$ are IID random vectors of the same form (2.2).

**A4.** $T^i(1)$ is distributed as $Z_1$ after multiplying $T^i(1)$:s payments by $c_i$, for constants $c_i > 0$. For method RDC we require that $c_i \equiv c$ for some $c$.

**A5.** There are subsets $G^{(\nu)} \in \mathcal{G}^{(\nu)}$ with $\mathrm{P}(G^{(\nu)}) > 0$ such that $c_i^2 \mathrm{E}[\tau_i^{(\nu)2} \,|\, G^{(\nu)}] = \tau_i^2$.

A5 states that all particular possible bootstrap upper triangle outcomes in $G^{(\nu)}$, as far as can be judged beforehand, give the same information on the MSEPs of reserves-ex-ante from actual reserves-ex-post. The sets $G^{(\nu)}$ would normally in applications make these upper triangle outcomes similar to the object triangle, after adjustment by the factor $c_i$. In practice this means that the bootstrapped claims shall not differ too much from object claims in the properties of the latter that are known and that can be judged to influence the MSEPs.

We need many claims reported in the first development period for A5 to hold, although this might not be sufficient. We discuss further requirements in Section 4.

Given A5 we can estimate the MSEPs. See Section 6.1, expressions (6.1)–(6.4).

We want the reserve-ex-post to have the reserve-ex-ante as expected value, i.e. $\widehat{R}_i$ should reflect the best-estimate values for $R_i$. We formulate it as a hypothesis, not as an assumption. Its truth depends on the situation and the reserving function used. For chain ladder, the Mack (1999) condition CL1 implies its truth. We describe how to test the hypothesis in Sections 4 and 6.

**H0:** $\mathrm{E}[R_i \mid \mathcal{G}] = \widehat{R}_i$ or equivalently $\mathrm{E}[Q_i \mid \mathcal{G}] = 1$ (application dependent truth)

#### 2.3.1. *Perspectives on assumptions*

BICH assumes IID claims, after inflation adjustment and segmentation. This is a natural and weak assumption for the intended use of bootstrapping individual claims, with many claims reported initially and no insurance exposure use. BICH predictions for claim period $i$ are in the price level of $i$. Predictions in the price level at the end of $n$ are obtained by multiplying the reserves and MSEP square roots by $\hat{c}_n/\hat{c}_i$ for estimates $\hat{c}_i$. See expression (6.2). If the price levels of the times of future payments are desired in the predictions (not implemented), then further assumptions on future inflation are needed.

Other authors mentioned in the introduction place various assumptions on aggregated triangle data. Formally the BICH assumptions are not weaker than the latter ones, or the other way around, since they are set up in quite different frameworks. One may ask whether a set of assumptions for aggregated triangle data found in the literature for a method M can be satisfied by a construction of IID claims, such that MSEP estimates computed by BICH are the same as those computed by M. If that is the case, then in a loose sense the BICH assumptions are weaker. As is seen in Example 1, this holds for the CL1, CL2, CL3 conditions of Mack (1999). Similarly, construction of IID claims such that triangle increments follow a specified GLM, as described in Björkwall et al. (2009), is also possible.

The Liu and Verrall (2009) assumptions for the Schnieper (1991) model are slightly different. Assumption 4 of the former states uncorrelatedness between development periods. In particular increments from new claim reportings are assumed to be uncorrelated. The BICH assumptions are not weak enough to cover this, since the total number of claims given the number reported now is negatively binomial with the latter as parameter. See Sections 3 and 7.4 below. BICH admits any kind of dependence between development periods for individual claims. But it cannot admit independence between development periods for the totality of all claims, unless all claims are reported in the first development period, i.e. unless $\mathrm{P}(W = 1) = 1$, where $W = W(r)$ as defined in (2.1).

However, in the Schnieper (1991) and Liu and Verrall (2009) model all distributions are conditional on the history, where the first development period $j = 1$ is known for all claim periods. So we can regard the first increment as non-stochastic. Thus the BICH and Liu and Verrall (2009) conditions can coexist, if all claims are reported in the claim occurrence period or the next period. In other words, we can employ BICH under the Liu and Verrall (2009) conditions if $\mathrm{P}(W \leq 2) = 1$. For e.g. most quarterly consumer insurance data, this covers the reporting delay.

The Larsen (2007) algorithm would resemble the one presented here, but since it would augment uncertainty from outcomes of experiments with parameters for Gamma, Pareto etc., with parameter estimation uncertainty, it would use stronger assumptions than we use here.

## 3. The BICH bootstrap procedure

We describe how to make a bootstrap image of the claims of $T$ using $Z$. The segmentation described in Section 5 does not appear in the notation of this section, but it is to be understood that it can be used. First we present an overview in five steps S1–S5, which is outlined in more detail below.

- S1. Identify a set $Z$ of finalized claims distributed as the object claims after inflation adjustment. Also identify possible suitable subsets $G^{(1)}$ as in Sections 4.1 and 4.2.
- S2. Draw a random sample, numbered $\nu$, with replacement from $Z$ such that the number of bootstrapped claims reported 'now' per claim period $i$ equals the corresponding number of object claims $M_i$.
- S3. Provided $G^{(\nu)}$ occurred, compute the bootstrap images of (2.3), (2.5)–(2.9) and add to sums that shall be used for estimates.
- S4. Go back to S2 until S2 and S3 have been repeated $B$ times with $G^{(\nu)}$ occurring. Set $B = 2,000$ for example.
- S5. Compute MSEP estimates, etc., and statistics for test of H0 in Section 6.

We make $B_0$ IID repetitions of this image. All performed repetitions are numbered consecutively with index $\nu$. By Assumption A5 we might want to use only some of these repetitions by conditioning on some sets $G^{(\nu)}$ of outcomes.

The number of used repetitions where $G^{(\nu)}$ occurred, which will be fixed, is denoted $B$. (Then $B_0$ is a random variable, having a negative binomial distribution $\mathrm{NB}(B,p)$ with moment generating function $\mathrm{E}[e^{tB_0}] = (pe^t/[1-(1-p)e^t])^B$, if the probability is $p$ in every repetition that the outcome will be used. Recall that the number of trials in IID repetitions until a given number of successes is obtained is a negative binomial variable.) The used repetitions are found in the subsequence $\nu_1, \nu_2, \ldots, \nu_B$. We index it by $t$, writing $\nu_t$ $(t = 1, \ldots, B)$. Since the repetitions are IID, the subsets $G^{(\nu)}$ should be defined in the same way, with all $\mathrm{P}(G^{(\nu)})$ equal.

In each repetition we draw, with replacement and separately for each $i$, successive claims from $Z$ until exactly $M_i$ of them have report period $W(r) \leq n-i+1$, i.e. have been reported 'now' in the bootstrap world. In the sequel, when writing 'now' we refer to this report condition in the bootstrap world. Writing now without quotes concerning reporting we refer to the object claims. Define

$$N_i^{(\nu)} = \text{total number drawn, of which } M_i \text{ are reported 'now'}. \tag{3.1}$$

By Assumption A4, it is immediate that the conditional distribution of $N_i^{(\nu)} \mid M_i$ is a bootstrap approximation of the conditional distribution of $N_i \mid M_i$ in the real world, since $N_i^{(\nu)}$ as well as $N_i$ is a negatively binomial random number of claims, realized from an infinite sequence of IID random entities, until $M_i$ of them have satisfied the reported 'now' condition.

As prediction we can use

$$\hat{N}_i = \frac{1}{B} \sum_{t=1}^{B} N_i^{(\nu_t)} \tag{3.2}$$

Now recall that $K$ is the number of claims that can be used for bootstrap. If we sample at random a claim from $\{Z_1, Z_2, \ldots, Z_K\}$ with probability $1/K$, then it can represent any object claim. Let for $i \in \{1, \ldots, n\}$ and $\nu \in \{1, 2, \ldots, B_0\}$

$$U_{i,1}^{(\nu)}, \ldots, U_{i,N_i^{(\nu)}}^{(\nu)} \tag{3.3}$$

be IID integer valued and uniformly distributed on $\{1, \ldots, K\}$. Namely,

$$\mathrm{P}(U_{1,1}^{(1)} = r) = 1/K, \; r = 1, \ldots, K. \tag{3.4}$$

We use the finalized claims

$$Z_{U_{i,1}^{(\nu)}}, \ldots, Z_{U_{i,N_i^{(\nu)}}^{(\nu)}}$$

as an image of the claim set $T^i$ of claims that occurred in claim period $i$.

In other words, we draw with replacement $\sum_{\nu=1}^{B_0} \sum_{i=1}^{n} N_i^{(\nu)}$ claims from $Z$ with equal probability for each $r = 1, \ldots, K$ in each drawing. For the examples of Section 7 this is about 5,000 million claims.

We will use the superindex $(\nu)$ for bootstrap variables, with (1) for the first variable of the repetitions as a representative of the sequence of IID bootstrap variables. Averages will have superindex (-).

A bootstrap image of the $Y^i(k,j)$ of $T^i$ is $Y(U_{i,k}^{(\nu)},j)$ = payment sum in development period $j$ of the $k$:th claim drawn from $Z$ to represent claim period $i$. See (2.1). Let $F'(i,r) = F(r)$ if $F(r) \leq n-i+1$, otherwise 0. The bootstrap image of $\mathcal{G}$ is

$$\mathcal{G}^{(\nu)} =$$

$$\sigma\{W(U_{i,k}^{(\nu)}), F'(i,U_{i,k}^{(\nu)}), Y(U_{i,k}^{(\nu)},1),\ldots,Y(U_{i,k}^{(\nu)},n-i+1);\ 1\leq i\leq n;\ W(U_{i,k}^{(\nu)})\leq n-i+1\}.$$

The reserve-ex-ante of this image is obtained by using the same function on $\mathcal{G}^{(\nu)}$ as was used on $\mathcal{G}$ to get $\widehat{R}_i$ in (2.3).

$$\widehat{R}_i^{(\nu)} = \text{ bootstrap reserve-ex-ante for claim period } i,\ i=1,\ldots,n \tag{3.5}$$

Bootstrap images of the random variables of (2.6) – (2.10) are

$$Y_{ij}^{(\nu)} = \sum_{k=1}^{N_i^{(\nu)}} Y(U_{i,k}^{(\nu)},j) \tag{3.6}$$

$$H_i^{(\nu)} = \sum_{j=1}^{n-i+1} Y_{ij}^{(\nu)} \tag{3.7}$$

$$R_i^{(\nu)} = \sum_{j=n-i+2}^{n+s} Y_{ij}^{(\nu)} \tag{3.8}$$

$$Q_i^{(\nu)} = R_i^{(\nu)}/\widehat{R}_i^{(\nu)} \tag{3.9}$$

$$\tau_i^{(\nu)2} = \mathrm{E}[(R_i^{(\nu)} - \widehat{R}_i^{(\nu)})^2 \mid \mathcal{G}^{(\nu)}], \tag{3.10}$$

where the last one is unobservable and determined by which events in $\mathcal{G}^{(\nu)}$ occur.

In Section 6 we will explain how to use this mean square deviation measure together with an adjustment factor $c_i$ reflecting possible inflation.

The approach is to generate prescribed numbers of claims reported 'now'. It works if those numbers are large enough. This eliminates the need for insurance exposures and too much modeling. For example, claim sizes could depend on reporting delays, but we can ignore this possible dependence. Say that 90 percent of claims finally below 10,000 EUR, but only 10 percent of claims finally at least 10,000 EUR, are reported immediately at occurrence. Regardless, the bootstrapped sample for any claim period will, under our assumptions, contain small and large claims proportionally to their frequencies in the object claim data set.

## 4. Application guidelines

We do not here distinguish between estimation error and prediction error. Since $Z$ forms a discrete empirical distribution that is an estimate of an underlying one, there is however an estimation error, which is hard to quantify. So A4 and A5 cannot be exactly true. For $K$ sufficiently large the empirical distribution is a good approximation, see Section 4.3.

In applications where $Z$ and $T$ are from the same line of business and where we have access to the full history and where we judge the claim and payment processes to be sufficiently time-homogeneous, we want to use all claims back to the earliest finalized claim period in order to make the best possible prediction, thus setting $s = 0$. Choosing $s > 0$ by not going back that far, BICH can however shed light on the performance of tail prediction methods.

The real $n+s$ might be large, but payments after some moderate $n+s$ negligible, e.g. less than one percent of the total claim cost.

Hypothesis H0 can be tested in the bootstrap world by computing the empirical distribution of $Q_i^{(\nu)} = R_i^{(\nu)}/\widehat{R}_i^{(\nu)}$ conditional on $G^{(\nu)}$. Let $Q_i^{(-)}$ be the mean of this distribution and $0.01s(Q_i^{(-)})$ its standard error. See (6.13) and (6.14) in Section 6.1 below. If the confidence interval $[Q_i^{(-)}-1.96\times0.01s(Q_i^{(-)}), Q_i^{(-)}+1.96\times0.01s(Q_i^{(-)})]$ contains 1, then H0 can be accepted at the 95 percent level, if $B$ is large enough for CLT use. Normally $B = 2,000$ suffices.

H0 states that the reserve prediction is the mean of the finally realized reserve. If true, then

$$\sqrt{\mathrm{Var}[R_i \mid \mathcal{G}]} = \tau_i \tag{4.1}$$

The purpose of $c_i$ in Assumption A4 is to adjust for inflation or deflation. The adjustment by a factor dependent only on claim-period is a simplification of reality, since the development period also could influence the degree of inflation in the payments. For the present purpose the simplified model normally suffices. If not, the claim payments can be adjusted before the BICH algorithm is run.

### 4.1. Subsets with reserves close to mean reserves

A class of subsets $G^{(\nu)}$ that is implemented in the program for BICH is the following: in a first run $\widehat{R}_i^{(-)}$, the empirical bootstrap mean reserves-ex-ante per claim period, are computed to sufficient closeness to their expectations, i.e. after sufficient convergence. See (6.7) below. In a second run two factors $b_1 < b_2$ are used to bound the outcomes. Namely, so that all outcomes $\nu$ are thrown away that do not have every $\widehat{R}_i^{(\nu)}$ $(i \in \{1,\ldots,n\})$, the bootstrap reserve-ex-ante per claim period, within the interval with endpoints $b_1\widehat{R}_i^{(-)}$ and $b_2\widehat{R}_i^{(-)}$. For example we can take $b_1 = 0.80$ and $b_2 = 1.25$. If $b_1\widehat{R}_i^{(-)}$ is positive then it is the left endpoint, and if it is negative it is the right endpoint. If the probability for an outcome to be used and not thrown away is $p$ then we have to make about $B_0 = B/p$ repetitions to obtain $B$ useful samples.

The randomness in this bootstrap can however, decrease too much with too narrow sets $G^{(\nu)}$, so this needs to be done carefully.

### 4.2. Subsets with fixed numbers of claims per known reporting period

Another kind of subsets $G^{(\nu)}$ is defined by letting all bootstrapped claim reporting numbers per known development period be equal to the corresponding numbers of the object claim set. Namely, letting $A_{iw}$ be the number of claims reported in development period $w \geq 1$, the corresponding bootstrapped numbers $A_{iw}^{(\nu)}$ in repetition $\nu$ should be these for $w \leq n - i + 1$. This is sometimes a reasonable bounding of the bootstrap outcomes to make them more like the object outcome. In other cases this bounding can decrease the randomness of bootstrap too much.

Since the waiting time between such bootstrap outcomes will be long, the actual procedure is as follows. Claims with $W = w \leq n-i+1$ are drawn until $A_{iw}$ have been obtained. After that they are rejected. Claims with $W > n-i+1$ are drawn until the number drawn with $W \leq n-i+1$, including rejected ones, is $M_i$. This creates a bootstrap distribution equal to the one obtained by using $G^{(\nu)}$ as literally defined.

### 4.3. When will BICH work?

When will BICH work, assuming A1 – A4 and that we try to determine sets $G^{(\nu)}$ for A5 as best as possible? That is, when can we make A5 hold? We have stated that, with at least a couple of hundred claims reported initially BICH outperforms using insurance exposures. This would not, however, always be sufficient for A5.

The number $K$ of finalized claims in $Z$ should be sufficiently large to adequately represent the claim distribution, considering its dispersion.

Take the extreme case of all payments being 1000 EUR and always made in development periods $W, W+1, W+2$, where $W$ is the development period of reporting. Then $K = 200$ would be sufficient.

On the other hand, suppose there are large claims above the 0.1 percentile, which constitute half of expected claim cost (= sum of payments) and are highly variable. If the payment made in the reporting period says little about the final claim cost, then $K$ might need to be at least 1 million. If the representativeness of $Z$ for large claims is in doubt, then a table of percentiles for $Z$ should be made and compared to percentile tables for several other business lines in addition to the one under consideration, or for several competing companies, such that this total business can be regarded as having about the same tail distribution as the sets $T^i$ of object claims. If the compared tail percentile tables are about the same, then $Z$ should do.

## 5. BICH segmentation

The finalized claims in $Z$ could have a different distribution of some background variables than the sets $T^i$ of object claims. For example, the proportion of historic claims coming from business line 7 might be 0.03 in the finalized claims, while it is 0.08 for claim period 12 in the set of object claims.

Another cause of such discrepancy can be seasonal variations. Say that claims occurring in August are on the average more expensive and their payments more drawn out than in other months. If the $Z_r$ are evenly distributed over the year, then bootstrapping from all these for the object claim period 8 (August) could be misleading if the time unit of the analysis is month.

Therefore BICH has an option for segmentation by a background variable, which possibly is a combination of many such variables. This works so that, separately per claim period, the proportions of the background variable values in the claims of the bootstrapped sample that have been reported 'now' will be equal to the proportions in the set of object claims reported now.

The model is then that the assumptions of Section 2 hold separately per segment. This is the non-parametric way to model dependence on background variables. It works if there are sufficiently many claims in each discrete segment. This is analogous to stratified survey sampling aiming to make estimates and forecasts more accurate.

However, there is a caveat in that we cannot have too few historic claims per segment. Also sufficiently many, say at least one hundred, object claims should be reported in the first development period. In the bootstrap, each segment is a discrete distribution of claims with finitely many points in the space of $Z$. The fewer points in this discrete distribution, the less similarity to the real corresponding distribution. The latter would normally be best represented by a distribution on an infinite set of points. In the extreme case that each segment has just one historic claim all randomness disappears. The percentile analysis recommended in Section 4.3 is appropriate for each segment separately.

The generalization of the procedure of Section 3 is the following:

$M_{vi}$ = number of claims in $T^i$ reported in segment $v, v \in \{1, \ldots, s_0\}$,
where $\sum_{v=1}^{s_0} M_{vi} = M_i$

$Z_{vr}$ $(v = 1, \ldots, s_0;\ r = 1, \ldots, K_v)$ are the claims of $Z$ in segment $v$,
where $K_v \geq 1$ and $\sum_{v=1}^{s_0} K_v \leq K$. Segment variable values in the historic claims not found in the object claims cannot be used. If there are such values, then $\sum_{v=1}^{s_0} K_v < K$.

The segmentation bootstrap is to draw $M_{vi}$ reported claims with replacement from the $Z_{vr}$. Totals $N_{vi}^{(\nu)} \geq M_{vi}$, including claims not reported 'now', are then used to compute reserves separately per segment. These are added and mean square deviation estimates are computed, as described in Section 6.1. The segmentation mechanism has not appeared in the preceding sections and will not appear in the following sections other than Section 7.6, but is to be understood to be available.

## 6. Inference from BICH bootstrap

### 6.1. Estimates

The idea is now to estimate $c_i^2 \mathrm{E}[\tau_i^{(1)2} \mid G^{(1)}]$, equal to $\tau_i^2$ under our assumptions, by sample statistics of the sample of $B$ repetitions $\nu_1, \ldots, \nu_B$ where $G^{(\nu)}$ occur.

We transfer the bootstrap mean square error estimates to MSEP estimates for the object reserve predictions $\widehat{R}_i$. Define the estimate

$$\widetilde{\tau}_i = \sqrt{\frac{1}{B}\sum_{t=1}^{B}\left(R_i^{(\nu_t)} - \widehat{R}_i^{(\nu_t)}\right)^2}. \tag{6.1}$$

To make the variance of $\widetilde{\tau}_i$ sufficiently small is only a matter of making $B$ sufficiently large and computations time.

Estimate the conditional prediction error of $R_i \mid \mathcal{G}$ by first making an inflation estimate equal to the ratio of payments, see (6.6) below, namely for chain ladder and Schnieper

$$\hat{c}_i = H_i / H_i^{(-)} \tag{6.2}$$

and, in accordance with Assumption A4, for RDC

$$\hat{c}_i = \sum_{r=1}^{n} H_r \Big/ \sum_{r=1}^{n} H_r^{(-)} \tag{6.3}$$

and then a transfer of the bootstrap estimate to the real world by

$$\widehat{\tau}_i = \hat{c}_i \widetilde{\tau}_i \tag{6.4}$$

for which we give an approximate 95 % confidence interval of

$$\tau_i = \sqrt{\widehat{\tau}_i^2 \pm 1.96\, d(\widehat{\tau}_i^2)}, \tag{6.5}$$

where, treating $\hat{c}_i$ as fixed due to its small variance, we compute

$$d(\widehat{\tau}_i^2) = \hat{c}_i^2 \sqrt{\frac{1}{B(B-1)}\sum_{t=1}^{B}\left[\left(R_i^{(\nu_t)} - \widehat{R}_i^{(\nu_t)}\right)^2 - \widetilde{\tau}_i^2\right]^2}.$$

The bootstrap distribution of $Q_i^{(\nu_t)} = R_i^{(\nu_t)} / \widehat{R}_i^{(\nu_t)}$ is an estimate of the distribution of the real world ratio of (reserve-ex-post)/(reserve-ex-ante). Björkwall et al. (2009) remark, in the context of triangle-only bootstraps, that the 99.5 percentile might be unreliable. BICH should yield better high percentile estimates, at least for consumer insurance with thousands of claims per year, provided our model assumptions are sufficiently satisfied, in particular so that sufficiently many large claims are represented in $Z$. See the discussion in Section 4.3.

For confidence intervals for quantiles, see Wilcox (1997), p. 87. We give a simple way to compensate for quantile uncertainty. Let $q_p$ be the $p$-quantile for $Q_i^{(\nu_t)}$ and let $\hat{q}_p$ be the empirical $p$-quantile obtained with $B$ samples. Let $p_0 = \mathrm{P}(Q_i^{(\nu_t)} \le \hat{q}_p)$. Let us regard $X = pB =$ number of observations $\le \hat{q}_p$ as random and $\hat{q}_p$ as fixed. Then $X$ is binomial $(B, p_0)$, and with a normal approximation we can state with approximately 95 % confidence that $p_0 \ge p_u = p - 1.6449\ \sqrt{p(1-p)}/\sqrt{B}$. This is equivalent to $q_{p_u} \le q_{p_0} = \hat{q}_p$ (95 %). Taking $p = 0.996034$ and $B = 10{,}000$ gives $p_u = 0.995$, so that the true 99.5 percentile is $\le$ the empirical 99.6034 percentile with $\sim 95$ % confidence. We can determine $B$ from $p_u$ and $p$ as $B = 1.6449^2 p(1-p)/(p-p_u)^2$. Taking e.g. $p_u = 0.995$ and $p = 0.9955$ yields $B = 48{,}481$.

Let $v(X)$ denote 100 times the CV (coefficient of variation) of a random variable $X$, i.e. the CV in percent. Let $\hat{v}(X)$ denote its estimate. The following averages, standard errors (except (6.10)) and CV-estimates to illuminate the model are then calculated with BICH:

$$H_i^{(-)} = \frac{1}{B}\sum_{t=1}^{B} H_i^{(\nu_t)} \tag{6.6}$$

$$\widehat{R}_i^{(-)} = \frac{1}{B}\sum_{t=1}^{B} \widehat{R}_i^{(\nu_t)} \tag{6.7}$$

$$R_i^{(-)} = \frac{1}{B}\sum_{t=1}^{B} R_i^{(\nu_t)} \tag{6.8}$$

$$\widehat{D}[\widehat{R}_i^{(\nu_1)}] = \sqrt{\frac{1}{B-1}\sum_{t=1}^{B}\left(\widehat{R}_i^{(\nu_t)} - \widehat{R}_i^{(-)}\right)^2} \tag{6.9}$$

$$\widehat{D}[R_i^{(\nu_1)}] = \sqrt{\frac{1}{B-1}\sum_{t=1}^{B}\left(R_i^{(\nu_t)} - R_i^{(-)}\right)^2} \tag{6.10}$$

$$\hat{v}(\widehat{R}_i^{(-)}) = 100\sqrt{\frac{1}{B}}\widehat{D}[\widehat{R}_i^{(\nu_1)}]/\widehat{R}_i^{(-)} \tag{6.11}$$

$$\hat{v}(R_i^{(-)}) = 100\sqrt{\frac{1}{B}}\widehat{D}[R_i^{(\nu_1)}]/R_i^{(-)} \tag{6.12}$$

$$Q_i^{(-)} = \frac{1}{B}\sum_{t=1}^{B} Q_i^{(\nu_t)} = \frac{1}{B}\sum_{t=1}^{B} R_i^{(\nu_t)}/\widehat{R}_i^{(\nu_t)} \tag{6.13}$$

$$s(Q_i^{(-)}) = 100\sqrt{\frac{1}{B(B-1)}\sum_{t=1}^{B}\left(Q_i^{(\nu_t)} - Q_i^{(-)}\right)^2}. \tag{6.14}$$

The ratio $Q_i^{(-)}$ of mean (reserve-ex-post)/(reserve-ex-ante) in bootstrap and its standard error $0.01s(Q_i^{(-)})$ serve to judge whether Hypothesis H0 is sufficiently true, and if not give ideas for improvement of the reserving function $\widehat{R}_i$.

### 6.2. Test for similarity of bootstrap and object triangles

The mean value estimate $\widehat{R}_i^{(-)}$ in (6.7) and the standard deviation estimate $\widehat{D}[\widehat{R}_i^{(\nu_1)}]$ in (6.9) can be used to judge if the realized bootstrap images $\nu_t$ are sufficiently like the outcome of $\mathcal{G}$, after multiplication by the constant $c_i$ of Assumption A4.

We propose also the following test. We define a metric of distance between increment triangles that admits different general levels, reflecting Assumption A4. Thus we measure the difference between two normalized triangles. Let $\nabla_1$ and $\nabla_2$ be triangles of known development, where

$$\nabla_1 = \{V_{ij},\ \ i \in \{1,\dots,n\},\ j \in \{1,\dots,n-i+1\}\},$$
$$\nabla_2 = \{W_{ij},\ i \in \{1,\dots,n\},\ j \in \{1,\dots,n-i+1\}\}.$$

Then set

$$\rho_n(\nabla_1,\nabla_2) = \sum_{i,j} \left| \frac{V_{ij}}{|\sum_{k,r} V_{rk}|} - \frac{W_{ij}}{|\sum_{k,r} W_{rk}|} \right|, \tag{6.15}$$

with sums over $\{i \geq 1,\ j \geq 1,\ i+j \leq n+1\}$ and $\{k \geq 1,\ r \geq 1,\ k+r \leq n+1\}$.

We define $\nabla_0$ as the object triangle and $\nabla_t$ as the bootstrap triangles, i.e.

$$\nabla_0 = \{Y_{ij},\ i \in \{1,\dots,n\},\ j \in \{1,\dots,n-i+1\}\}$$
$$\nabla_t = \left\{Y_{ij}^{(\nu_t)},\ i \in \{1,\dots,n\},\ j \in \{1,\dots,n-i+1\}\right\} \text{ for } t \in \{1,2,\dots,B\}.$$

The arithmetic mean bootstrapped triangle, which is computed as the triangle of arithmetic means of the elements, is

$$\nabla_- = \frac{1}{B}\sum_{t=1}^{B} \nabla_t \approx \mathrm{E}[\nabla_1] \tag{6.16}$$

where the $\approx$ holds for sufficiently large $B$ by the strong law of large numbers.

We would ideally like to have $\nabla_t = \nabla_- = \nabla_0$ for all $t$, since all relevant randomness in the real world is conditional on $\mathcal{G}$, of which $\nabla_0$ is a function. In applications however, this would destroy the randomness of the lower future triangles $\Delta_t$. And volatility is also necessary to judge parameter uncertainty. We tried experiments with $G^{(\nu)}$ defined by $\rho_n(\cdot)$, but they were not successful.

We propose to compare $\rho_n(\mathrm{E}[\nabla_1],\nabla_0)$ to the distribution of $\rho_n(\mathrm{E}[\nabla_1],\nabla_t)$. If the former is less than e.g. the 95 % percentile of the latter, then $\nabla_0$ can be assumed to have been drawn from the distribution of $\nabla_t$.

We have to compute the value $\rho_n(\nabla_-,\nabla_0)$ and percentiles for the empirical distribution of $\rho_n(\nabla_-,\nabla_t)$ in two bootstrap stages. The first one for computing $\nabla_-$ to sufficient closeness to its expectation $\mathrm{E}[\nabla_1]$, and the second one for percentiles.

## 7. Data and examples with tables of BICH applications

### 7.1. Real data

BICH requires detailed data of possibly many millions lines of claim info, in contrast to the triangle-only methods. This presupposes that large claim tables with one line per payment are available, which we strongly recommend at least for middle-size to large insurance companies.

### 7.2. Simulated data

We give four examples. For secrecy reasons only simulated data are used. The bootstrapped claims had the same distribution as the object claims. The number of bootstrap repetitions was $B = 10,000$ for all examples. The subsets $G^{(\nu)}$ with boundaries $b_1$ and $b_2$, described in Section 4.1, had no effect for Examples 1, 3 , 4 so we set $G^{(\nu)}$ to the whole sample space of experiment $\nu$. For Example 2, the mechanism generating $G^{(\nu)}$ described in Section 4.2 was used.

The values of $\rho_n(\triangledown_-, \triangledown_0)$ were as expected for simulated data which obey our assumptions, i.e. most below the 90 % percentile.

We have also computed examples with real data (from Länsförsäkringar Alliance) for change of incurred, where claims-handler practice had changed over time. For those the value of $\rho_n(\triangledown_-, \triangledown_0)$ was above the max-value of the distribution of $\rho_n(\triangledown_-, \triangledown_t)$. So for those cases the $\rho_n$-test proved its power to show when BICH is not appropriate. These examples are not rendered here.

Table 3 gives a comparison of standard errors by Mack (1999) and by BICH, and of $\sqrt{\text{MSEP}}$s by the Schnieper (1991), Liu and Verrall (2009) method and by BICH.

We give detailed tables only for Example 1 with chain ladder. Then in Section 7.8 we compare chain ladder, the Schnieper (1991) method and RDC side by side. We give reserve predictions, MSEP square roots and, for Example 4, the ratios of mean (reserve-ex-post)/(reserve-ex-ante). In all cases the confidence interval widths and standard errors are so small that the differences between methods are certain, except for the $\sqrt{\text{MSEP}}$s 233,891, 231,280, 1,347,084 and 1,359,985 in Table 6.

The time period is month. For Examples 1, 3, 4 we simulated 1,000,000 claims with probability 1/24 for each one of the claim occurrence months 2008-01, ... , 2009-12. For Example 2 we simulated 50,000 claims for each one of the claim occurrence months 2008-01, ... , 2009-12. The claims of 2009 are the object claims and those of 2008 are the bootstrap claims. Thus $n = 12$.

For claim periods $i$ such that all claim reportings are known, i.e. such that the largest possible value of $W$ is $\leq n - i + 1$, the Schnieper (1991) reserves are the same as the chain ladder reserves. Thus we do not give Schnieper results for those $i$.

The exposures required for the Schnieper reserves are for all examples the number of claims reported in the first development period, as given by (7.7), also when $\mathrm{P}(W > 2) > 0$. This should be the best choice, even if insurance exposures are known.

### 7.3. Example 1 satisfying the Mack assumptions

Tables 1 and 2 give output from BICH applying chain ladder. The assumptions of Mack (1999) hold with $\alpha = 1$ and $w_{ik} = 1$.

Report month was supposed to be the same as occurrence month. I.e. no IBNR (taken as distinct from RBNS) was constructed, so that $N_i = M_i$. (The Mack assumptions cannot be satisfied for an individual claim reported in a later period than the claim period without all payments being 0 with probability 1.) No tail was assumed, i.e. all claims are finalized within the first twelve development periods.

For each claim, 12 increments for development periods $j = 1, \ldots, 12$ paid in the months claim-month$+j-1$ were recursively simulated with expectations and variances determined by these vectors

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_j$ | 1.60 | 1.50 | 1.40 | 1.35 | 1.30 | 1.25 | 1.20 | 1.15 | 1.10 | 1.07 | 1.01 |
| $\sigma_j^2$ | 60 | 50 | 40 | 35 | 30 | 25 | 20 | 15 | 10 | 7 | 1 |

in this way. The first increment $C_{i1}$ was drawn from a uniform distribution on (0,100), i.e. with mean 50. Then for $j = 1, \ldots, 11$ the increment $C_{i,j+1} - C_{i,j}$ was drawn from a lognormal distribution with mean $(f_j - 1)C_{i,j}$ and variance $\sigma_j^2 C_{i,j}$. Thus the development factors $F_{i,j} = C_{i,j+1}/C_{i,j}$ and the cumulative amounts $C_{i,j}$ satisfy the conditions (CL1), (CL2), (CL3), with $f_j$ and $\sigma_j^2$ having the same meaning, in Mack (1999).

Table 1 gives object variables and the transferred prediction error estimate $\widehat{\tau}_i$ by (6.4) from the bootstrap world.

Table 2 gives bootstrap variables. The two rightmost columns can be made arbitrarily small by taking $B$ arbitrarily large. Hypothesis H0 is true. Hence $\widehat{\tau}_i$ is also a standard error, i.e. an estimate of the standard deviation of $R_i$ conditional on $\mathcal{G}$.

The program also gives a table of percentiles of the predictive distribution of the ratio (reserve-ex-post)/(reserve-ex-ante), but it is not rendered here.

### 7.4. Example 2 satisfying the Liu and Verrall assumptions

The Schnieper method is to partition the increments in claim period $i$ and development period $j$ into increments $N_{ij}$ from new claims reported in $j$ and increments $-D_{ij}$ from claims reported before $j$, where $1 \leq i, j \leq n$. (The minus sign of the latter reflects the concrete reinsurance case considered in Schnieper (1991), where the claims-handler reserve presumably typically was set too high for an individual claim.) The total increment is then $N_{ij} - D_{ij}$. Assumptions for means and variances of the distributions of $N_{ij}$ and $D_{ij}$, and for dependence and correlation, are stated. An exposure number $E_i$ per claim period $i$ is supposed to be available and used for the distribution of $N_{ij}$. For details, see Schnieper (1991) and Liu and Verrall (2009).

As remarked in Section 2.3, the BICH and Liu and Verrall (2009) conditions can coexist, if all claims are reported in the claim occurrence period or the next period. So, for Example 2, we assume that $\mathrm{P}(W \leq 2) = 1$. We can

TABLE 1

EXAMPLE 1 OBJECT STATISTICS AND ESTIMATES

| $i$ | $M_i$ (2.4) | $\hat{N}_i$ (3.2) | $H_i$ (2.7) | $\widehat{R}_i$ (2.3) | $\widehat{\tau}_i$ (6.4) | Conf.interval for $\widehat{\tau}_i$ (6.5) | |
|---|---|---|---|---|---|---|---|
| 1 | 41,697 | 41,697 | 25,264,318 | 0 | 0 | 0 | 0 |
| 2 | 41,363 | 41,363 | 24,632,778 | 246,737 | 6,483 | 6,392 | 6,572 |
| 3 | 41,578 | 41,578 | 23,180,702 | 1,854,974 | 16,733 | 16,497 | 16,964 |
| 4 | 41,689 | 41,689 | 21,384,930 | 4,017,889 | 25,061 | 24,717 | 25,400 |
| 5 | 41,602 | 41,602 | 18,332,901 | 6,707,639 | 32,381 | 31,926 | 32,829 |
| 6 | 41,962 | 41,962 | 15,604,248 | 9,960,468 | 42,949 | 42,350 | 43,540 |
| 7 | 41,862 | 41,862 | 12,230,428 | 12,810,982 | 51,732 | 51,005 | 52,448 |
| 8 | 41,914 | 41,914 | 9,524,612 | 15,816,812 | 63,980 | 63,106 | 64,843 |
| 9 | 41,631 | 41,631 | 6,935,343 | 17,988,816 | 77,279 | 76,204 | 78,340 |
| 10 | 41,623 | 41,623 | 4,971,429 | 20,041,740 | 93,787 | 92,489 | 95,068 |
| 11 | 41,861 | 41,861 | 3,339,478 | 21,842,293 | 116,764 | 115,141 | 118,365 |
| 12 | 41,614 | 41,614 | 2,075,695 | 22,931,216 | 144,441 | 142,462 | 146,393 |
| TT | 500,396 | 500,396 | 167,476,861 | 134,219,568 | 314,510 | 310,079 | 318,879 |

TABLE 2

EXAMPLE 1 BOOTSTRAP STATISTICS AND ESTIMATES

| $i$ | $H_i^{(-)}$ (6.6) | $\widehat{R}_i^{(-)}$ (6.7) | $R_i^{(-)}$ (6.8) | $\widehat{D}[\widehat{R}_i^{(\nu_1)}]$ (6.9) | $Q_i^{(-)}$ (6.13) | $s(Q_i^{(-)})$ (6.14) | $\hat{v}(\widehat{R}_i^{(-)})$ (6.11) | $\hat{v}(R_i^{(-)})$ (6.12) |
|---|---|---|---|---|---|---|---|---|
| 1 | 25,220,947 | 0 | 0 | 0 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 2 | 24,772,169 | 244,265 | 244,228 | 4,812 | 1.0002 | 0.0267 | 0.0197 | 0.0198 |
| 3 | 23,275,940 | 1,872,850 | 1,872,685 | 15,287 | 0.9999 | 0.0090 | 0.0082 | 0.0095 |
| 4 | 21,215,806 | 4,002,615 | 4,002,171 | 27,468 | 0.9999 | 0.0062 | 0.0069 | 0.0080 |
| 5 | 18,405,531 | 6,757,415 | 6,757,088 | 44,368 | 1.0000 | 0.0048 | 0.0066 | 0.0074 |
| 6 | 15,469,705 | 9,912,876 | 9,912,563 | 62,433 | 1.0000 | 0.0043 | 0.0063 | 0.0070 |
| 7 | 12,347,380 | 12,973,470 | 12,973,036 | 80,089 | 1.0000 | 0.0040 | 0.0062 | 0.0070 |
| 8 | 9,506,985 | 15,842,589 | 15,842,691 | 92,538 | 1.0000 | 0.0040 | 0.0058 | 0.0066 |
| 9 | 6,991,668 | 18,187,886 | 18,187,235 | 103,609 | 1.0000 | 0.0043 | 0.0057 | 0.0067 |
| 10 | 4,995,341 | 20,177,470 | 20,176,366 | 109,673 | 0.9999 | 0.0047 | 0.0054 | 0.0067 |
| 11 | 3,347,400 | 21,971,247 | 21,971,925 | 104,154 | 1.0000 | 0.0053 | 0.0047 | 0.0066 |
| 12 | 2,080,615 | 23,089,631 | 23,091,563 | 81,190 | 1.0001 | 0.0063 | 0.0035 | 0.0065 |
| TT | 167,629,487 | 135,032,312 | 135,031,551 | 359,368 | 1.0000 | 0.0023 | 0.0027 | 0.0024 |

deduce the values of $\lambda_2$ and $\sigma_2^2$, as defined in $A_1'$ of Schnieper (1991) and (2.3), (2.5) of Liu and Verrall (2009). We can also set suitable values for $E_i$. This will enable us to construct a claim set obeying the BICH and Liu and Verrall (2009) assumptions and to compute the exact parameter and MSEP values for this set. Let

$$A_{iw} = \text{number of claims reported in development period } w,\ w \in \{1,2\} \tag{7.1}$$

Let $p = \mathrm{P}(W = 1)$. Then $A_{i1} + A_{i2} \mid A_{i1}$ is negatively binomial $\mathrm{NB}(A_{i1}, p)$. From the properties of this distribution (see Section 3 after the overview) we get

$$\mathrm{E}[A_{i2} \mid A_{i1}] = A_{i1}(1-p)/p \qquad \mathrm{Var}[A_{i2} \mid A_{i1}] = A_{i1}(1-p)/p^2. \tag{7.2}$$

Let $X_{ik},\ k \in \{1, \dots, A_{i2}\}$, be the payments for those claims in claim period $i$ and development period 2 that have $W = 2$. In the BICH model, these are IID and independent of the past history. Let $\mu = \mathrm{E}[X_{ik}]$ and $\sigma^2 = \mathrm{Var}[X_{ik}]$. Then

$$N_{i2} = \sum_{k=1}^{A_{i2}} X_{ik}, \tag{7.3}$$

$$\mathrm{E}[N_{i2} \,|\, A_{i1}, N_{i1}] = \mathrm{E}[\mathrm{E}[N_{i2} \,|\, A_{i2}] \,|\, A_{i1}, N_{i1}] = \mathrm{E}[\mu A_{i2} \,|\, A_{i1}, N_{i1}] = \mu A_{i1}(1-p)/p, \tag{7.4}$$

$$\mathrm{Var}[N_{i2} \mid A_{i1}, N_{i1}] = \mathrm{Var}[\mathrm{E}[N_{i2} \mid A_{i2}] \mid A_{i1}, N_{i1}] + \mathrm{E}[\mathrm{Var}[N_{i2} \mid A_{i2}] \mid A_{i1}, N_{i1}]$$
$$= \mathrm{Var}[\mu A_{i2} \mid A_{i1}, N_{i1}] + \mathrm{E}[\sigma^2 A_{i2} \mid A_{i1}, N_{i1}] = \mu^2 A_{i1}(1-p)/p^2 + \sigma^2 A_{i1}(1-p)/p. \tag{7.5}$$

Thus

$$\begin{cases} \mathrm{E}[N_{i2} \mid A_{i1}, N_{i1}] = A_{i1}\mu(1-p)/p \\ \mathrm{Var}[N_{i2} \mid A_{i1}, N_{i1}] = A_{i1}(\mu^2 + p\sigma^2)(1-p)/p^2 \end{cases} \tag{7.6}$$

If we proceed like this for $N_{i3}, N_{i4}, \dots$ when $\mathrm{P}(W \geq 3) > 0$, we get similar but more complicated expressions showing how $N_{i1}, N_{i2}, \dots, N_{in}$ are dependent.

Now we regard $A_{i1}$ and $N_{i1}$ as non-stochastic, as described at the end of Section 2.3. So if we take

$$E_i = A_{i1} \tag{7.7}$$

the conditions for the Schnieper (1991) and Liu and Verrall (2009) model are satisfied. Here $\lambda_2$ and $\sigma_2^2$ are given by the coefficients for $A_{i1}$ in (7.6).

A similar construction cannot be applied for $j \geq 3$ when $\mathrm{P}(W \geq 3) > 0$.

We let $\mathrm{P}(W = 1) = p = 0.75$ and $\mathrm{P}(W = 2) = 1 - p = 0.25$. All claims had their last payment in development period 12. The payment at development period $W$ was uniformly distributed on (50,70). At subsequent development periods $j$, the incremental payment had a uniform distribution with mean $-\delta_j X$ and variance $\tau_j^2 X$, where $X$ was the claim's cumulative payment before $j$. The

$\delta_j$ and $\tau_j^2$ are given below and have the same meaning as in $A_1'$ of Schnieper (1991) and (2.3), (2.5) of Liu and Verrall (2009).

| $j$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\delta_j$ | -0.40 | 0.10 | -0.06 | -0.07 | 0.05 | 0.06 | -0.03 | -0.03 | 0.02 | -0.02 | 0.01 |
| $\tau_j^2$ | 0.70 | 0.60 | 0.50 | 0.40 | 0.30 | 0.20 | 0.10 | 0.07 | 0.05 | 0.03 | 0.00 |

Using (7.6) we compute:

| $j$ | 1 | 2 |
|---|---|---|
| $\lambda_j$ | $\mu = 60$ | $\mu(1-p)/p = 20$ |
| $\sigma_j^2$ | $\sigma^2 = 400/12 = 33.3333...$ | $(\mu^2 + p\sigma^2)(1-p)/p^2 = 1611.1111...$ |

Results are given only in Table 3. Section 7.8 does not deal with Example 2. The Schnieper method should be best here, but no significant difference between it and the chain ladder could be found, even with 40,000 bootstrap repetitions. No significant differences could either be found between the Schnieper and RDC methods, although the latter presupposes only the general BICH assumptions and contains no mean or variance structures.

The exact MSEP values were obtained by using the real parameter values above for $\lambda_j$, $\sigma_j^2$, $\delta_j$ and $\tau_j^2$ in the MSEP formulas of Liu and Verrall (2009), version "L & V Original". The reserves themselves were computed using estimates.

Version "L & V Original" was used also for the MSEP estimates. Version "L & V with adjustment" gave the same results after rounding to integers. Version "Mack's Approximation" differed only, after rounding to integers, in the value for $i = 12$. That version's $\sqrt{\text{MSEP}}$ was 8,234, hardly distinguishable from 8,239.

### 7.5. Example 3

With probability 0.6 a claim is reported in period 1, i.e. in the claim occurrence month. With probability 0.2 in period 2, with probability 0.1 in period 3 and with probability 0.1 in period 4. (This gives IBNR factors 1.6667, 1.2500 and 1.1111 for the last claim month, next last month and the month before that.)

The total claim amount for a claim was simulated in two steps. First a random mean claim $\mu$ was drawn from a uniform distribution on (0,100), i.e. with mean 50. Then a lognormal claim amount $X$ with mean $\mu$ and variance $\mu^2$ was generated. These were then simulated to be paid in $Q$ equal payments $X/Q$ in the months {report-month, report-month+1, . . . ,report-month+$Q-1$ }, with $P(Q = j) = 1/9$ $(j = 1, \dots, 9)$. Thus no payment was made later than month 12.

### 7.6. Example 4

Two segments of claim were simulated, with the segment 1 claims occurring in months 01–03 and the segment 2 claims occurring in months 04–12. Thus, we use segmenting with $s_0 = 2$, as described in Section 5. For the RDC method to work there should be at least some finalized object claims with the last possible $F$ in each segment. Otherwise there will be a negative bias. So we let the segment 1 claims have $F \leq n = 12$, making the month 01 claims finalized. And we let the segment 2 claims have $F \leq n - 3 = 9$, making the month 04 claims finalized.

First we simulated the reporting delay $W$. Then given $W$, we simulated the life length $L = F - W + 1$. The probabilities were these. P under $w$ means $\mathrm{P}(W = w)$ and P under $\lambda$ means $\mathrm{P}(L = \lambda)$.

| Seg | $w$ | 1 | 2 | 3 | 4 | $\lambda$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | P | 0.60 | 0.20 | 0.10 | 0.10 | P | 0.00 | 0.00 | 0.30 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.20 |
| 2 | P | 0.70 | 0.15 | 0.15 | | P | 0.00 | 0.60 | 0.10 | 0.10 | 0.05 | 0.05 | 0.10 | | |

The monthly payments, counted with index $h \geq 1$ from reporting, are $Y(r, h + W - 1)$. We follow here expression (2.1) for bootstrap claims. Analogously for object claims. They were constructed recursively from a sequence of random means $\mu_h$ to be lognormal with mean $\mu_h$ and variance $\mu_h^2$, conditional on $\mu_h$.

First $\mu_1$ was drawn from a uniform distribution, conditional on $L$.

$\mu_1 \sim U(0, 100\sqrt{L-1})$ for segment 1,
$\mu_1 \sim U(0, 100/\sqrt{L-1})$ for segment 2.

Then for $h = 2, \ldots, L$

$\mu_h = Y(r, h - 1 + W - 1)(1 + 0.1(W - 1))$ for segment 1,
$\mu_h = Y(r, h - 1 + W - 1)(1 - 0.1(W - 1))$ for segment 2.

### 7.7. Benchmark tests of BICH

For Example 1, Table 3 compares the standard errors $\widehat{\tau}_i$ with those of Mack (1999) and with the exact standard deviations. The latter were computed by using the real $f_j$ and $\sigma_j^2$, not their estimates, in the Mack formulas. Condition CL1 of Mack (1999) is satisfied, so the standard deviations are the same as the MSEP square roots.

For Example 2, the comparison is between $\widehat{\tau}_i$, the Liu and Verrall (2009) MSEP square roots for the Schnieper (1991) reserves and the exact values, as described in Section 7.4. The total over all claim periods for the latter was omitted, due to the computational complexity.

It is seen that $\widehat{\tau}_i$ is closer to the true values than the estimates of Mack (1999) and Liu and Verrall (2009), respectively. This is a natural consequence of the former being computed using millions of lines of detailed data and the latter being computed using only a few aggregated numbers. If only a few hundred claims had been available, the comparison could have been reversed.

The value of Table 3 is as a benchmark for the BICH method. The implication, albeit vague, is that we trust BICH to give correct prediction errors in other situations with about the same number of claims and the same variation of payments.

TABLE 3

COMPARING STANDARD ERRORS AND MSEPS

| | Example 1 | | | Example 2 | | | |
|---|---|---|---|---|---|---|---|
| $i$ | $\widehat{\tau}_i$ (6.4) | Mack s.e.$(\hat{C}_{in})$ | Exact standard deviation | Schnieper reserve | $\widehat{\tau}_i$ BICH $\sqrt{\text{MSEP}}$ | Liu and Verrall $\sqrt{\text{MSEP}}$ | Exact $\sqrt{\text{MSEP}}$ |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 6,483 | 3,646 | 6,992 | -37,713 | 0 | 0 | 0 |
| 3 | 16,733 | 9,273 | 17,250 | 36,386 | 407 | 214 | 404 |
| 4 | 25,061 | 24,530 | 25,253 | -39,280 | 647 | 543 | 648 |
| 5 | 32,381 | 25,595 | 33,337 | 70,669 | 858 | 590 | 856 |
| 6 | 42,949 | 40,981 | 42,934 | 177,981 | 1,086 | 630 | 1,085 |
| 7 | 51,732 | 47,473 | 52,506 | -49,094 | 1,456 | 1,127 | 1,466 |
| 8 | 63,980 | 55,950 | 64,592 | -247,514 | 1,871 | 1,481 | 1,865 |
| 9 | 77,279 | 65,797 | 77,637 | 13,427 | 2,182 | 1,966 | 2,226 |
| 10 | 93,787 | 81,563 | 94,404 | 224,620 | 2,592 | 2,302 | 2,631 |
| 11 | 116,764 | 107,167 | 116,842 | -165,387 | 3,151 | 2,877 | 3,131 |
| 12 | 144,441 | 141,901 | 146,029 | 1,477,940 | 8,263 | 8,239 | 8,448 |
| TT | 314,510 | 281,812 | 316,165 | 1,462,034 | 11,154 | | |

### 7.8. Using BICH to compare reserving methods

For Examples 1, 3, 4 we now compare chain ladder, Schnieper and RDC using their BICH MSEPs. Example 2 showed no significant MSEP differences. For RDC we used $q_0 = 500$ for the number of quantile intervals of paid up to 'now'. The difference in results between $q_0 = 100$ and $q_0 = 500$ was not large. No upper limit $w_0$, as described in Section A.2, was set for Examples 1 and 4. For Example 3, a couple of BICH bootstraps showed that we should set $w_0 = 1$, which is in line with the claim simulation construction.

As can be seen from Table 4, chain ladder is only slightly better than RDC for Example 1, although the simulated claims were tailor-made for chain ladder. When we made a PPCF type computation, using only finalized claims for mean

TABLE 4

Example 1 comparison

| | Chain ladder | | RDC | |
|---|---|---|---|---|
| $i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ |
| 1 | 0 | 0 | 0 | 0 |
| 2 | 246,737 | 6,483 | 247,337 | 6,510 |
| 3 | 1,854,974 | 16,733 | 1,853,444 | 17,244 |
| 4 | 4,017,889 | 25,061 | 4,010,803 | 25,847 |
| 5 | 6,707,639 | 32,381 | 6,710,155 | 34,172 |
| 6 | 9,960,468 | 42,949 | 9,955,997 | 43,979 |
| 7 | 12,810,982 | 51,732 | 12,831,583 | 53,120 |
| 8 | 15,816,812 | 63,980 | 15,848,494 | 65,843 |
| 9 | 17,988,816 | 77,279 | 18,006,636 | 80,318 |
| 10 | 20,041,740 | 93,787 | 20,018,086 | 96,109 |
| 11 | 21,842,293 | 116,764 | 21,862,177 | 119,753 |
| 12 | 22,931,216 | 144,441 | 22,940,208 | 146,503 |
| TT | 134,219,568 | 314,510 | 134,284,919 | 325,508 |

TABLE 5

Example 3 comparison

| | Chain ladder | | Schnieper | | RDC | |
|---|---|---|---|---|---|---|
| $i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ |
| 1 | 0 | 0 | | | 0 | 0 |
| 2 | 2,758 | 257 | | | 2,440 | 18 |
| 3 | 10,752 | 590 | | | 10,020 | 156 |
| 4 | 30,383 | 1,055 | same | | 30,503 | 354 |
| 5 | 77,833 | 1,723 | as | | 79,320 | 667 |
| 6 | 157,909 | 2,611 | chain | | 160,583 | 1,183 |
| 7 | 278,019 | 3,821 | ladder | | 274,719 | 1,888 |
| 8 | 448,487 | 5,360 | | | 443,958 | 2,788 |
| 9 | 666,052 | 7,434 | | | 672,901 | 3,844 |
| 10 | 979,091 | 10,957 | 979,568 | 10,651 | 969,195 | 7,051 |
| 11 | 1,332,668 | 15,485 | 1,331,496 | 14,230 | 1,327,760 | 9,850 |
| 12 | 1,707,369 | 24,276 | 1,708,036 | 19,355 | 1,713,257 | 13,349 |
| TT | 5,691,321 | 36,237 | 5,691,292 | 32,109 | 5,684,656 | 20,929 |

payment estimates, we obtained total prediction error about 535,000, which is much worse than 325,508 for RDC. For Examples 3 and 4 RDC was better. The Schnieper method performed between chain ladder and RDC.

TABLE 6

EXAMPLE 4 COMPARISON

| | CHAIN LADDER | | | SCHNIEPER | | | RDC | | |
|---|---|---|---|---|---|---|---|---|---|
| $i$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $Q_i^{(-)}$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $Q_i^{(-)}$ | $\widehat{R}_i$ | $\widehat{\tau}_i$ | $Q_i^{(-)}$ |
| 1 | 0 | 0 | | | | | 0 | 0 | |
| 2 | 781,236 | 639,083 | 1.1615 | | | | 743,633 | 635,319 | 1.1967 |
| 3 | 1,951,032 | 928,295 | 1.0388 | | | | 2,146,695 | 886,785 | 1.0591 |
| 4 | 0 | 0 | 0.0000 | | same | | 0 | 0 | 0.0000 |
| 5 | 16,185 | 8,858 | 1.1186 | | as | | 18,791 | 9,567 | 2.0326 |
| 6 | 73,082 | 19,854 | 1.0229 | | chain | | 71,237 | 21,881 | 1.2023 |
| 7 | 448,671 | 56,890 | 1.0057 | | ladder | | 444,991 | 51,257 | 1.0249 |
| 8 | 1,057,292 | 99,351 | 1.0022 | | | | 1,052,013 | 90,488 | 1.0089 |
| 9 | 2,072,212 | 136,817 | 1.0008 | | | | 2,168,135 | 123,484 | 1.0056 |
| 10 | 4,496,636 | 171,564 | 1.0003 | | | | 4,466,886 | 156,713 | 1.0016 |
| 11 | 8,463,515 | 213,716 | 1.0002 | 8,471,478 | 209,846 | 1.0006 | 8,487,404 | 190,951 | 1.0008 |
| 12 | 15,771,619 | 242,177 | 1.0000 | 15,817,473 | 233,891 | 1.0004 | 15,888,240 | 231,280 | 0.9999 |
| TT | 35,131,481 | 1,347,084 | 1.0002 | 35,185,297 | 1,359,985 | 1.0014 | 35,488,026 | 1,295,050 | 1.0042 |

Run time for RDC was about five hours, while chain ladder takes less than an hour. For MSEP estimate precision it will not in practice be necessary to make $B = 10,000$ repetitions, which we made here to be certain of our results. For these examples $B = 2,000$ repetitions should suffice. Larger $B$ might be needed for quantiles, see Section 6.

In Table 6, for Example 4 with segmentation, we give also the mean ratios $Q_i^{(-)}$ of reserve-ex-post to reserve-ex-ante. These were not close to 1 for all claim periods, as they were for Examples 1 and 3.

## 8. CONCLUSION

We have shown that BICH gives good prediction error estimates under the natural conditions of essentially IID claims with sufficiently many finalized ones. A new method RDC, also derived for IID claims, is integrated into BICH and compared with chain ladder and the Schnieper method in BICH bootstraps.

## ACKNOWLEDGEMENTS

Thanks are due to two referees and the editor Prof. Mario Wüthrich for many valuable suggestions, which helped to improve the presentation considerably.

## Appendix A: RDC method

### A.1. The RDC framework

We describe the RDC (Reserve by Detailed Conditioning) method for reserving using customer waiting-for-report periods $W$, finalization periods $F$ and payments $Y(\cdot)$ of individual claims. Payments on open and settled claims are pooled in a way that we surmise is optimal or near optimal. The assumption is that all claims in $Z$ and $T$, respectively, are IID. RDC can be used without bootstrapping, in which case $Z$ is not needed. In other words we require that $c_i \equiv c$ for some $c$ in Assumption A4. In practice this means that payments must have been adjusted for inflation beforehand. We also assume no tail, i.e. $\mathrm{P}(F \leq n) = 1$. We do neither consider the possibility of reopening a finalized claim.

No other assumptions are made. No likelihood expressions will appear in the sequel. We do not even prescribe a mean value or variance structure. The method is intended to be unbiased and give small mean square prediction errors in any situation with IID claims without tails, possibly after segmentation as described in Section 5. If special assumptions are applicable – such as independence of individual payments or independence of the sequence of individual payments and $F$ – algorithms using these will be better. But we aim to show that the RDC method will then only be marginally worse while notably better in other situations.

Besides offering an alternative to chain ladder and other methods at the aggregate level, a purpose of RDC is to offer a much enhanced PPCF method for reserving individual claims, using both finalized and non-finalized claims for estimates of claim life length and payment parameters, while avoiding the bias that a naive use of non-finalized claims can entail. Thus, RDC can replace claims-handler reserves for claim types with some volume, not too long development and not too large payments. In the sequel payments are primarily intended to mean payments, not changes in incurred (payment sum plus claims-handler reserve).

Broadly, we break down the parameters of the distribution of a claim, as formulated in (A.1) below, in many small details while conditioning on observable variables with many combinations. The form of the probability estimates in Section A.3 and Theorem A.1 for mean payments show that RDC gives consistent reserve predictions for large numbers of claims. Still, with finite samples we might risk instability from overparametrization. However, we contend that the adding of many individual claim reserves will, by the law of large numbers, often make the total RDC reserve for a claim period more stable than other reserving functions such as chain ladder and Schnieper's method. BICH bootstraps will help to set parameters for maximal stability. The contention was corroborated by BICH bootstraps on simulated data in Section 7.8.

Consider a claim in the object claim set $T$ as described by (2.2). We suppress the claim period superscript $i \in \{1, \ldots, n\}$ and the claim number $k$ below. Thus

$$\{W, F, Y(1), \ldots, Y(F)\} \tag{A.1}$$

is a claim with some distribution, the parameters of which we attempt to infer. Recall that $W = 1$ if the claim is reported in the claim occurrence period. Likewise $F = 1$ if finalized in the claim occurrence period. Define

$$L = F - W + 1 = \text{ life length of claim.} \tag{A.2}$$

The following terminology will be used:

- $W \leq n - i + 1$: the claim is **reported**;
- $L \leq n - i - W + 2$: the claim is **finalized**, i.e. settled. Then it is also reported;
- $W \leq n - i + 1$ and $L > n - i - W + 2$: the claim is reported but still **open**.

Define the sum of amounts paid up to and including period $t$ from reporting as

$$H(t) = \sum_{h=1}^{t} Y(h + W - 1), \quad t \in \{0, 1, \ldots, n\}, \tag{A.3}$$

where $h$ is counted from reporting with the reporting period $W$ having $h = 1$.

Then $H(0) = 0$, $H(L)$ is the total claim cost and $H(n - i - W + 2)$ is the payment sum up to and including the now development period $n - i + 1$. It follows that $H(L) - H(n - i - W + 2)$ is the remaining payment sum for a reported open claim after development period $n - i + 1$, i.e. the reserve-ex-post.

We want to predict the expected remaining payment sum from the known sum. Consider this expression.

$$\mathrm{E}[H(L) - H(t) \mid L > t, H(t), W] \tag{A.4}$$

For $t = n - i - W + 2$ an estimate of (A.4) gives the RBNS (Reported But Not Settled) reserve of a reported open claim. For $t = 0$ we obtain the IBNR (Incurred But Not Reported) reserve per claim.

The assumption in Wüthrich and Merz (2008), Chapter 10.1.2 under Predicting Reported Open Claims, the $\stackrel{(d)}{=}$ expression, is that, in our notation, $H(t)$ without $Y(\cdot)$ suffices for inference. This might or might not be true, but any model using the individual payments $Y(\cdot)$ would need special and questionable assumptions to be workable. Hence we will not use the individual payments $Y(\cdot)$ summing to $H(t)$. If the Wüthrich and Merz (2008) assumption is not true our calculus is valid anyway, but the reserve estimates could have smaller variances if $Y(\cdot)$ were used individually.

In the sequel, any ratio with denominator 0 is defined to be 0.

### A.2. Parameters and observable variables used in conditioning

There is no parametric model in the RDC method for the dependence of the reserve on $H(t)$. Therefore we must approximate the conditioning variable $H(t)$ with its empirical quantile intervals for reported claims in $T$ that have $L > t$.

Let $q_0$ be the number of even-spaced quantile intervals fixed in advance. Then $q_0 = 1$ means no conditioning with respect to $H(t)$. With $q_0 = 2$ we condition with respect to whether the quantile of $H(t)$ is at most 0.5 or not. With $q_0 = 3$ with respect to which one of the intervals $(0, \frac{1}{3}], (\frac{1}{3}, \frac{2}{3}], (\frac{2}{3}, 1]$ the quantile of $H(t)$ belongs to, etc. In simulations we have used up to $q_0 = 500$. Let $Q_0 \equiv 1$ and for $t > 0$

$$Q_t = \text{ interval number of the quantile of } H(t) \text{ for } L > t, \quad Q_t \in \{1, \ldots, q_0\} \tag{A.5}$$

At some point the prediction error $\tau_i$ in (2.10) will stop decreasing noticeably with increasing $q_0$. Taking note of Wüthrich and Merz (2008), last paragraph of Chapter 10, we risk overparametrization with too large $q_0$. In our simulated examples with a clear dependence on $H(t)$ we have, however, not found any $q_0^{\max}$ such that $\tau_i$ increases for $q_0 > q_0^{\max}$.

For $W$ it might not be feasible to condition with respect to all its possible values, if the distribution of the sequence $Y(W), Y(W+1), \ldots, Y(W+L-1)$ does not depend much or at all on $W$. Therefore we might want to fix a number $w_0 \geq 1$ in advance and condition with respect to $W \wedge w_0$, the minimum of $W$ and $w_0$. BICH bootstraps will indicate the proper value of $w_0$. If we do not fix $w_0$ in advance, then $w_0$ will be the largest number $w$ such that $\mathrm{P}(W = w) > 0$. Anyway $w_0 \leq n$, since we assume no tail. The distribution of $W$ itself is not studied or used in any other way than via chain ladder predictions of future numbers of claims per $W$, see (A.31).

We define the underlying reserve for a claim as

$$R(q, w, t) = \mathrm{E}[H(L) - H(t) \mid L > t, Q_t = q, W \wedge w_0 = w], \tag{A.6}$$

whose value we wish to estimate. For

$$0 \leq t \leq n-1 \qquad t+1 \leq \lambda \leq n \qquad t+1 \leq h \leq \lambda$$

define probabilities and expected payments

$$p_\lambda(q, w, t) = \mathrm{P}(L = \lambda \mid L > t, Q_t = q, W \wedge w_0 = w), \tag{A.7}$$

$$\mu_{\lambda h}(q, w, t) = \mathrm{E}[Y(h + W - 1) \mid L = \lambda, Q_t = q, W \wedge w_0 = w]. \tag{A.8}$$

Then, we have

$$R(q, w, t) = \sum_{\lambda=t+1}^{n} \sum_{h=t+1}^{\lambda} p_\lambda(q, w, t) \mu_{\lambda h}(q, w, t). \tag{A.9}$$

## A.3. Probability estimates

Probability estimates are obtained via estimates of the probabilities for finalization in a period, given that the claim was not finalized before that. I.e. with

$$r_\lambda(q, w, t) = \mathrm{P}(L = \lambda \mid L \geq \lambda, Q_t = q, W \wedge w_0 = w) \tag{A.10}$$

it holds, with an empty product defined as 1,

$$p_\lambda(q,w,t) = r_\lambda(q,w,t) \prod_{k=t+1}^{\lambda-1} [1 - r_k(q,w,t)]. \tag{A.11}$$

This is shown by chaining successive conditional survival probabilities. Namely, suppressing $q$ and $w$ we can write (A.11) in this equivalent way.

$$\mathrm{P}(L=\lambda \mid L>t) = \left(\prod_{k=t+1}^{\lambda-1} \mathrm{P}(L>k \mid L \geq k)\right) \mathrm{P}(L=\lambda \mid L \geq \lambda) \tag{A.12}$$

An estimate of (A.10) is obtained from making the observations

$$\begin{cases} I^{\mathrm{F}}_\lambda(q,w,t) = \text{number of finalized claims with } L=\lambda, Q_t=q, W{\wedge}w_0=w \\ J_\lambda(q,w,t) = \text{number of reported claims with } L \geq \lambda, Q_t=q, W{\wedge}w_0=w \end{cases} \tag{A.13}$$

and using as estimates

$$\begin{cases} \hat{r}_n(q,w,t) = 1 \\ \hat{r}_\lambda(q,w,t) = I^{\mathrm{F}}_\lambda(q,w,t)/J_\lambda(q,w,t), \quad \lambda < n, \end{cases} \tag{A.14}$$

which give the estimates $\hat{p}_\lambda$ of $p_\lambda$

$$\hat{p}_\lambda(q,w,t) = \hat{r}_\lambda(q,w,t) \prod_{k=t+1}^{\lambda-1} [1 - \hat{r}_k(q,w,t)]. \tag{A.15}$$

This indirect way allows us to use both finalized and open claims. If only finalized claims were used, the estimates would have larger variance than possible.

The discrete distributions defined above is a way to structure our observations into an empirical distribution.

### A.4. Mean payment estimates

For estimates of $\mu_{\lambda h}(q,w,t)$ we shall combine payments from open and finalized claims. Only known payments, i.e. with $h \leq n-i-W+2$, are used. The sums of $Y(h+W-1)$ is over all reported claims in all claim periods $i \in \{1,\ldots,n\}$. With

$$Y^{\mathrm{F}}_{\lambda h}(q,w,t) = \sum_{\{L \leq n-i-W+2, L=\lambda, Q_t=q, W\wedge w_0=w\}} Y(h+W-1) \tag{A.16}$$

an estimate of $\mu_{\lambda h}(q,w,t)$ using only finalized claims is $Y^{\mathrm{F}}_{\lambda h}(q,w,t)/I^{\mathrm{F}}_\lambda(q,w,t)$.

But we also want to use the open claims. Thus we observe the following variables derived from open claims for $t \le n-2$. They are not defined and not used for $t = n-1$, where (A.22) alone determines (A.27). For

$$0 \le t \le n-2 \qquad t+1 \le r \le n-1 \qquad t+1 \le h \le r$$

define claim numbers and payment sums for open claims known to be open at period $r$, i.e. $L > r$, (counted from reporting) but not known to be open later than $r$.

$$I^{\mathrm{O}}_r(q,w,t) = \text{number of claims with } n\text{–}i\text{–}W\text{+}2 = r, L>r, Q_t = q, W \wedge w_0 = w, \quad \text{(A.17)}$$

$$Y^{\mathrm{O}}_{rh}(q,w,t) = \sum_{\{n-i-W+2=r, L>r, Q_t=q, W\wedge w_0 = w\}} Y(h+W-1). \quad \text{(A.18)}$$

Below we drop $(q,w,t)$ from the notation, since these parameters are fixed in equations (A.19) – (A.42).

We compute the predicted number of open claims at $r$ with $L=\lambda$ as

$$I^{\mathrm{O}}_{r\lambda} = \frac{\hat{p}_\lambda}{\hat{p}_{r+1} + \ldots + \hat{p}_n} I^{\mathrm{O}}_r, \quad \lambda = r+1, \ldots, n. \quad \text{(A.19)}$$

It is easy to see that it holds, with the degree of approximation depending on the precision of the estimates $\hat{p}_\lambda$,

$$\mathrm{E}[Y^{\mathrm{O}}_{rh} \mid I^{\mathrm{O}}_r] \approx \sum_{\lambda=r+1}^{n} I^{\mathrm{O}}_{r\lambda}\, \mu_{\lambda h}. \quad \text{(A.20)}$$

Now we have a non-trivial problem in distributing $Y^{\mathrm{O}}_{rh}$ among the possible $L$-values $r+1, \ldots, n$. Namely, we need predicted payment sums $Y^{\mathrm{O}}_{r\lambda h}$ with $L=\lambda$ such that

$$\sum_{\lambda=r+1}^{n} Y^{\mathrm{O}}_{r\lambda h} = Y^{\mathrm{O}}_{rh}. \quad \text{(A.21)}$$

We propose an involved procedure, which is unfortunately hard to understand. But it is one that has shown itself to yield good results in BICH bootstraps. We define claim number sums $I^{(r)}_\lambda$ (normally not integers but real numbers) and payment sums $Y^{(r)}_{\lambda h}$ that are computed recursively in $r$, starting at $r = \lambda$ and going down to $r = h$. In each step the results of the previous steps are used. We use the appropriate $I^{(r)}_\lambda$ and $Y^{(r)}_{\lambda h}$ for an estimate of $\mu_{\lambda h}$ that uses both finalized and open claims in a way that can be presumed optimal or near optimal. The recursion is downwards because $Y^{\mathrm{O}}_{r\lambda h}$ can be predicted with higher precision for larger $r$. For $r = n-1$ only $\lambda = n$ is possible, so that a prediction using

$Y^{O}_{n-1,n,h}/I^{O}_{n-1,n}$ is as good as one obtained from finalized claims with the same number of claims in the denominator. And so on.

Define for

$$0 \le t \le n-1 \qquad t+1 \le \lambda \le n \qquad t+1 \le h \le \lambda$$

initial values determined by finalized claims

$$\begin{cases} I^{(\lambda)}_{\lambda} = I^{F}_{\lambda} \\ Y^{(\lambda)}_{\lambda h} = Y^{F}_{\lambda h} \end{cases} \tag{A.22}$$

and the recursion

$$\begin{cases} I^{(r)}_{\lambda} = I^{(r+1)}_{\lambda} + I^{O}_{r\lambda} \\ Y^{(r)}_{\lambda h} = Y^{(r+1)}_{\lambda h} + Y^{O}_{r\lambda h} \end{cases} \qquad r = \lambda-1,\ \lambda-2,\ \ldots,\ h \tag{A.23}$$

where, with

$$\beta_{rh} = Y^{O}_{rh}\left(\sum_{\nu=r+1}^{n} Y^{(r+1)}_{\nu h} I^{O}_{r\nu}/I^{(r+1)}_{\nu}\right)^{-1} \tag{A.24}$$

we distribute $Y^{O}_{rh}$ so that, for fixed $h$, a new mean payment estimate that can be made at this step, separately for open claims at this $r$ alone, is proportional to the previous estimate using all $r$ gone through from the top so far, namely

$$\frac{Y^{O}_{r\lambda h}}{I^{O}_{r\lambda}} = \beta_{rh}\frac{Y^{(r+1)}_{\lambda h}}{I^{(r+1)}_{\lambda}}, \qquad \text{which gives} \qquad Y^{O}_{r\lambda h} = \beta_{rh}\frac{Y^{(r+1)}_{\lambda h}}{I^{(r+1)}_{\lambda}} I^{O}_{r\lambda}, \tag{A.25}$$

provided $\beta_{rh}$ is defined, i.e. at least one of its terms within $()^{-1}$ has both numerator and denominator not 0. If this is not the case, we let all new mean payment estimates, separately for this $r$ alone, for fixed $h$ be equal, namely

$$Y^{O}_{r\lambda h} = \frac{\hat{p}_{\lambda}}{\hat{p}_{r+1} + \ldots + \hat{p}_{n}} Y^{O}_{rh} \quad (\text{if } \beta_{rh} \text{ is not defined}). \tag{A.26}$$

The constructions (A.25) and (A.26) satisfy (A.21).

The final mean payment estimates will be

$$\hat{\mu}_{\lambda h} = \frac{Y^{(h)}_{\lambda h}}{I^{(h)}_{\lambda}}, \quad h = t+1, \ldots, n; \quad \lambda = h, \ldots, n \tag{A.27}$$

We now state that, under natural conditions, the estimate $\hat{\mu}_{\lambda h}$ converges almost surely (with probability one) to $\mu_{\lambda h}$ as the sample of claims increases to infinity.

If this is true, convergence of $\mathrm{E}[\hat{\mu}_{\lambda h}]$ to $\mu_{\lambda h}$ follows, provided we assume some reasonable bounding of each absolute payment $|Y(j)|$, as defined by (A.1), that guarantees that convergence almost surely implies convergence in mean, by the dominated convergence theorem.

Let $N = N_1 + \ldots + N_n$ be the number of claims as defined by (2.5). When we write $N \to \infty$ we let all $N_i \to \infty$. With a.s. is meant convergence almost surely.

**Theorem A.1.** *Assume that one or both hold of these two conditions*

*(i)* $\sum_{\nu=r+1}^{n} p_\nu \mu_{\nu h} \neq 0$ *for* $h \leq r \leq \lambda$

*(ii)* $P(Y(h+W-1) \geq 0 \mid L=\nu, Q_t=q, W{\wedge}w_0=w) = 1$ *for* $\nu > h$

*Then* $\lim_{N\to\infty} \hat{\mu}_{\lambda h} = \mu_{\lambda h}$ *almost surely.*

The proof is given in Section A.6.

## A.5. Total reserve per claim and claim period

We use the building blocks computed to get the final reserve estimate per claim

$$\widehat{R}(q,w,t) = \sum_{\lambda=t+1}^{n} \sum_{h=t+1}^{\lambda} \hat{p}_\lambda(q,w,t)\hat{\mu}_{\lambda h}(q,w,t). \tag{A.28}$$

We can now obtain the IBNR reserve $\widehat{R}_i^{\mathrm{I}}$ and the RBNS reserve $\widehat{R}_i^{\mathrm{R}}$ for claim period $i$. These together give the reserve per claim period as defined by (2.3)

$$\widehat{R}_i = \widehat{R}_i^{\mathrm{I}} + \widehat{R}_i^{\mathrm{R}}. \tag{A.29}$$

### A.5.1. *IBNR reserve*

$$\widehat{R}(1, w{\wedge}w_0, 0) = \text{ reserve of an unreported claim with } W = w. \tag{A.30}$$

We predict the number of such claims per claim period $i$ very simply with chain ladder applied to claim reportings. More precisely, a claim contributes 1 to the increment of development period $W$ in the chain ladder algorithm. With

$$\begin{cases} A_{iw} = \text{number of claims reported in development period } w \\ \widehat{A}_{iw} = \text{chain ladder prediction of } A_{iw} \text{ for } w \geq n-i+2, \end{cases} \tag{A.31}$$

we obtain the IBNR reserve for claim period $i$ as

$$\widehat{R}_i^{\mathrm{I}} = \sum_{w=n-i+2}^{n} \widehat{A}_{iw}\widehat{R}(1, w{\wedge}w_0, 0). \tag{A.32}$$

A breakdown into development periods $j$ is obtained by taking $h+w-1=j$ i.e. $h=j-w+1$ in $\hat{\mu}_{\lambda h}$ of (A.28), giving for $j \geq n-i+2$

$$\widehat{R}^{\mathrm{I}}_{ij} = \sum_{w=n-i+2}^{j} \sum_{\lambda=j-w+1}^{n} \widehat{A}_{iw} \hat{p}_\lambda(1, w \wedge w_0, 0) \hat{\mu}_{\lambda, j-w+1}(1, w \wedge w_0, 0). \tag{A.33}$$

A.5.2. *RBNS reserve*

For RBNS in claim period $i$ we have

$$\widehat{R}(Q_{n-i-W+2}, W \wedge w_0, n-i-W+2) = \text{ reserve of an open claim.} \tag{A.34}$$

Let for claim period $i$

$$I^{\mathrm{O}}(i,q,w) = \text{number of open claims with } W=w \text{ and } Q_{n-i-W+2}=q. \tag{A.35}$$

Then the RBNS reserve for claim period $i$ is

$$\widehat{R}^{\mathrm{R}}_{i} = \sum_{w=1}^{n-i+1} \sum_{q=1}^{q_0} I^{\mathrm{O}}(i,q,w) \widehat{R}(q, w \wedge w_0, n-i-w+2). \tag{A.36}$$

Like the IBNR reserve, a breakdown into development periods $j$ is obtained from $h=j-w+1$ in (A.28), giving for $j \geq n-i+2$

$$\widehat{R}^{\mathrm{R}}_{ij} = \sum_{w=1}^{n-i+1} \sum_{q=1}^{q_0} \sum_{\lambda=j-w+1}^{n} I^{\mathrm{O}}(i,q,w) \hat{p}_\lambda(q, w \wedge w_0, n-i-w+2) \hat{\mu}_{\lambda, j-w+1}(q, w \wedge w_0, n-i-w+2). \tag{A.37}$$

## A.6. Proof of Theorem A1

We use induction in the form of downwards recursion. Namely, we shall show that (A.39) implies (A.42). First note that

$$\lim_{N\to\infty} \frac{Y^{\mathrm{O}}_{rh}}{I^{\mathrm{O}}_{r}} = \sum_{\nu=r+1}^{n} \frac{p_\nu}{p_{r+1}+\ldots+p_n} \mu_{\nu h} \quad \text{a.s.} \tag{A.38}$$

Assume

$$\lim_{N\to\infty} \frac{Y^{(r+1)}_{\lambda h}}{I^{(r+1)}_{\lambda}} = \mu_{\lambda h}, \quad \lambda = r+1, \ldots, n \quad \text{a.s.} \tag{A.39}$$

If (i) holds $\beta_{rh}$ will be defined for $N$ large enough. Then

$$\lim_{N\to\infty} \beta_{rh} = \lim_{N\to\infty} \frac{Y^{\mathrm{O}}_{rh}}{I^{\mathrm{O}}_{r}} \left( \sum_{\nu=r+1}^{n} \frac{Y^{(r+1)}_{\nu h}}{I^{(r+1)}_{\nu}} \frac{\hat{p}_\nu}{\hat{p}_{r+1}+\ldots+\hat{p}_n} \right)^{-1} = 1 \quad \text{a.s.} \tag{A.40}$$

Hence by (A.25)

$$\lim_{N\to\infty} \frac{Y^{\text{O}}_{r\lambda h}}{I^{\text{O}}_{r\lambda}} = \mu_{\lambda h} \quad \text{a.s.} \tag{A.41}$$

Thus, suppressing $r, \lambda, h$ in the right side of (A.23) and writing its terms as functions of $N$ only, we have $I^{(r)}_{\lambda} = I_{1N} + I_{2N}$ and $Y^{(r)}_{\lambda h} = Y_{1N} + Y_{2N}$ where $Y_{1N}/I_{1N} \to \mu_{\lambda h}$ and $Y_{2N}/I_{2N} \to \mu_{\lambda h}$ a.s. Hence

$$\lim_{N\to\infty} \frac{Y^{(r)}_{\lambda h}}{I^{(r)}_{\lambda}} = \mu_{\lambda h} \quad \text{a.s.} \tag{A.42}$$

Now assume that (i) is not true for some $r = r_0$ but that (ii) holds. Then for $\nu > r_0$ the probability is 0 of observing any non-zero payment with life length $L = \nu$, and we must have $\mu_{\nu h} = 0$. Then (A.42) holds with value 0.

By the strong law of large numbers (A.39) holds for $r+1 = \lambda$, since the numerator and denominator are given by (A.22) and use finalized claims only.

Going down to $r = h$ finishes the induction proof. □

## References

Björkwall, S., Hössjer, O. and Ohlsson, E. (2009) Non-parametric and parametric bootstrap techniques for age-to-age development factor methods in stochastic claims reserving. *Scandinavian Actuarial Journal*, **2009:4**, 306–331.

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

Fisher, W. H. and Lange, J. T. (1973) Loss reserve testing: a report year approach. *Proceedings of the Casualty Actuarial Society*, **60**, 189–207.

Hachemeister, C. A. and Stanard, J. N. (1975) *IBNR claims count estimation with static lag functions*. Spring meeting of the Casualty Actuarial Society.

Kuang, D., Nielsen, B. and Nielsen, J. P. (2009) Chain-ladder as maximum likelihood revisited. *Annals of Actuarial Science*, **4(1)**, 105–121.

Larsen, C. R. (2007) An individual claims reserving model. *ASTIN Bulletin*, **37(1)**, 113–132.

Liu, H. and Verrall, R. J. (2009) Predictive distributions for reserves which separate true IBNR and IBNER claims. *ASTIN Bulletin*, **39(1)**, 35–60.

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

Mack, T. (1999) The standard error of chain ladder reserve estimates: recursive calculation and inclusion of a tail factor. *ASTIN Bulletin*, **29(2)**, 361–366.

Norberg, R. (1993) Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin*, **23(1)**, 95–115.

Norberg, R. (1999) Prediction of outstanding liabilities II. Model variations and extensions. *ASTIN Bulletin*, **29(1)**, 5–25.

Sawkins, R. W. (1979) Methods of analysing claim payments in general insurance. *Transactions of the Institute of Actuaries of Australia*, 435–519.

Schnieper , R. (1991) Separating true IBNR and IBNER claims. *ASTIN Bulletin*, **21(1)**, 111–127.

Taylor, G., McGuire, G. and Sullivan, J. (2008) Individual claim loss reserving conditioned by case estimates. *Annals of Actuarial Science*, **3**, 215–256.

Taylor, G. (2011) Maximum likelihood and estimation efficiency of the chain ladder. *ASTIN Bulletin*, **41(1)**, 131–155.

VERRALL, R., NIELSEN, J. P. and JESSEN, A. (2010) Prediction of RBNS and IBNR claims using claim amounts and claim counts. *ASTIN Bulletin*, **40(2)**, 871–887.
WILCOX, R. (1997) *Introduction to Robust Estimation and Hypothesis Testing.* Academic Press.
WÜTHRICH, M. V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance.* Wiley.
ZHAO X. and ZHOU, X. (2010) Applying copula models to individual claim loss reserving methods. *Insurance: Mathematics and Economics*, **46(2)**, 290–299.

STIG ROSENLUND
*Länsförsäkringar Alliance, Sweden (retired)*
*Stockholm*
*Sweden*
*E-mail: stig.rosenlund@sverige.nu*

**Errors of print in the printed article**
Page 303 formula (6.4): $\widehat{\tau}_i = \hat{c}_i\widehat{\tau}_i$ shall be $\widehat{\tau}_i = \hat{c}_i\widetilde{\tau}_i$.
Page 324 line 2: **40(1)** shall be **40(2)**.

Some questions I have received are answered in the following four sections, which were not included in the printed paper.

## Appendix B: Elaborations

### B.1. (A.32) is consistent with (A.33) in the RDC model

This is an elaboration of Bootstrapping Individual Claim Histories, ASTIN Bulletin 2012(1), 291-324. Apart from verbal deliberations, I mainly elaborate Section A.5.

On p. 317 I state:

"For $W$ it might not be feasible to condition with respect to all its possible values, if the distribution of the sequence $Y(W), Y(W+1), \ldots, Y(W+L-1)$ does not depend much or at all on $W$. Therefore we might want to fix a number $w_0 \geq 1$ in advance and condition with respect to $W \wedge w_0$, the minimum of $W$ and $w_0$. BICH bootstraps will indicate the proper value of $w_0$. If we do not fix $w_0$ in advance, then $w_0$ will be the largest number $w$ such that $\mathrm{P}(W = w) > 0$. Anyway $w_0 \leq n$, since we assume no tail."

The development distribution can thus depend on $W$ for $W \geq w_0$, but to condition with respect to such values individually might still give an increase of MSEP due to overparametrization.

Overparametrization is present when there are parameters in the model whose estimates have so large variances that the MSEP is larger with these parameters than without them. This can be the case even if the use of the *true* parameter values would decrease or at least not increase the MSEP. Consider polynomial regression. If this model is true and the true degree of the regression curve is $p$, the use of the true parameters also for degrees $> p$ will be as good as the use of the parameters for degrees $\leq p$ only, since all coefficients for degrees $> p$ will be zero. But the estimates of these coefficients will generally not be zero, which entails larger MSEP. By a continuity argument it can be seen that we can have non-zero true coefficients for high degrees, but still get larger MSEP with estimated coefficients for these high degrees than with the use of zero coefficients for them.

The proper value of $w_0$ depends on the degree of dependence of the development distribution on $W$ and the number of claims per $W$. A strong dependence speaks for a large $w_0$. So do large numbers of claims for large $W$ by decreasing the parameter estimate variances.

The simulated claim set of Example 3 in Section 7.5, p. 310, has the development distribution independent of $W$. Here $n = 12$. Since $\mathrm{P}(W \leq 4) = 1$, i.e. the highest $W$-value is 4, and there is no tail, we have $\mathrm{P}(L \leq 9) = 1$ for the life length $L$. In the paper I set $w_0 = 1$ for this example. No claim can have a life length $L = n - w_0 + 1 = n = 12$. If we change this to $\mathrm{P}(L = 12) > 0$, the obvious solution is to include more periods by setting $n \geq 15$, e.g. $n = 24$ to even out seasonal variations if period length is one month.

In this context the question arises whether the expression (A.32) for the IBNR reserve for claim period $i$ is consistent with the breakdown into development periods given by expression (A.33). I give a derivation showing that this is indeed so.

**Proposition**. *With* $\widehat{R}_i^{\mathrm{I}}$ *given by* (A.32) *and* $\widehat{R}_{ij}^{\mathrm{I}}$ *given by* (A.33), *it holds*

$$\widehat{R}_i^{\mathrm{I}} = \sum_{j=n-i+2}^{n} \widehat{R}_{ij}^{\mathrm{I}}.$$

*Proof.* By (A.33)

$$\sum_{j=n-i+2}^{n} \widehat{R}_{ij}^{\mathrm{I}} = \sum_{j=n-i+2}^{n} \sum_{w=n-i+2}^{j} \sum_{\lambda=j-w+1}^{n} \widehat{A}_{iw}\, \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,j-w+1}(1, w\wedge w_0, 0).$$

Change the summation order in the two first $\sum$ of the right side. (Draw a graph depicting the values of $j$ and $w$ contributing to the sum.) Namely

$$\sum_{j=n-i+2}^{n} \sum_{w=n-i+2}^{j} = \sum_{w=n-i+2}^{n} \sum_{j=w}^{n}.$$

Hence

$$\sum_{j=n-i+2}^{n} \widehat{R}_{ij}^{\mathrm{I}} = \sum_{w=n-i+2}^{n} \widehat{A}_{iw} \left( \sum_{j=w}^{n} \sum_{\lambda=j-w+1}^{n} \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,j-w+1}(1, w\wedge w_0, 0) \right).$$

Change the summation order in the inner double sum. Make the variable substitution $h = j - w + 1, \;\; j = h + w - 1$.

$$\begin{aligned} &\sum_{j=w}^{n} \sum_{\lambda=j-w+1}^{n} \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,j-w+1}(1, w\wedge w_0, 0) \\ &= \sum_{\lambda=1}^{n} \sum_{j=w}^{n\wedge(\lambda+w-1)} \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,j-w+1}(1, w\wedge w_0, 0) \\ &= \sum_{\lambda=1}^{n} \sum_{h=1}^{n\wedge\lambda} \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,h}(1, w\wedge w_0, 0) \\ &= \sum_{\lambda=1}^{n} \sum_{h=1}^{\lambda} \hat{p}_\lambda(1, w\wedge w_0, 0)\hat{\mu}_{\lambda,h}(1, w\wedge w_0, 0). \end{aligned}$$

By (A.28) we recognize the last expression as $\widehat{R}(1, w\wedge w_0, 0)$. Thus

$$\sum_{j=n-i+2}^{n} \widehat{R}_{ij}^{\mathrm{I}} = \sum_{w=n-i+2}^{n} \widehat{A}_{iw} \widehat{R}(1, w\wedge w_0, 0) = \widehat{R}_i^{\mathrm{I}}.$$

□

### B.2. Bootstrap ex-post reserve as object reserve

I held a lecture at the Swedish Actuarial Association on 2013-10-22, describing my article Bootstrapping Individual Claim Histories, ASTIN Bulletin 2012(1), 291-324. A question was asked whether the mean bootstrap ex-post reserve

$$R_i^{(-)} = \frac{1}{B}\sum_{t=1}^{B} R_i^{(\nu_t)}$$

given in expression (6.8) on p. 304 could be used as object reserve. It could, but its expectation and almost sure limit as $B \to \infty$ would be better. And that can be computed. It is a PPCF (expected *Payments Per Claim Finalized*) reserve, since all claims in the set $Z$ used for bootstrap must be finalized. Furthermore, they must all belong to claim periods where all claims are finalized. Newly occurred and finalized claims cannot be used. Thus the MSEP will be larger than necessary. In contrast, the chain ladder, Schnieper and RDC methods treated in the paper all use payments on non-finalized claims as well.

To compute this PPCF reserve, consider the historic claim set $Z$ with $K$ claims defined in Section 2.1, p. 294, and pretend that they all occurred in claim period $i \in \{1, \dots, n\}$. For simplicity, assume no segmentation and no inflation. Also assume that $G^{(\nu)}$ of Assumption A5 is the whole sample space of experiment $\nu$, so that $\nu_t = t$. Define

$$E_i = \sum_{r=1}^{K} \sum_{j>n-i+1} Y(r,j) = \text{total ex-post reserve in } Z,$$

$$K_i = \sum_{r=1}^{K} \mathbf{1}_{\{W(r) \le n-i+1\}} = \text{number of claims reported 'now' in } Z.$$

In the BICH bootstrap procedure we draw a random number $N_i^{(\nu)}$ of claims until $M_i$ are reported 'now'. Here $M_i$ is the number of object claims reported now, see (2.4) p. 295. Then obviously $N_i^{(\nu)}$ has the negative binomial distribution NB$(M_i, K_i/K)$, with mean $M_i K/K_i$.

For each randomly drawn claim, pretended to have occurred in claim period $i$, its expected ex-post reserve is of course $E_i/K$. Since $N_i^{(\nu)}$ is a stopping time (see below for definition), by Wald's theorem we thus obtain

$$\mathrm{E}[R_i^{(-)}] = \mathrm{E}[R_i^{(\nu)}] = \mathrm{E}[N_i^{(\nu)}]E_i/K = \frac{M_i K}{K_i}\frac{E_i}{K} = \frac{M_i E_i}{K_i}.$$

This is the PPCF reserve that could be used. It is unbiased but not suitable. The bootstrap sample $Z$ of claims is just a sample of the real claim distribution.

By $N_i^{(\nu)}$ being a stopping time is meant that the event $\{N_i^{(\nu)} = k\}$ is independent of the sequence of random drawings (if they were to continue) n:os $k{+}1,\ k{+}2,\ k{+}3 \dots$.

**A simple account**

Pretending that all claims in $Z$ occurred in period $i$, we can compute the remaining payment sum $E_i$ and the number $K_i$ of claims reported 'now', i.e. known. Then $E_i/K_i$ is the mean reserve per known claim. Hence

$$\text{(number of known object claims)} \times \text{(mean reserve per known claim)} = M_i E_i / K_i$$

is the object reserve with this PPCF method.

**Available claims**

There are not more claims available for these reserves, since we could include the claims used for bootstrap also in the set $T$used for object (actual) reserves. I state on p. 294, line 4 from bottom, that "The claims for $i$ that are finalized could be part of $Z$." The converse is also true, namely that the claims of $Z$ could be part of $T$. We do not always, however, want to include all claims of $Z$ in $T$, due to practice changes.

### B.3. Addendum to paper

In (6.10) I gave the standard deviation estimate

$$\widehat{D}[R_i^{(\nu_1)}] = \sqrt{\frac{1}{B-1}\sum_{t=1}^{B}\left(R_i^{(\nu_t)} - R_i^{(-)}\right)^2}.$$

Since the mean is known, a better estimate is

$$D^*[R_i^{(\nu_1)}] = \sqrt{\frac{1}{B}\sum_{t=1}^{B}\left(R_i^{(\nu_t)} - \frac{M_i E_i}{K_i}\right)^2},$$

if $G^{(\nu)}$ of Assumption A5 is the whole sample space of experiment $\nu$.

### B.4. From (A.6) to (A.9)

Suppress $q, w, t, L > t$ in the notation. They do not change from (A.6) to (A.9). Now (A.3) gives

$$H(L) - H(t) = \sum_{h=t+1}^{L} Y(h+W-1).$$

Then

$$R = \mathrm{E}[H(L) - H(t)] = \sum_{h=t+1}^{L} \mathrm{E}[Y(h+W-1)],$$

$$p_\lambda = \mathrm{P}(L = \lambda),$$

$$\mu_{\lambda h} = \mathrm{E}[Y(h+W-1) \mid L = \lambda].$$

We want to show

$$R = \sum_{\lambda=t+1}^{n} \sum_{h=t+1}^{\lambda} p_\lambda \mu_{\lambda h}.$$

A change in the order of summation gives

$$\sum_{\lambda=t+1}^{n} \sum_{h=t+1}^{\lambda} p_\lambda \mu_{\lambda h} = \sum_{h=t+1}^{n} \sum_{\lambda=h}^{n} p_\lambda \mu_{\lambda h} =$$

$$\sum_{h=t+1}^{n} \sum_{\lambda=h}^{n} \mathrm{P}(L = \lambda)\mathrm{E}[Y(h+W-1) \mid L = \lambda] =$$

$$\sum_{h=t+1}^{n} \sum_{\lambda=1}^{n} \mathrm{P}(L = \lambda)\mathrm{E}[Y(h+W-1) \mid L = \lambda] =$$

$$\sum_{h=t+1}^{n} \mathrm{E}[\mathrm{E}[Y(h+W-1) \mid L]] = \sum_{h=t+1}^{L} \mathrm{E}[Y(h+W-1)],$$

where the change from $\lambda{=}h$ to $\lambda{=}1$ from the 3:rd to the 4:th equation member, and the change from $n$ to $L$ from the 5:th to the 6:th member, can be done since $Y(h+W-1) = 0$ for $L < h$. ■

### B.5. Modification of quantile intervals for RDC

In Rapp I have changed the grouping into quantile intervals so that each group per $q, t, w$ contains at least one finalized claim. If a group initially contains only open claims it is merged with the group with the same $t, w$ and nearest lower $q$ with finalized claims, if such a group exists. If not, it is merged with group with the same $t, w$ and the nearest higher $q$ with finalized claims. If not even such a group exists, $Q_t$ is set to 1.

## B.6. Corrections of bugs in Rapp concerning RDC

Heartfelt thanks are due to (in alphabetical order) Ari Dwi Hartanto, Mujiati Dwi Kartikasari and Rinjani Pebriawan at Universitas Gadjah Mada, Yogyakarta, Indonesia, for exposing bugs, thereby enabling me to correct them. The latest of them was corrected 2016-04-28. It was an error in the implementation of equation (A.36).