Rapp

# The reserving methods BICH and RDC

Lecture in Indonesia november 2014
Stig Rosenlund

The methods of this talk are implemented in the programming language Rapp. To find Rapp on the internet, search **free actuarial language.**

About me: PhD in Mathematical Statistics at the University of Göteborg in 1983. Actuary at the Swedish insurance company Länsförsäkringar 1983-2010. After retirement in 2010, I have resumed my research, now focusing on insurance mathematics. So far I have in the 2010's had three papers published - two in tariff analysis and one in reserving.

Please interrupt me with questions at any time!

**BICH = Bootstrapping Individual Claim Histories**

**RDC = Reserve by Detailed Conditioning**

The methods are described in

Rosenlund, S. (2012).
Bootstrapping Individual Claim Histories.
*ASTIN Bulletin* **42**(1), 291-324.

It is included in the manual as an appendix.

I will produce results in real time for the methods. The input data are two fictitious simulated claim files, one with payments and with changes in known claim cost. Both have 1,710,629 lines with 75 characters per line.

BICH performs stochastic reserving ie calculates

**MSEP = Mean Square Error of Prediction**

for some primary reserving method. By primary I mean one that makes reserve point predictions. BICH is implemented in Rapp for these primary methods:

- Benktander (alias Benktander-Hovinen)
- Bornhuetter-Ferguson
- Cape-Cod
- Chainladder
- RDC
- Schnieper

Terminology according to

Wüthrich, M. V. and Merz, M. (2008).

*Stochastic Claims Reserving Methods in Insurance.*

424 pages + preface. Wiley.

Benktander, Bornhuetter-Ferguson and Cape Cod are special by requiring premiums or àpriori predictions for final claim cost to be able to be better than Chainladder. How close to the expected outcomes should the àpriori predictions be in the bootstrap? A subjective element risks being introduced into the calculation. There is a need for further simulation conditions that can be verified objectively. I do not treat those methods today.

Besides stochastic reserving with a given method, you can with

BICH compare different methods with respect to the MSEP and then choose the one with minimum MSEP - the best. Useful even if you do not need to report MSEP.

In Schnieper's method the payment increments, alternatively changes in known claim cost for the period i and development period j, are separated:

- from new claims reported during j
- from claims reported before j

The method requires exposure per claim period, but I indicate in my paper that the number of claims reported in the first period of development is better as exposure than the number of policies or premium volume for the cases simulated in paper. Namely, with many claims reported in the first reporting period. Thus one can apply Schnieper with access to only claim data.

In my simulations Schnieper has been better than Chainladder, but one can imagine situations where Schnieper gives overparametrization with poorer results than Chainladder.

With overparametrization I mean that you have parameters in the model that would not increase and may decrease the MSEP *if they were known*, but where the estimates of them have so large variances that MSEP increases.

My new method RDC turned out in my simulations to be for the most part better than Chainladder and Schnieper. The exception was a case that fully meet the Mack Chainladder conditions, where Chainladder (which then equals Schnieper) was slightly better than RDC. Barely noticeable, but significant. (Table 4, page 313 in the paper.) It can be seen as overparametrization in RDC, since RDC has many parameters while Chainladder has few. Other cases, not satisfying some special conditions like Mack's, can surely also be found that give disadvantage to RDC due to overparametrization.

But the clever thing with BICH is that you can identify those situations through method comparisons.

Furthermore, you can vary the number of parameters to be estimated in

RDC, and with the help of MSEP comparisons choose an appropriate number. For example, you condition in RDC with respect to quantile intervals for hitherto paid. With 500 quantile intervals you get a very detailed conditioning. If you put the number of quantile intervals to 1, you do not condition at all with respect to hitherto paid.

## Conditions and methods for BICH and RDC

The conditions are that the claims must be independent and identically distributed after possible inflation adjustment and breakdown by different homogeneous groups which I call segments. Also the number of claims reported in the initial development period should be sufficient, e.g. at least a few hundred, for each claim period. And for BICH a mathematical condition, which I call A5, which provides a selection criterion implying that we get bootstraps of historical settled claims that are similar to the claims to be reserved, in a sense.

<u>The approach in BICH</u>

A collection Z of historical claims from entirely settled claim periods is available with reporting date, claim date and all pay dates and payment amounts. Furthermore, we have a collection of current claims to be reserved. .

For each claim period i we do the following: Draw a random claim with replacement (= bootstrap) from Z. Move the claim period ahead so that we pretend that it occurred in claim period i. We note if it would then be reported at the present time point. Repeat this until the number of pretended claims reported now equals the number of reported claims now in the current collection for this claim period i. In addition, we have gotten a feigned number of claims that not would have been known and reported at the present time point.

When we're done with this, we have a bootstrapped image of the current claim collection. So we check with the selection condition according to A5 if we shall keep or discard the image. If it is retained, we calculate reserves on it per claim period, with e.g. Schnieper or RDC. Moreover, and this is the clever part, we can calculate what the actual pretended outcome was because everything is known for the pretended claims. And so we take the difference between them and square it.

Then we do all this again, for example B = 10,000 times excluding the wasted images. (That's the number of repetitions used in my paper.) Every time we have a squared difference whose mean we calculate over the 10,000 times. And it is our MSEP estimator. It is reported in Rapp's BICH procedure and furthermore a confidence interval for the estimate is reported. Plus some other quantities, among others one that shows if the reserving method (eg Schnieper) can be assumed to be unbiased. The exact formulation of this hypothesis is H0 in the paper page 296.

## The method RDC

It is to calculate the conditional expected value of the remaining payment amount (or remaining change of known claim cost. ie paid amount plus claimshandler's reserve) given a certain segment and

- hitherto paid
- delay between the occurrence and reporting

Two main steps for this. Lifetime is the time between reporting and settlement.

1. Compute an estimate of the probability distribution of lifetime with the conditioning above.
2. Compute an estimate of the expected value of the remaining payment amount per future development period with the same conditioning and additionally conditional on lifetime.

Then we can multiply together the results from 1 and 2 and summarize, and so we get a reserve.

The probability distribution of the lifetime is calculated using survival probabilities chained together for lifetime probabilities (the Kaplan-Meier estimator).

Step 2 is extremely complicated, because I use payments on both settled and open claims. It has to be difficult to not give negative bias. To use only payments on settled claims is much simpler but gives greater MSEP, in a described case much larger. (The simpler method is very old and is called in the literature PPCF = expected Payments Per Claim Finalized.) Do not even start reading this section of the paper, if you do not have very good patience and much time!

The conditioning with respect to hitherto paid is approximated by a quantile interval partitioning, as I mentioned above, with any number of intervals.

The conditioning with respect to the delay between occurrence and reporting can be reduced or abolished, if it is found that it does not reduce the MSEP.

The claims need to be adjusted for inflation. If not, you can enter in Rapp a file with price index to make the inflation adjustment.

The calculation gives a statistical reserve for each individual reported claim and an IBNR reserve - reserve for Incurred But Not Reported claims - per claim period. You can specify a file where the reserves for each claim shall be written fall, and then you get reports in Excel with IBNS divided on IBNR- and RBNS reserves per segment and claim period. The sum of statistical reserves is the RBNS reserve for Reported But Not Settled claims.

**Combining RDC and GLM**

It is natural to use a GLM model (log link) to get factors for a statistical reserve for a collection of background variables, such as diagnosis, customer's age, etc. A natural combination of RDC and such a GLM model is to calculate reserves both ways and multiply GLM reserves by a factor per segment and claim period to get the sum of them for each segment and claim period equal to the RDC/RBNS reserve per segment and claim period. Rapp's menu Reserves has an application for this.

Thus. BICH can

- Give MSEP for forecasts under weak conditions.
- Compare primary reserve methods.

If the claims are not independent and identically distributed after inflation adjustment and segmentation, then MSEP as such for forecasts can become quite unusable. But the relationships between the MSEPs for the reserve methods Chainladder, RDC and Schnieper can still be about right, so that the one that according to BICH is best also probably is de facto best.

Now

- Bootstrap results with BICH for Schnieper's method are produced.
- RDC reserves are computed, including statistical reserves per claim with RDC/GLM combination. MMT (the marginal method) is used.

**BICH-runs**

| Demo name | Method | Outnamepart | Time for B = 2000 |
|---|---|---|---|
| DemoChl | Chainladder | C | c:a 20 seconds |
| DemoSch | Schnieper | S | c:a 20 seconds |
| DemoRDC | RDC / w0 = 1 | R | c:a 6 minutes |

Choice S1 = partial segmentation in the menu.

w0 = 1: No conditioning with respect to the delay between occurrence and reporting. With maximum such conditioning RDC gave much worse results than other methods.

B is the number of times that an entire set of claims for each claim period is drawn.

With the simulated claim files, and certain selection criteria in the demo parmfiles, B = 2000 means that some claim is used for reserve calculation about 225 million times.

DemoSch is run as an example. Because of the long execution time (9 minutes), previously computed results are shown for RDC with BICH. But a single calculation of primary reserves with RDC is fast.

There are many parameters in the menus. Among other things, an extensive apparatus for discounting and inflation adjustments. And for tail computation. All is explained exhaustively, in the manuals and in the menu info boxes.