# A three-way credibility approach to loss reserving

G. VENTER
The *Workers Compensation Reinsurance Bureau, New York, NY 10038, USA*

Two ways of estimating ultimate incurred losses are loss based and exposure based. The former uses a factor applied to losses emerged to date; the latter bases the loss estimate on the exposure, and ignores emerged losses. A hybrid is the Bornheutter-Ferguson method, which adds the emerged losses to an appropriate percentage of the exposure based estimate. This paper considers a weighting of these three estimators using credibility concepts. It generalizes the results of I. Robbin's paper, 'A Bayesian Credibility Formula for IBNR Counts' in PCAS 1986.

***Keywords:*** Loss reserving, Credibility concept,

## 1. Introduction

Let N denote the ultimate incurred loss being estimated, M be the losses emerged to date, and ***R* = *N* - *M*** be the losses yet to emerge. Here 'emerged' could be taken as 'paid' or 'reported', for example. It is assumed that every year a vector of parameters u is drawn from some fixed distribution, and that ***R*** and M are then drawn independently from distributions determined by u. The following conditional moments are also specified:

n(u)=E(NIu),
q(u)=E(RIu)+E(NIu),
s2(u) =Var(Mju).

For notational simplicity, the '(u)' will usually be omitted. Independence of n and ***q*** is also assumed.
Thus, for instance, EM = EE(M 1 u) = E(n - ***nq)*** = E(1 - q)En.

The problem will be formulated as estimating ***R*** having observed M. Three possible estimators are::

(i) development factor estimate: M ***X*** (ER/EM);
(ii) exposure estimate: EN - M;
(iii) Bornheutter-Ferguson estimate: EN - EM.

The credibility weights for these three estimators will turn out to be proportional, respectively, to the three variance components of Var M below:

Var M= Var(n)[E(l - ***q)12*** + Var(q)E(n') + E(.s2)
=t
2
+ u2 + w2.

These components derive from the usual expression
VarM=VarE(Mlu)+EVar(Mlu)
=Var[n(l-q)] +w2,
and the variance product formula for independent random variables Var XY = Var XE( Y2) + Var Y(EX)2.

The linear estimate of *R* from M having the smallest expected squared error is, by general least squares theory,

R=ER+ (M-EM)Cov(R, M)
Var M

EM and Var M have already been expressed in terms of the moments of q, n, and s2. This is done for *Cov(R, M)* as follows:

*Cov(R,* M)=ECov(R, Mlu)+Cov[E(Rlu),E(Mlu)]
= 0 + Cov[nq, **n(1 -** q)] (by conditional independence)
= E(n2)E((I - q)q) - (En)2E(I - q)E(q) (by definition of covariance)
- E( n2)(1 - q)Eq + E( n2)E(1 - ***q)Eq*** (equals zero)
= E( n')[ (Eq)2 - E( q2)] + Var( n)E(l - ***q)Eq***
= ***-u2*** + ***t2Eq*** + E(1 - ***q)***
= ***-u2*** + ***t2ER*** + EM (multiplying ratio by EN).

Plugging this in above then gives

R= *(t2+u2+w2)ER+(M-EM)(-u2+t2ER+EM)*
*t2* + *u2* + *w2*
*(t2* + *u2* + w2)ER + *t'M[ER* + *EM]* - *t2ER* + u2EM - *u2M*
=
*t2* + *Id2* + *w2*
= *t'M[ER* + EM] + u2[EN- *M]* + w2(EN- EM)
t2 + u2 + w2

f

which is the credibility estimator for *R*.

**2. Discussion**

If Var(n) is high, then EN is not a particularly good estimator of N. In this case, t2 is higher, and so the development factor estimate gets more weight. This is the only estimator that does not use EN. If Var(q) is higher, then there is less stability in the development pattern, and u2 is higher, placing more weight on the exposure based estimator, which does not use the development factors. Finally, if w2 is higher, then ***M*** is less representative of its conditional mean, and so less useful as a predictor of ***R.*** Thus

the Bornheutter-Ferguson estimator, which does not use *M*, receives more weight.

3. **Parametric example**

Each year parameters ***a, b,*** and c are drawn, i.e., u = ***(a, b, c),*** and ***a*** is independent of ***b*** and c. Here 0 < ***a -C*** 1 and ***b, c*** > ***0. R*** is then gamma distributed in ***b, ac*** and ***M*** is gamma in ***b,*** (1 - ***a)~,*** with *R* and ***M*** drawn independently given these parameters. Then N is gamma in ***b, c,*** and so n = ***bc.*** It can then be verified that ***q*** = ***a,*** and ***q*** is thus independent of n, and that s2 = (1 - ***a)b2c.***

**4. Estimating parameters**

The components t2, u2, and w2 are estimated below under some fairly strong assumptions, most of which can probably be relaxed with appropriate modifications. First, the parametric model above will be

assumed, although not all aspects of it are needed. Second, it is assumed that the ultimate losses N are known for several older years, and that the process has stability, so that the parameters for the different

years can be thought of as arising from the same distribution. Finally, it is assumed that estimates of q are

available for each of the older years, and so the distribution of q has already been estimated.

For the older years attention is focussed on one particular stage of development, say second report, and

M will thus denote the emerged losses at that point. The exposure estimate of N is taken as EN, so the sample variance can be computed. By the model this can be expressed as

E(N-EN)'=EVar(NIu)+VarE(Nlu)=E(b*c)+Var(n).

Using l/l - q as the development factor, the squared difference between N and the developed losses can be computed for each year, and the average taken over the older years available. This average then becomes an estimator of

E(N-&)'=ZE+-_E(b*c) (equality demonstrated below).

Combining these estimators, gives an estimate for Var(n). This can be negative, and if so Var(n) is estimated to be 0. In this model, E(s*) = E(1 - q)E( b*c), so an estimate for E(s*) is also provided.

The development of the above equality is as follows:

E(N-&)2=E(N2-2M(ER) + (I!;)2 N2-2M(fl_tqR) + (I:;)*

***(l-a)b2(c+(1-a)C*)+a(l-a)b2C2***

***l-a***

+ (l-a)b*(C+(1--a)c*)

(1 - 0)" I

+ ***b*c**** , l = 2E &E(b'c).

**5. Alternate estimation**

The above estimation can probably be made more practical, for instance by using the known development factor l/E[l - q]. An easier method is to use regression of actual emerged losses against the

three estimates, using the data from the older years. The regression coefficients, which are interpreted as

the three credibilities, can be made to sum to unity by taking them to be of the form ***a/(a + b + c),*** ***b/( a + b + c),*** and c/( ***a*** + ***b*** + ***c).*** In fact, any multiple of (a, ***6, c)*** will give the same credibilities, so it suffices to set c = 1. Rather than regressing ***R*** against the three estimates of ***R, N*** can be regressed against

M plus the three estimates with the same results. These three estimates of N will be denoted by Ni, N2 and N3_ Adding a subscript i for year observed, then, gives the regression formulation:

N;= a NE +
***b***
***a+b+l a+b+l N,, + a+i+lN,i+ci.***

A non-linear regression package could be used to find the values of ***a*** and ***b*** that minimize Cc:. Alternatively, the partials with respect to ***a*** and ***b*** can set to zero, yielding the following equations for ***b*** in

terms of ***a.*** Here, X=N-N,, Y=N-N,, Z=N-N,, U=N,-N,, v=N,-N,, and w=N,-N,, where the subscript ***i*** is omitted but implied.

b= _ ***a*CXU+aC(ZZJ-XV)- CZV***
***aCYU- CYV***

These can be equated, and solved for a using a root finder method (e.g., secant rule). If no non-negative real root exists, ***a*** or b can be set to zero, and the other solved for. The equations are b = ***-CZV + CYV*** and ***a = -CZW + CXW.*** These can be compared to see which minimizes CC:. An advantage of this method is that N can be the latest estimate, rather than actual losses, for the old years. In this case, a weighted regression could be used, where the weights are inversely proportional to the variance of the latest estimates. Another possible advantage is robustness: the weights derived by regression may work well in practice, even if some of the model assumptions are violated. If there is a changing pattern over

Table 1
Workers' Compensation Reinsurance Bureau accident year losses as a percentage of the exposure base.

| Act year | (1) actual |
|---|---|
| 60 | 5.2 |
| 61 | 3.8 |
| 62 | 5.0 |
| 63 | 4.2 |
| 64 | 5.0 |
| 65 | 3.2 |
| 66 | 5.1 |
| 61 | 4.1 |
| 68 | 4.9 |
| 69 | 7.3 |
| 70 | |
| 71 | |
| 72 | |
| 73 | |
| 74 | |
| 75 | |
| 76 | |
| 77 | |
| 78 | |
| 79 | |
| 80 | |
| 81 | |
| 82 | |
| 83 | |
| 84 | |
| 85 | |
| 86 | |
| 87 | |

5.3
8.9
9.0
9.0
11.1
12.3
15.2
9.2
10.8
11.4
9.1
6.7
8.3
7.5
9.0
11.3
8.2
9.9
(2)
exposure
base
estimate
(3)
(2) - (1)
(4)
estimate
developed
from 1st rpt
(5)
(4) - (1)
(6)
estimate
developed
from 7th rpt
(7)
(6)-(l)
(8)
estimate
developed
from 15th rpt
(9)
(8) - (1)
5.1 -0.1
4.2 0.4
4.4 -0.6
4.1 0.5
4.8 -0.2
4.2 1.0
4.6 - 1.1
5.0 0.9
5.4 0.5
6.2 -1.1
6.8 1.5
1.4 -1.5
8.8 - 0.2
9.7 0.7
10.9 -0.2
12.8 0.5
14.1 - 1.1
9.7 0.5
10.6 -0.2
11.6 0.2
9.4 0.3
1.4 0.7
8.1 - 0.2
8.8 1.3
9.5 0.5

10.0 -1.3
10.4 2.2
10.8 0.9
8.8 3.6
6.9 3.1
4.9 -0.1
6.5 2.3
5.4 0.4
1.4 -1.8
3.6 -2.1
2.4 - 1.7
4.9 0.0
4.0 - 3.3
4.1 -1.2
8.2 -0.7
7.0 -2.0
10.8 1.8
15.2 4.1
11.2 -1.1
15.9 0.7
5.1 - 4.1
10.7 -0.1
11.3 -0.1
8.3 -0.8
8.2 1.5
11.2 2.9
1.9 0.4
6.0 -3.0
9.9 -1.4
1.3 -0.9
6.9 1.7 5.5
4.5 0.7 4.3
4.6 -0.4 5.2
5.0 0.8 4.1
6.1 1.1 5.4
2.5 -0.7 2.1
3.6 -2.1 4.1
5.0 0.9 4.0
4.8 -0.1 5.3
5.3 -2.0 6.4
4.9 -0.4 5.1
8.1 -0.8 9.1
6.7 -2.3 9.5
1.3 -1.7
12.5 1.4
12.4 0.1
16.2 1.0
9.7 0.5
12.2 1.4
12.0 0.6
9.8 0.7
0.3
0.5
0.2
0.5
0.4
-0.5
-1.6
-0.1
0.4
-0.9
-0.2
0.8
0.5
Average of absolute value
of difference 0.7 1.7 1.0 0.5

Table 2
Credibility weights - regression estimate.
Development Bomhuetter-
factor Ferguson
Exposure
1st Report
2nd Report
3rd Report
4th Report
5th Report
6th Report
7th Report
8th Report
9th Report
10th Report
11 th Report
12th Report
13th Report
14th Report
15th Report
16th Report
17th Report
18th Report
19th Report
0.30
0.31
0.28
0.28
0.33
0.00
0.00
0.45
0.64
0.46
1.00
0.57
0.75
0.89
0.24 0.66
0.17 0.83
0.37 0.45
0.18 0.82
0.16 0.84
0.41 0.29
0.40 0.30
0.22 0.49
0.27 0.45
0.31 0.36
0.40 0.60
0.50 0.50
0.28 0.27
0.00 0.36
0.38 0.16
0.00 0.00
0.43 0.00
0.00 0.25
0.00 0.11
**O.OOOJ**
60 61 62 63 64 65 66 67 68 69 70 71 72
- ACTUAI - EXP.EST -- RPT 1 EST .... RPT 7 EST .~~~~~~' RPT 15 EST
Fig. 1. WCRB - Accident year losses, 1960-1972 (actual vs. exposure and loss development methods).
0.18
1'~
0.06
I'
0.04
0.02
0.00.
73 74 75 76 77 78 79 80 ai a2 a3 84 a5 86 a7

– ACTUAL –'– EXP. EST - - RPT 1 EST .... RPT 7 EST

Fig. 2. WCRB - Accident year losses, 1973-1987 (actual vs. exposure and loss development methods).

08 ""'

-l 0 /,,/:,/,j/,,,,,,,,,,,,,,,,,

60 62 64 66 68 70 72 74 76 78 80 82 84 86

61 63 65 67 69 71 73 75 77 79 81 83 85 87

ACCDENT YEAR

Fig. 3. WCRB - Accident year losses, 1960-1987 (actual vs. exposure and loss development methods).

0.18

0.16

0.14

0.12

_/

0.06

0.04

a

2

B

-.l 01, , , , , , , , , , , , , , , , , , , , , , , , , , , , ,

60 62 64 6.6 68 70 72 74 76 78 80 82 84 86

61 63 65 67 69 71 73 75 77 79 81 t33 85

AWDENTYEAR

Fig. 4. WCRB - Accident year losses, 1960-1986 (actual vs. loss development niethod at first report).

.......... ACT"AL

e D., ............

..........

..........

............

......................

-l 0 ..],l,,.,,,,,,,l,,,,/,

60 62 64 66 68 70 72 74 76 78 80

61 63 65 67 69 71 73 75 77 79

ACCOEM YEAR

Fig. 5. WCRB - Accident year losses, 1960-1980 (actual vs. loss development method at seventh report),

gq==J 1

l! a12

0.1

P

F 0.08

2 0.06

a

L!J 0.04

VI

!j 0.02

1 - 1

IRPTI~EST ]

- 0' , , , , , , , , , , , , , ,

60 62 64 66 68 70 72

61 63 6.5 67 69 71

ACZDENTYEAR

**Fig. 6. WCRB - Accident year losses, 1960-1972 (actual vs. loss development method at 15th report).**

time, a non-weighted regression may be reasonable, in that the later years, though less certain, may be more relevant.

An example illustrating this method is included. The losses are workers compensation excess, taken as a ratio to an exposure base (Table 1). Credibilities are estimated by regression for estimates at 1st through 19th report of losses (Table 2). Graphs are included illustrating the accuracy of the exposure estimates (Figures l-3) and the development factor method at lst, 7th, and 15th report (Figures 4-6). The resulting

credibilities illustrate a problem with this method: there is no particular connection between estimates at
different maturities.