General Re-Rating Formula

Muhamed Borogovac, Ph.D.

Enterprise Risk Management Symposium
Society of Actuaries

Chicago, IL

March 28-30, 2007

Abstract

In this paper, a re-rating formula which calculates proposed rates directly by means of fully developed and trended loss costs and current class differentials has been introduced.

The new formula eliminates the need to calculate the following, previously necessary variables:

- the overall rate change,
- classification parameters' loss ratios,
- proposed differentials that apply to the rate classification parameters
- balance back factor.

Surprisingly, to calculate new rates by our formula (18) we don't even need to know the current rates.

This model significantly reduces the amount of data manipulation in the P&C re-rating area, as well as in group insurance. As rates are measures of risk of an insurance product, the formula (18) can be considered as a risk management tool in the above practice areas.

1. Introduction

A rate change involves the following **three-step process**:

Step 1. Determine the overall average rate change,
Step 2. Change all differentials according to the experience projected to the future period of new rates,
Step 3. Balance back to overall rate change indicated, see [1].

The objective of this paper is to replace the above process with a single formula for the proposed rates, given the experience data inputs. This is formula (18) (i.e., formula (17) in the case when the number of classification parameters is n=3).

The idea of the proof of formula (18) (i.e., (17)) is to create a somewhat complicated "**one-step model.**" In that model all variables are properly introduced into a single formula. We find that some of the variables cancel out, leaving us with a relatively simple formula (17) (i.e., (18)) for the proposed rates $\bar{r}_{ijk}$. The formula calculates exactly the same proposed rates as the above **three-step process**.

2. Notation and Derivation of the Formula

In order to simplify derivation, let us limit ourselves to only three rate classification parameters—**Territory, Class** and **Industry**—which is not a loss of generality. Let $L=(l_{ijk})_{m\times n\times p}$ denote the tensor (i.e., three-dimensional array) of **fully developed and trended losses (FDTL)**. We usually obtain that tensor by multiplying the **tensor of current losses** by the **development factor** and the **trend factor**.

The element in the cell (i,j,k) represents a projected loss (i.e., FDTL) in the Class "i", Territory "j", and Industry "k", where i=1,2,…,m, j=1,2,…n, k=1, 2, …, p. Let us denote by L the total of FDTL, i.e.

$$L := \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} l_{ijk}$$

By the same token let us introduce the following notation:

Tensor E = $(e_{ijk})_{mxnxp}$, where e_{ijk} is the number of exposure units in Class "i", Territory "j" and Industry "k".

Tensor R=$(r_{ijk})_{mxnxp}$, represents current manual rates in the cells (i,j,k), i=1,2,…,m, j=1,2,…n, k=1, 2, …p.

In this setting we have two sets of differentials. Let the vectors X = (x_1, x_2, … , x_m), T = (t_1, t_2, …, t_n), and F = (f_1, f_2, …, f_p) represent **current** Class, Territory and Industry differentials, respectively, and let the vectors Y = (y_1, y_2, … ,y_m), U = (u_1, u_2, … ,u_n), and G = (g_1, g_2, …, g_p) represent **proposed** Class, Territory and Industry differentials, respectively.

In this notation, **base rate** = r_{111}, x_1=t_1= f_1=1, and $r_{ijk}=r_{111}*x_i*t_j*f_k$.

We will need the following equations in the proof:

$$x_i := \frac{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{ijk}e_{ijk}}{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{1jk}e_{1jk}} = \frac{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{i11}t_j f_k e_{ijk}}{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{111}t_j f_k e_{1jk}} = \frac{r_{i11}}{r_{111}} . \tag{1.1}$$

Similarly $$t_j := \frac{r_{1j1}}{r_{111}}, \tag{1.2}$$

and $$f_k := \frac{r_{11k}}{r_{111}} . \tag{1.3}$$

As every unit of exposure in the cell (i,j,k) is expected to cost t_j*f_k times the cost of the unit of exposure in the cell (i,1,1), we can introduce the following concept:

$$E_i^C = \sum_{j=1}^{n}\sum_{k=1}^{p} e_{ijk}t_j f_k \text{ , i=1, 2, …, m.} \tag{2.1}$$

Similarly,

$$E_j^T = \sum_{i=1}^{m} \sum_{k=1}^{p} e_{ijk} x_i f_k \text{ , j=1, 2, …, n.} \quad (2.2)$$

$$E_k^I = \sum_{i=1}^{m} \sum_{j=1}^{n} e_{ijk} x_i t_j \text{ , k=1, 2, …, p.} \quad (2.3)$$

A remark about generalization: If N>3 denotes the number of rate classification parameters (e.g., Class, Territory, etc.) then we would have N relations (2.1), (2.2), …(2,N) rather than 3, one for each E^1, E^2, …, E^N. The summation on the right-hand side of each equation would go over N-1 indexes, i.e., there would be N-1, rather than 2, sums on the right-hand side of each of the equations (2.1), (2.2), , (2.N). The summation on the right-hand side would go over all indexes, except the one that represents the index of the particular classification parameter. For example, the summation in the relation (2.2) runs over all indexes except index "j" which corresponds to the Territory classification parameter. Similar remarks about generalization hold for the following relations (3) and relations (5).

In the general formulae, we consider Industry to be N^{th} rate classification parameter, i.e. we consider that the additional rate classification parameters, if any, are listed between Territory and Industry. This means, N rate classification parameters are:
(Class, Territory, …, Industry).

Even though all derivation is performed with only three rate classification parameters (Class, Territory and Industry), it is now obvious how the formulae would look like in the general situation of N rate classification parameters (Class, Territory, …, Industry).

The **loss cost for Class i adjusted for heterogeneity** becomes:

$$L_i^C = \frac{\sum_{j=1}^{n} \sum_{k=1}^{p} l_{ijk}}{E_i^C} \text{ , i=1, 2, …, m.} \quad (3.1)$$

The **loss cost for Territory j adjusted for heterogeneity** becomes:

$$L_j^T = \frac{\sum_{i=1}^{m}\sum_{k=1}^{p} l_{ijk}}{E_j^T}, \text{ j=1, 2, …, n.} \tag{3.2}$$

The **loss cost for Industry k adjusted for heterogeneity** becomes:

$$L_k^I = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} l_{ijk}}{E_k^I}, \text{ k=1, 2, …, p.} \tag{3.3}$$

From these equations it follows

$$L := \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} l_{ijk} = \sum_{i=1}^{m} L_i^C E_i^C = \sum_{j=1}^{n} L_j^T E_j^T = \sum_{k=1}^{p} L_k^I E_k^I \tag{4}$$

The **base rate for Class i** (i.e. Class i, Territory 1, Industry 1 rate) can be expressed as:

$$R_i^C = \frac{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{ijk} e_{ijk}}{E_i^C} = \frac{\sum_{j=1}^{n}\sum_{k=1}^{p} r_{i11} t_j f_k e_{ijk}}{\sum_{j=1}^{n}\sum_{k=1}^{p} e_{ijk} t_j f_k} = r_{i11}, \text{ i=1, 2, …, m.} \tag{5.1}$$

Similarly,

$$R_j^T = r_{1j1}, \text{ j=1, 2, …, n.} \tag{5.2}$$

$$R_k^I = r_{11k}, \text{ k=1, 2, …, p.} \tag{5.3}$$

Then **earned premiums at current rates** can be written as

$$P = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} r_{ijk} e_{ijk} = \sum_{i=1}^{m} R_i^C E_i^C = \sum_{j=1}^{n} R_j^T E_j^T = \sum_{k=1}^{p} R_k^I E_k^I \quad (6)$$

The **fully developed and trended loss ratio for Class i** is

$$LR_i^C = \frac{\sum_{j=1}^{n}\sum_{k=1}^{p} l_{ijk}}{\sum_{j=1}^{n}\sum_{k=1}^{p} e_{ijk} r_{ijk}} = \frac{L_i^C E_i^C}{R_i^C E_i^C} = \frac{L_i^C}{R_i^C} \text{, i=1, 2, …, m.} \quad (7.1)$$

Similarly

$$LR_j^T = \frac{L_j^T}{R_j^T} \text{, j=1, 2, …, n.} \quad (7.2)$$

$$LR_k^I = \frac{L_k^I}{R_k^I} \text{, k=1, 2, …, p.} \quad (7.3)$$

For the proposed Class differentials we have

$$y_i = x_i \frac{LR_i^C}{LR_1^C} = \frac{r_{i11}}{r_{111}} \frac{\frac{L_i^C}{R_i^C}}{\frac{L_1^C}{R_1^C}} = \frac{L_i^C}{L_1^C} \text{, i=1, 2, …, m.} \quad (8.1)$$

Similarly

$$u_j = \frac{L_j^T}{L_1^T} \text{, j=1, 2, …, n.} \quad (8.2)$$

$$g_k = \frac{L_k^I}{L_1^I} \text{, k=1, 2, …, p.} \quad (8.3)$$

Remark 1: Note that the relations (8) prove that the "**loss ratio method**" is equivalent to the "**loss cost method**" if we adjust losses and exposures for heterogeneity.

Remark 2: Usually for the **Step 2** calculation of proposed differentials, **experience loss ratio** is defined as **current losses divided by earned premiums.** Instead, in the above derivation losses are fully developed and trended, i.e., the above loss ratios are different than the experience loss ratios. However, the final results for y_i, u_j, g_k are correct because the development and trend factors are the same in the numerators and the denominators, and therefore they cancel out in the equations (8).

In order to find the overall rate change we need first to find the:

Expected effective loss ratio = fully developed and trended losses divided by earned premiums at current rates

In the above notation it is

$$LR = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} l_{ijk}}{\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} r_{ijk} e_{ijk}} \qquad (9)$$

From (4) and (6) it follows

$$LR = \frac{\sum_{i=1}^{m} L_i^C E_i^C}{\sum_{i=1}^{m} R_i^C E_i^C} = \frac{\sum_{j=1}^{n} L_j^T E_j^T}{\sum_{j=1}^{n} R_j^T E_j^T} = \frac{\sum_{k=1}^{p} L_k^I E_k^I}{\sum_{k=1}^{p} R_k^I E_k^I} \qquad (10)$$

Then the overall **rate change** is equal to

$$RC = \frac{LR}{PLR} - 1,$$

where **PLR** stands for "**permissible loss ratio**" sometimes called "expected or target loss ratio."

If we are changing only one set of differentials, say Class differentials, i.e., if we **assume that the Territory and Industry differentials do not change**, then, in the above notation the **balance back factor** is

$$BBF = \frac{\sum_{i=1}^{m} y_i E_i^C}{\sum_{i=1}^{m} x_i E_i^C} \tag{11}$$

From **Steps 1,2,3**, it follows that for the **proposed rate** $\bar{r}_{i11}$ for the cell (i,1,1) it holds

$$\bar{r}_{i11} = \frac{r_{111} y_i LR}{PLR * BBF} \tag{12}$$

Substituting the equations (10) and (11) into (12) and then (1.1), (8.1) and (5.1),

$$\bar{r}_{i11} = \frac{r_{111} \sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m} x_i E_i^C}{PLR \sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m} y_i E_i^C} y_i = \frac{r_{111} \sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m} \frac{r_{i11}}{r_{111}} E_i^C}{PLR \sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m} \frac{L_i^C}{L_1^C} E_i^C} \frac{L_i^C}{L_1^C} =$$

$$= \frac{\sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m} r_{i11} E_l^C}{PLR \sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m} L_i^C E_i^C} L_i^C = \frac{L_i^C}{PLR} \tag{13}$$

Finally,

$$\bar{r}_{ijk} = \bar{r}_{i11} t_j f_k = \frac{L_i^C}{PLR} t_j f_k \,, \tag{14}$$

which is the formula we intended to prove.

In the developed form:

$$\bar{r}_{ijk} = \frac{(\sum_{j=1}^{n}\sum_{k=1}^{p} l_{ijk})t_j f_k}{E_i^C PLR} \tag{15}$$

This proves the following theorem:

Theorem 1: If **Territory and Industry differentials do not change**, then the proposed rate for the Class i, Territory j and Industry k is equal to product of the total FDTL of the Class i adjusted for heterogeneity and the current differentials of the Territory j, t_j, and Industry k, f_k divided by permissible loss ratio.

Let us now remove the assumption above that the Territory and Industry differentials do not change.

All the above formulae (1) through (10) still hold. The formula (11) for BBF changes into

$$BBF = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} y_i u_j g_k e_{ijk}}{\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k}^{p} x_i t_j f_k e_{ijk}} \tag{16}$$

Then, substituting equations (10), (16) into (12), and after that (8.1), (8.2), (8.3), (1.1) and (6) we get for the proposed rate:

$$\bar{r}_{ijk} = \frac{r_{111}\sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} x_i t_j f_k e_{ijk}}{PLR\sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} y_i u_j g_k e_{ijk}} y_i u_j g_k =$$

$$= \frac{r_{111}\sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} \frac{r_{i11}}{r_{111}} t_j f_k e_{ijk}}{PLR\sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} \frac{L_i^C}{L_1^C}\frac{L_j^T}{L_1^T}\frac{L_k^I}{L_1^I} e_{ijk}} \frac{L_i^C}{L_1^C}\frac{L_j^T}{L_1^T}\frac{L_k^I}{L_1^I} =$$

$$= \frac{\sum_{i=1}^{m} L_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} r_{ijk} e_{ijk}}{PLR\sum_{i=1}^{m} R_i^C E_i^C \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} L_i^C L_j^T L_k^I e_{ijk}} L_i^C L_j^T L_k^I = \frac{LL_i^C L_j^T L_k^I}{PLR\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} L_i^C L_j^T L_k^I e_{ijk}} \qquad (17)$$

The advantage of this formula for the proposed rate $\bar{r}_{ijk}$ over the **three-step process** is that it significantly simplifies the re-rating process. The previous steps to calculate the overall rate change; Class, Territory and Industry loss ratios; proposed Class, Territory and Industry differentials and balance back factor are no longer needed. There is no need to know even the current base rate. The only essential data are the fully developed and trended losses for each Class, Territory and Industry cell, l_{ijk}, and the number of exposure units for each cell, e_{ijk}. Given that, we only have to calculate the Class, Territory and Industry claim costs adjusted for heterogeneity, L_i^C, L_j^T, L_k^I by means of formulae (3.1), (3.2) and (3.3), respectively, and then to substitute it into formula (17), i.e., in formula (18) in general situation when N>3.

In general case, when there are N>3 classification parameters Class, Territory, …, and Industry, (i.e. Industry is Nth rather than third classification parameter) then the re-rating formula looks like:

$$\bar{r}_{ij\ldots k} = \frac{LL_i^C L_j^T \ldots L_k^I}{PLR\sum_{i=1}^{m}\sum_{j=1}^{n}\ldots\sum_{k=1}^{p} L_i^C L_j^T \ldots L_k^I e_{ij\ldots k}} \qquad (18)$$

3. Numerical Example

In the Excel workbook provided it has been demonstrated how the formula (18) works in case of two classification parameters Class and Territory. In that case the formula (18) simplifies to:

$$\bar{r}_{ij} = \frac{L L_i^C L_j^T}{PLR \sum_{i=1}^{m} \sum_{j=1}^{n} L_i^C L_j^T e_{ij}} \qquad (19)$$

The input data needed for the calculation is in both **bold** and ***italic*** format, and painted in blue in the workbook. Note that the base rate r_{11} does not play any role, i.e., we can enter any number, including 0, and the proposed rates remain the same. It means that the proposed rates are totally independent of the past rates, and depend exclusively on loss costs and permissible loss ratio.

Example:

Base Rate r_{11} =	100
Trend*Develop. Factor	***1.409***
PLR =	***0.80***

Matrix of Exposures E

12000	***3000***
4500	***2000***

Matrix of Current Losses

840,000.00	***300,000.00***
500,000.00	***250,000.00***

Matrix of FDT Losses L

		Total FDTL
1,183,602.74	422,715.26	L
704,525.44	352,262.72	2,663,106.16

Vector Row of Territory Diff. T

t_1	t_2
1	***1.15***

Vector Column of Class Diff. X

x_1	***1***
x_2	***1.1***

Matrix of Current rates

100.00	115.00
110.00	126.50

Vector Row of Territory Loss Costs

L_1^T	L_2^T
111	149

Vector Column of Class Loss Costs

L_1^C	103.97
L_2^C	155.41

11,582	15,495
17,312	23,161

Matrix ($L_i^C L_j^T e_{ij}$)

		$\sum_{i=1}^{m}\sum_{j=1}^{n} L_i^C L_i^T e_{ij}$
138,978,000	46,484,731	
77,902,843	46,322,826	309,688,400

Matrix of Proposed Rates

124.49	166.56
186.09	248.97

4. Conclusion

Many variables canceled out in the proof of the formula (17). The favorable consequence is that there is no need to calculate the following items any more:

1. The overall rate change,

2. Classification parameters' (Class, Territory, …, Industry) loss ratios,

3. Proposed Class, Territory, …, Industry differentials

4. Balance back factor.

It is well known in the re-rating field that it is tedious and sometimes impossible to calculate all the above variables. By formula (17), i.e., the general formula (18), those calculations are unnecessary. In addition, we do not need the current rates in order to calculate the new rates.

References

Brown, R.L., and Gottlieb, L.R. 2001. *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*, 2d edition. Winstead, CT: ACTEX Publications, Inc.