An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car & discussion of paper

Excess, Deductible, and Individual Risk Rating Tasks:

5. Analyze an experience rating plan.

By using data from the Canadian private passenger auto merit rating plan, Bailey & Simon demonstrate that a single car's experience for a single year has significant and measurable credibility for experience rating. They also show that the credibility of individual risk experience within a class varies based on how narrowly the class is defined. They aren't trying to set rates; they are looking backwards to show that the year Y-1 experience should be given some credibility in predicting the year Y experience.

Merit Rating Terminology

The merit rating system in Canada is similar to an experience rating plan, where the merit rating for an insured is based on the # of full years since the most recent accident (or if the insured has had no accidents, the # of years since the insured became licensed).

The merit ratings are:

1. 'A' - 3 or more years
2. 'X' - 2 years
3. 'Y' - 1 year
4. 'B' - 0 years

For example, if an insured had a merit rating of X at the start of year 1 but had 1 (or more) claims during year 1, then that insured would start year 2 as a B rating. If that insured instead had no claims in year 1, they would start year 2 with an A rating.

In the rating algorithm, this would be used as:

Premium = Base Rate $\times$ Merit Factor $\times$ Territory Factor $\times$ (any other variables)

The merit factors at the time of the paper were 0.65 for 'A', 0.80 for 'X', 0.90 for 'Y', and 1.00 for 'B'. This means that the base rate would correspond with 'B' risks.

In the paper, Bailey & Simon also use 'A + X' to mean 2 or more years and 'A + X + Y' as 1 or more years.

Experience Rating Formula

What Bailey & Simon do is ask: from a theoretical standpoint, what does the merit rating plan data imply about the credibility of experience from a single car for 1 year? They essentially do this by pretending that merit rating didn't exist, but instead, the rating plan had an experience modification factor. In this case, the premium formula would be:

Premium = Base Rate × Experience Mod × Territory Factor × (any other variables)

In experience rating, the experience mod is a credibility-weighted factor with credibility given to the experience of an individual risk and the complement of credibility given to the experience of the class of risks containing the individual risk. To put this in factor form, the experience is expressed relative to the class total experience. For example, using loss ratios as the measure of experience:

Cred-Wtd Individual Risk Loss Ratio = Z × Loss Ratio of Individual Risk + (1 - Z) × Loss Ratio of Class

Dividing all terms by the class loss ratio gives us the Mod formula (for a no-split plan):

$$\text{Experience Mod} = Z \times \frac{\text{Loss Ratio of Risk}}{\text{Loss Ratio of Class}} + (1 - Z)$$

Bailey & Simon denote the relative loss experience as R, so the formula becomes:

Mod = ZR + (1 - Z)

Normally in experience rating, we use historical data and established credibility values to calculate the Mod that will apply to a risk in the future policy term. Bailey & Simon will look at this backwards:

- They re-arrange the formula to solve for **Z** = (**Mod** - 1) / (**R** - 1).
- They start with the current policy term data to figure out what the ideal **Mod** would have been.
- Next they infer or estimate what **R** would have been for the prior policy term (they only use 1 year of experience since what they are trying to prove is that 1 year of data has some credibility, but in practice, experience rating often uses multiple years of historical data).
- Lastly, they plug **Mod** and **R** into the formula to solve for **Z**. The value of **Z** will then be the appropriate credibility for that individual risk for one year.

Bailey & Simon note that using relative loss ratios for **Mod** and **R** is usually problematic because severity across risks is generally too volatile and thus unreliable. While they do give an example of their calculation with relative loss ratios in Table 4 of their paper, the rest of their paper will omit severity and just use relative frequency TO PREMIUM (this would be equivalent to using relative loss ratios where severity is assumed to be constant across all risks, so it would cancel out).

Also, the premiums used in the relative frequency calculation should have 2 adjustments:

- Premiums should be on-leveled, so we don't double count the impact of past rate changes when calculating the experience Mod.
- Premiums should have the current merit rating factors backed out, since we will be replacing the merit rating factors with the experience Mod. Since 'B' ratings have a factor of 1.00, this is equivalent to saying premiums will be at 'B' rates.

Exposure Correlation

So why didn't Bailey & Simon calculate frequency as claim counts per exposure, the way we normally use the term? The reason is because of exposure correlation (what the paper calls "maldistribution"). Specifically, the authors were concerned about the possible exposure correlation between the territory rating variable and the merit rating variable.

As an example, if exposure correlation did exist between these variables, you might see more 'B' ratings in a given territory and more 'A' ratings in a different territory. This would mean that the frequency to car-years SHOULD differ between merit ratings solely because you'd expect some territories to have higher frequencies (to car-years) than other territories. However, when frequencies are defined relative to on-level premiums at 'B' rates, these premiums will include the current territory factors. If those territory factors are priced accurately, using premiums will adjust for the impact of exposure correlation.

To summarize the above using the paper's terminology, Hazam says a premium base (for frequency) only eliminates maldistribution if:

1. **High frequency (to car-year) territories are also high average premium territories.**
2. **Territorial (rate) differentials are proper.** One sign of this would be equal loss ratios across territories.

Calculating the Mod

We want to use data in the CURRENT period (2 policy years combined in the Bailey & Simon paper) to calculate the ideal **Mod** for a given risk. Instead of looking at individual risk data, Bailey & Simon calculate the ideal average Mod for all risks with a given current merit rating as:

$$\text{Mod} = \frac{(\text{\# of claims with rating}) \;/\; (\text{on-level earned premium for rating at 'B' rates})}{(\text{\# of claims in total for class}) \;/\; (\text{class total on-level earned premium at 'B' rates})}$$

For example, the Mod applied for 'B' ratings in class 1 on page 160 of the paper is 1.476. This is calculated based on numbers in Table 1 on page 162 as follows:

$$1.476 = \frac{(\text{\# of claims in class 1 from 'B' ratings}) \;/\; (\text{OLEP in class 1 from 'B' ratings at 'B' rates})}{(\text{\# of claims in total for class '1'}) \;/\; (\text{class '1' total OLEP at 'B' rates})} = \frac{(37{,}730)/(17{,}226)}{(288{,}019)/(194{,}106)} = \frac{2.190}{1.484}$$

Bailey & Simon calculate Mods for 'A' and 'B' ratings, and instead of calculating Mods for 'X' and 'Y' ratings, Bailey & Simon calculate them for 'A + X' and 'A + X + Y'. This is because they want to show the impact of adding additional years claims-free.

This approach results in ideal Mod values for all ratings that, if used in the rating algorithm instead of merit rating factors, would produce equal ratios of frequency to earned premium across all merit ratings.

Note that if there were no other variables in the rating plan besides merit rating (i.e., no territory variable), there would be no issue with exposure correlation, and using premium or exposures as the denominators for the relative frequency calculation would give you identical results for the Mod values.

Calculating Mod Example

Suppose we have the following rates and data for a book of business:

Premium = 1,000 (base rate) × Territory Factor × Merit Rating Factor

Merit Rating	Factor
A	0.65
X	0.80
Y	0.90
B	1.00

Territory	Factor
1	1.00
2	0.75

Insured	Territory	Merit Rating at Start of Year 2	Year 2 Claim Count	Year 2 Prem	Year 2 Prem at B rates
1	1	A	0	650	1,000
2	1	B	1	1,000	1,000
3	1	X	0	800	1,000
4	1	B	2	1,000	1,000
5	2	B	1	750	750
6	2	B	0	750	750
7	2	X	1	600	750
8	2	Y	0	675	750

If pricing was "perfect", we would have equal ratios of claim counts to premium (not at B rates) across all merit ratings. We can determine the ideal mods (that will replace merit factors) needed to get this perfect premium by looking at frequency to premium that backs out the merit factors (i.e., at B rates). If desired, we can use these ideal mods and Premium = 1,000 (base rate) × Territory Factor × Mod to calculate what would have been "perfect" premiums for year 2, and we can confirm these are perfect by checking that the frequency to premium ratios are flat across all merit ratings:

(1) Year 2 Merit Rating	(2) Year 2 Prem at B rates	(3) Year 2 Claims	(4)=(3)/(2) Freq to Prem at B rates	(5)=(4)/(4tot) Relative Frequency	(6)=(2)×(5) Perfect Premium	(7)=(3)/(6) Frequency to Perfect Prem
A	1,000	0	0	0	0	n/a
X	1,750	1	0.0006	0.8	1,400	0.000714
Y	750	0	0	0	0	n/a
B	3,500	4	0.0011	1.6	5,600	0.000714
Total	7,000	5	0.0007	1	7,000	0.000714

For their purposes, Bailey & Simon instead present the data using A+X and A+X+Y, as they will use this to show the impact of adding additional years claim-free (Total is now A+X+Y + B):

(1) Year 2 Merit Rating	(2) Year 2 Prem at B rates	(3) Year 2 Claims	(4)=(3)/(2) Freq to Prem at B rates	(5)=(4)/(4tot) Relative Frequency aka Ideal **Mod**	(6)=(2)×(5) Perfect Premium	(7)=(3)/(6) Frequency to Perfect Prem
A	1,000	0	0	0	0	n/a
A+X	2,750	1	0.0004	0.509	1,400	0.000714
A+X+Y	3,500	1	0.0003	0.4	1,400	0.000714
B	3,500	4	0.0011	1.6	5,600	0.000714
Total	7,000	5	0.0007	1	7,000	0.000714

Calculating R

To obtain **R**, we need the ratios of relative frequency to premium from LAST YEAR'S experience for each individual risk. For any risks with CURRENT ratings of 'A', 'X', or 'Y', by definition they had no claims last year, so R = 0 for these risks (the same logic applies when defining risks as 'A + X' and 'A + X + Y'). For these risks, the Mod formula reduces to **Mod = 1 - Z**, and we can already solve for **Z = 1 - Mod**.

Any risk with a current 'B' rating would have had at least 1 claim last year (unless it was newly licensed). If we actually had data for last year's experience for current 'B' ratings, we would calculate last year's relative frequency to premium for current 'B' ratings in order to obtain **R**. However, if we don't have this data (and in the paper we don't), we can approximate **R** for 'B' ratings as follows:

- Assume that the class total claim frequency to earned car-years is the same each year.
- Assume that claim counts are Poisson distributed.

For a Poisson distribution with mean λ, the formula for the proportion of insureds with k accidents is:

$$Pr(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \qquad Pr(X=0)=\frac{\lambda^0 e^{-\lambda}}{0!}=e^{-\lambda} \qquad Pr(X\geq 1)=1-Pr(X=0)=1-e^{-\lambda}$$

Remember that all claims in the class last year only come from current 'B' ratings. To determine R for these insureds, Bailey & Simon use relative claim frequency to earned-car years, which we can estimate with the above assumptions. In this case, **R** is approximated as:

$$\text{R} = \frac{\text{(\# of claims last year from current 'B' ratings) / (earned car years last year of current 'B' ratings)}}{\text{(\# of claims last year for class) / (class total earned car years last year)}} = \frac{N\lambda/[N(1-e^{-\lambda})]}{N\lambda/N} = \frac{1}{1-e^{-\lambda}}$$

where λ = class total claim frequency = $\dfrac{\text{\# of claims from class in current year}}{\text{earned car years* of insureds in class in current year}}$

and N is the total earned car-years in the class last year.

*Note that for λ, car-years is used as the denominator instead of premium.

The **R** formula above can also be described as a ratio of the average number of claims produced by 'B' rating insureds to the average number of claims produced by any insured in the class. For example, the R applied for 'B' ratings in class '1' in page 160 of the paper is $\frac{1.044}{0.087}=12$. This is calculated based on the R formula above with λ derived from Table 1 on page 162 as follows:

$$\lambda = \frac{\text{\# of claims from class '1'}}{\text{earned car years of insureds in class '1'}} = \frac{288{,}019}{3{,}325{,}714} = 0.087 \qquad \text{Avg \# claims from B ratings} = \frac{0.087}{1-e^{-0.087}} = 1.044$$

Finally, for 'B' ratings, we can re-arrange the Mod formula to solve for **Z** = (**Mod** - 1) / (**R** - 1). Continuing with the example from the prior page, we can get λ = 5 claims / 8 risks = 0.625, and then solve for Z:

Merit Rating	Mod	**R**	**Z** = (**Mod** - 1) / (**R** - 1)
A	0	0	1
A+X	0.509	0	0.491
A+X+Y	0.4	0	0.6
B	1.6	$\frac{1}{1-e^{-0.625}}$ = 2.152	0.521

What the Analysis Shows

Bailey & Simon show the implied credibilities for 'A' (3+ years), 'A + X' (2+ years), and 'A + X + Y' (1+ years) ratings for all classes in Table 2 of their paper. The credibility values for the 1 year ('A + X + Y' ratings) range between 0.038 and 0.071, which does imply that some credibility is warranted for the experience of a single car for a single year of claims-free experience (interestingly, you get negative credibility values if you just look at 'X' or 'Y' ratings by themselves, but this is because the Mod is relative to the class average, which is heavily influenced by 'A' ratings).

The 2nd to last column of Table 2 also shows the claim frequency for each class. All else being equal, higher frequency (more claim counts) would imply more data and thus greater credibility. However, when we look at the last column in Table 2 that shows the ratio of 3+ year credibility to claim frequency, we see that it is significantly higher for class 1. This is because experience rating credibility depends not just on the volume of data, but also the variance of loss experience within a class. Since classes 2 through 5 are more narrowly defined than class 1, risks within those classes will be more similar to each other than risks within class 1. As a result, experience rating, which distinguishes the individual risk from the class average risk, will have less credibility in classes 2-5.

Table 3 of the paper shows relative credibilities obtained by dividing the credibilities in Table 2 by the 1 year column in that table (e.g., dividing the credibility for A ratings by the credibility for A+X+Y). Bailey & Simon state that the closer the credibilities for 2 and 3 years of experience are to 2 and 3 times the 1 year credibility, then the less variation in an insured's probability of an accident for that class. They use an example to demonstrate this:

- Assume all risks have a mean claim frequency λ of either 0.05, 0.10, or 0.20.
- Assume frequency has a Poisson distribution, so $Pr(X = 0) = e^{-\lambda}$.

Claim	Expected number of risks with t years claims-free			
Frequency	$t = 0$	$t = 1$	$t = 2$	$t = 3$
0.05	100,000	95,123 = 100,000 $\times e^{-0.05}$	90,484 = 95,123 $\times e^{-0.05}$	86,071 = 90,484 $\times e^{-0.05}$
0.10	100,000	90,484 = 100,000 $\times e^{-0.10}$	81,873 = 90,484 $\times e^{-0.10}$	74,082 = 81,873 $\times e^{-0.10}$
0.20	50,000	40,937 = 50,000 $\times e^{-0.20}$	33,516 = 40,937 $\times e^{-0.20}$	27,441 = 33,516 $\times e^{-0.20}$
Total	250,000	226,543	205,873	187,593

Claim	Expected number of claims in year following t from above risks			
Frequency	$t = 0$	$t = 1$	$t = 2$	$t = 3$
0.05	5,000 = 100,000 $\times$ 0.05	4,756 = 95,123 $\times$ 0.05	4,524	4,304
0.10	10,000 = 100,000 $\times$ 0.10	9,048 = 90,484 $\times$ 0.10	8,187	7,408
0.20	10,000 = 50,000 $\times$ 0.20	8,187 = 40,937 $\times$ 0.20	6,703	5,488
Total	25,000	21,992	19,415	17,200
Claim Frequency	0.10000 = 25,000 / 250,000	0.09708 = 21,992 / 226,543	0.09430	0.09169
Freq relative to $t = 0$	1	0.9708	0.9430	0.9169
Z = 1 - Rel Freq		0.0292	0.0570	0.0831
Relative Credibility		1	1.948	2.843

The relative credibilities of 1.948 and 2.843 are close to 2 and 3, respectively. However, Table 3 of the paper showed lower values, which the authors said could be due to risks entering/exiting the portfolio or risk characteristics changing over time. In the discussion of the paper, Hazam also points out that the credibility increases closely in proportion to the # of years only for low credibility values.

Recap of Conclusions

Bailey & Simon come to 3 conclusions:

1. The experience of a single car for 1 year has significant and measurable credibility for experience rating.
2. Individual risk experience is more credible when there is more variance in loss experience within a risk class, which occurs in less refined risk classification systems.
3. The credibilities for varying years of experience should increase in proportion to the # of years of experience, though this is only true if several conditions hold: credibilities are low, risks aren't entering/exiting the book, and risk characteristics aren't changing over time.

Note on Bühlmann Credibility

Hazam mentions the formula for Bühlmann credibility in his paper. We already briefly discussed this in the prior section on Experience Rating, but it's worth touching on here again briefly.

Suppose X is a random variable (e.g., # of claims) with some distribution with parameter Θ (e.g., λ for a Poisson distribution), and Θ itself is a random variable with some distribution and additional parameters. In that case, the credibility of a sample of n observations from X is given by:

$$Z = \frac{n}{n+k}$$

n = # of claims in sample

$$k = \frac{E[Var(X|\Theta)]}{Var(E[X|\Theta])}$$

$E[Var(X|\Theta)]$ is the Expected Value of Process Variance (EPV)
$Var(E[X|\Theta])$ is the Variance of Hypothetical Means (VHM)

For example, if X is distributed Poisson(λ) and λ is distributed Normal(μ,σ^2), then:

$$k = \frac{E[Var(X|\lambda)]}{Var(E[X|\lambda])} = \frac{E[\lambda]}{Var(\lambda)} = \frac{\mu}{\sigma^2}$$

Note that k only needs to be calculated once for given random variables X and Θ, and remains constant for different samples of n sizes taken from the X variable. In the paper, Hazam backs into k by assuming $n = 100$ gives $Z = 0.046$, and then re-calculates Z for different values of n:

$$Z_{100} = 0.046 = \frac{100}{100+k} \qquad \implies k = 2{,}074$$

$$Z_{200} = \frac{200}{200+2,074} = 0.088 \qquad Z_{200}/Z_{100} = 0.088/0.046 = 1.912$$

$$Z_{300} = \frac{300}{300+2,074} = 0.126 \qquad Z_{300}/Z_{100} = 0.126/0.046 = 2.747$$

This shows that even theoretically, the relative credibilities for 2 or 3 times the experience should be less than 2 and 3 times the original (single year) credibility.

Problem Knowledge Checklist

1. Experience Rating Formula
    - Be able to state whether you should use exposures or premium as the denominator of frequency when calculating the Mod, and why.
    - Be able to state the 2 conditions the paper lists for using premium to account for maldistribution (i.e., exposure correlation) between merit rating and territory.
    - Be able to calculate the Mod, R, and Z for each rating (including knowing the Poisson formulas for 'B' ratings).
    - Be able to derive and calculate R for B ratings for non-Poisson distributions.
2. What the Analysis Shows
    - Be able to briefly explain why individual risk experience is more credible when there is more variance in loss experience within a risk class.
    - Be able to use ratios of credibilities for 2 and 3 years claim-free to 1 year claim-free to compare the stability between books of business.
    - Be able to state why the ratios for 2 and 3 years would be less than 2 and 3, respectively.
3. Note on Bühlmann Credibility
    - Be able to calculate EPV and VHM to obtain k.
    - Be able to calculate the credibility using the Bühlmann credibility formula.