

## EMPIRICAL BAYESIAN CREDIBILITY FOR WORKERS' COMPENSATION CLASSIFICATION RATEMAKING

GLENN MEYERS

### *Abstract*

This paper demonstrates how a company can derive accurate classification relativities. The method uses an empirical Bayesian credibility formula as taken from the paper "Credibility for Loss Ratios" by Buhlmann and Straub and modified by the ISO Credibility Subcommittee.

The data required for this method can be purchased from the National Council. A classification review is performed on three years of live data. Relativities predicted by both this method and the present ratemaking formula are compared with the actual relativities from a fourth year of data.

### I. INTRODUCTION

Workers' Compensation has traditionally been a highly regulated line of insurance. Rates are usually recommended by the National Council on Compensation Insurance and, with regulatory approval, become the industrywide standard. While many states permit deviations, insurers have generally adhered to the standard rates. Insurers compete on price by offering various dividend plans.

With the creation of the model law for competitive rating in Workers' Compensation, this is rapidly changing. In order to promote a better business climate, many states have passed competitive rating laws.

Under a uniform pricing system, it is not necessary to have rates equal to the expected cost of writing the policy. But in a competitive environment, many economists, such as Paul Samuelson [1], assert that the price will be equal to the expected cost of writing the policy. While the present ratemaking formula, which is described by Kallop [2], makes no systematic deviation from expected cost pricing (on an underwriting basis), it is not obvious that these rates are the best estimates of the expected cost. The present ratemaking method has held up for a long time under a system of uniform ratemaking, but it remains to be seen how long it will hold up under the increased pressure of open competition.

In most states, all insurers report their experience to the National Council. This reporting takes two forms. First, insurers report their aggregate premium and loss experience. Since rates are uniform, it is not necessary to adjust premiums to a common rate level. Thus it is easier to estimate the overall needed rate change with this data. Second, insurers report loss and exposure experience for each insured on a policy year basis. While this data is not as timely as the financial aggregate data, it is more detailed. Because of its fine breakdown, it can be used for deriving class relativities.

The broad-based experience reported for Workers' Compensation should be compared to the experience reported for other lines. In private passenger automobile insurance, for example, many policies are written by independent insurers who do not report their experience. Many different classification systems and rating plans are used. Thus, combining experience is difficult, if not impossible. Because of this, it is difficult for many insurers to set accurate rates.

It can be argued that reporting experience on a standard basis can enhance competition by making it easier for insurers to enter the market. But the need to report experience on a standard basis can discourage insurers from trying innovative classification systems and rating plans. Clearly, some compromises must be made in order to obtain the greatest benefits from competitive rating.

To summarize, the economic incentive to calculate accurate rates for Workers' Compensation is stronger than ever before, and the volume and quality of data are better than in any other line of insurance. Also, methods of data processing are becoming cheaper and more flexible. Under these conditions, improvements in the accuracy of ratemaking can surely be made.

This paper addresses the problem of determining accurate classification relativities. The method used to derive classification relativities differs from the present method in its use of an empirical Bayesian credibility formula.

We begin with a description of the empirical Bayesian credibility formula. We then compare the accuracy of the classification relativities predicted using this formula with those predicted by the present ratemaking formula.

The theory described in this paper is applicable to both loss ratio and pure premium ratemaking. However, it makes no sense to credibility weight the pure premium of a class with a thirty cent rate with the pure premium of a class with a thirty dollar rate. This is frequently the case in Workers' Compensation. Thus, we describe the theory in terms of loss ratios.

The loss ratios are based on Unit Statistical Plan data. Since the overall rate change is determined externally (the National Council uses financial aggregates), these loss ratios are used to determine class relativities.

## 2. INFORMATION AND ESTIMATION

A general principle in statistical estimation theory is that more information about a certain quantity leads to a better estimate of that quantity. A goal of statistical estimation theory is to develop ways of using all sources of relevant information in arriving at an estimate. In this section we shall show how this principle applies to Bayesian estimation and credibility theory.

Our problem is to estimate the loss ratio for a class of insureds. We consider two sources of information that can be used to estimate the loss ratio.

First, we can use the historical loss ratios for the class. While this information has a direct relationship to the quantity being estimated, it can be subject to random fluctuation because of small volume.

Second, we can use the loss ratio for a group of similar classes. Because of the greater volume of experience, this information has less random fluctuation. However, it has a less direct relationship to the quantity being estimated. The classes in the group may simply have different loss ratios.

Each of these sources of information is relevant to the quantity being estimated. The problem we want to address becomes the following: how can one use both sources of information to derive an estimate of the loss ratio for a class?

We seek a mathematical solution to this problem. To solve this problem we must first specify a model that we feel resembles the situation. We must then specify the information that we have available. We then mathematically derive the best estimate of the loss ratio.

We begin by making the following assumptions.

1. The expected loss ratio,  $\mu$ , is randomly selected from a distribution with mean  $M$  and variance  $\tau^2$ .
2. Each loss ratio,  $X$ , is randomly selected from a distribution with mean  $\mu$ , and variance  $\sigma^2$ .

This model bears a fair resemblance to our situation. We observe a class loss ratio,  $X$ , which fluctuates around the class's expected loss ratio,  $\mu$ . Our second source of information is the loss ratio,  $M$ , for a group of classes. The

possibility that classes in this group may have different loss ratios is represented by selecting  $\mu$  at random from a specified distribution.

The problem is to estimate the true loss ratio for a given class. We now describe some solutions to this problem.

### *The Bayesian Solution*

The Bayesian solution to this problem is to calculate the average  $\mu$  for all classes with observed loss ratio  $X$ . We write this as  $E[\mu|X]$ . One must have a complete description of the distributions for  $X$  and  $\mu$  to perform this calculation. For example, if we know that  $X$  and  $\mu$  are normally distributed, it is demonstrated by Hoel [3] that

$$E[\mu|X] = \frac{\tau^2}{\tau^2 + \sigma^2} \cdot X + \frac{\sigma^2}{\tau^2 + \sigma^2} \cdot M.$$

Hewitt [4] and Mayerson [5] give the Bayesian solution for other distributional assumptions.

It should be noted that the Bayesian solution given above is a linear function of the observed loss ratio,  $X$ . While this is also true for many other Bayesian solutions, it is not true for all Bayesian solutions. Hewitt [6] gives an example where the Bayesian solution is not linear.

### *The Credibility Solution*

The credibility solution, given by Buhlmann [7], is to use the linear approximation to the Bayesian solution which minimizes the expected squared error. As noted above, in many cases the credibility solution is identical to the Bayesian solution. While the credibility solution may not be as accurate as the Bayesian solution, it does not require as much information. One need not have a complete description of the distribution of  $X$  and  $\mu$ . One need only have the values of  $M$ ,  $\tau^2$  and  $\sigma^2$ . We will denote the credibility solution by  $C[\mu|X]$ .

The credibility solution can be stated as follows. Let

$$C[\mu|X] = A \cdot X + B.$$

We want to choose  $A$  and  $B$  so that

$$E[(C[\mu|X] - E[\mu|X])^2]$$

is minimized. The solution can be written in the following form.

$$C[\mu|X] = \frac{\tau^2}{\tau^2 + \sigma^2} \cdot X + \frac{\sigma^2}{\tau^2 + \sigma^2} \cdot M.$$

Define the credibility factor,  $Z$ , as follows:

$$Z = \frac{\tau^2}{\tau^2 + \sigma^2}$$

The credibility solution now takes the more familiar form:

$$C[\mu|X] = Z \cdot X + (1 - Z) \cdot M.$$

The credibility factor can be viewed as a measure which compares the variance of  $X$  with the variance of  $\mu$ . A credibility factor close to zero indicates that the random fluctuations of individual class loss ratios are large compared to the true differences in loss ratios between classes in the group. A credibility factor close to one indicates just the opposite. Philbrick [8] discusses this aspect of credibility theory in detail.

A major problem with the credibility solution is that, in real life situations, one does not know  $M$ ,  $\tau^2$  or  $\sigma^2$ . While it is possible to choose the unknown parameters by judgment, American actuaries have used a more direct approach; they choose the entire estimation formula by judgment. These formulas are generally referred to as the "classical" credibility formulas. The rationale for these formulas is given by Longley-Cook [9].

While the Bayesian and the credibility solutions provide considerable insight into the estimation process, one more step is needed. We must be able to form our estimates entirely from observations. This is the essence of the empirical Bayesian solution.

### 3. EMPIRICAL BAYESIAN CREDIBILITY

We begin our discussion of empirical Bayesian credibility with a description of the solution given by Buhlmann and Straub [10] in their landmark paper "Credibility for Loss Ratios." This solution has been amplified and modified by the Credibility Subcommittee of Insurance Services Office. Much of the following development is taken from a report written by the Credibility Subcommittee [11].

We begin by specifying the model underlying the empirical Bayesian credibility formula. Next, we give the credibility formula in terms of the parameters of the model. Finally, we show how to estimate the parameters of the model.

### The Model

The formula requires the following data.

1.  $T$  years of experience for  $N$  classes.
2. The premium for class  $i$  in year  $t$  (denoted by  $P_{it}$ ).
3. The loss ratio for class  $i$  in year  $t$  (denoted by  $X_{it}$ ).

We make the following assumptions.

1. The expected loss ratio for class  $i$ ,  $\mu_i$ , is randomly selected from a distribution with mean  $M$  and variance  $\tau^2$ .
2. Each loss ratio,  $X_{it}$ , is randomly selected from a distribution with mean  $\mu_i$  and variance  $V_i^2/P_{it}$ .

Most actuaries would agree that the variability of a class loss ratio decreases as the size of the class increases. The assumption that the variance of the loss ratio is inversely proportional to the premium (i.e.,  $\text{Var}[X_{it}] = V_i^2/P_{it}$ ) is a simple way to approximate this relationship. Note that the constant of proportionality,  $V_i^2$ , can be different for each class.

It is unlikely that this relationship is precise. Meyers and Schenker [12] propose a model of the loss process in which the variance of the loss ratio is not inversely proportional to the premium. In this model the variance of the loss ratio can be written in the form  $\text{Var}[X_{it}] = \alpha/P_{it} + \beta$ . The constant term,  $\beta$ , is positive when there are additional, but unidentified, sources of variation. Examples of this could include changing economic conditions, or increased emphasis on loss control. Meyers [13] discusses how a positive constant term affects the credibility formula.

### The Credibility Formula

For a given class,  $j$ , we want to find an estimate,  $\hat{\mu}_j$ , of the expected loss ratio,  $\mu_j$ . Here, we present the formula given by Buhlmann and Straub [14].

The estimate is of the following form.

$$\hat{\mu}_j = \sum_i \sum_t A_{it} \cdot X_{it}$$

$A_{it}$  is chosen to minimize  $E[(\hat{\mu}_j - \mu_j)^2]$ , subject to the constraint that  $E[\hat{\mu}_j] = M$ .

Note that all the observed loss ratios,  $X_{it}$ , contain some information about the expected loss ratio  $\mu_j$ . The exact nature of this information is specified by

the assumptions listed above and the accompanying mathematics. It should be noted that since the  $X_{ji}$ 's contain more information about  $\mu_j$  than the other  $X_{ii}$ 's, the  $A_{ii}$ 's depend upon  $j$ .

Using the method of Lagrange multipliers, one can solve for the  $A_{ii}$ 's. Buhlmann and Straub went one step further by algebraically manipulating the solution so as to express it in a form which resembles a standard credibility formula.

$$\text{Let } P_i = \sum_i P_{ii} \quad (\text{total class premium}),$$

$$\bar{X}_i = \sum_i P_{ii} \cdot X_{ii}/P_i \quad (\text{premium weighted average of } X_{ii}),$$

$$\Sigma^2 = E[V_i^2]$$

$$K = \Sigma^2/\tau^2 \quad (\text{credibility constant}),$$

$$Z_i = P_i/(P_i + K) \quad (\text{credibility factor}), \text{ and}$$

$$\hat{M} = \sum_i Z_i \cdot \bar{X}_i / \sum_i Z_i \quad (\text{credibility weighted average of } \bar{X}_i).$$

$$\text{Then } \mu_j = Z_j \cdot \bar{X}_j + (1 - Z_j) \cdot \hat{M}.$$

There is one point that should not be overlooked. The complement of credibility is assigned to the *credibility-weighted* average loss ratio and not the premium-weighted average loss ratio as many would assume. The reason for this is simply that it is the solution to the minimization problem. It should be noted that  $\hat{M}$  has some very nice properties.

First, it can be demonstrated [15] that

$$\sum_i \sum_i P_{ii} \cdot \hat{\mu}_i = \sum_i \sum_i P_{ii} \cdot X_{ii}.$$

This means that the estimates of the class loss ratios are "in balance" with the overall loss ratio.

Second, it can be demonstrated [16] that  $\hat{M}$  is the minimum variance unbiased estimate of  $M$ .

#### *Estimating the Parameters*

The following estimators of  $\Sigma^2$  and  $\tau^2$  were derived by Buhlmann and Straub [17].

Let  $P_{..} = \sum_i \sum_t P_{it}$  (total premium),

$$P2 = \sum_i P_i^2,$$

$\bar{X}_{..} = \sum_i \sum_t P_{it} \cdot X_{it} / P_{..}$  (premium-weighted average of  $X_{it}$ ), and

$$W = \sum_i P_i \cdot (\bar{X}_i - \bar{X}_{..})^2 / (N - 1)$$

Then estimates for  $\Sigma^2$  and  $\tau^2$  are given by

$$\hat{\Sigma}^2 = \frac{\sum_i \sum_t P_{it} \cdot (X_{it} - \bar{X}_i)^2}{N \cdot T - N} \quad \text{and}$$

$$\hat{\tau}^2 = \frac{(W - \hat{\Sigma}^2) \cdot (N - 1) \cdot P_{..}}{P_{..}^2 - P2}.$$

Buhlmann and Straub then used  $\hat{K} = \hat{\Sigma}^2 / \hat{\tau}^2$  as their estimate of the credibility constant. The credibility of a class loss ratio becomes the following:

$$\hat{Z}_i^1 = \frac{P_i}{P_i + \hat{K}}.$$

The ISO Credibility Subcommittee modified this formula for the following reason. Even though  $\hat{\Sigma}^2$  is an unbiased estimate of  $\Sigma^2$ , and  $\hat{\tau}^2$  is an unbiased estimate of  $\tau^2$ , it turns out that  $\hat{Z}_i^1$  is a biased estimate of  $Z_i$ . The modified formula, which attempts to correct for this bias, can be written as follows.

$$\hat{Z}_i = \frac{P_i}{P_i + \hat{K}} \cdot \frac{N - 3}{N} + \frac{3}{N}$$

This modification is identical to that given by Morris and Van Slyke [18]. A derivation of this modification is given by ISO [19]. This derivation makes a number of simplifying assumptions in addition to those already stated. They are as follows.

1.  $X_{it}$  is normally distributed.
2.  $\mu_i$  is normally distributed.
3.  $\Sigma^2$  is known.



Since these assumptions are somewhat restrictive, this correction for bias should be regarded as only approximate.

Under the above assumptions, it is not possible to correct for this bias when  $N < 3$ . Thus, one should not use this empirical Bayesian formula when there are three or fewer classes.

Note that the minimum credibility that is possible in this formula is  $3/N$ .

It is possible for the estimate,  $\hat{\tau}^2$ , to be negative. This can be disconcerting to those who think that estimates of a variance should be positive. However, this phenomenon does have a natural interpretation. If we assume that the  $X_{it}$ 's are normally distributed in addition to our stated assumptions, it is possible to test the hypothesis that all the  $\mu_i$ 's are equal. This test is referred to as analysis of variance (ANOVA), and is described by Freund and Littell [20]. This test calculates a statistic called the  $F$  statistic. Abnormally high values of the  $F$  statistic indicate that we should reject the hypothesis that all  $\mu_i$ 's are equal, while lower  $F$  values indicate failure to reject this hypothesis.

It turns out in our case that  $F = W/\hat{\Sigma}^2$ . Thus we have that  $\hat{\tau}^2$  is negative if and only if  $F$  is less than one. Since under the null hypothesis,  $E[F] = (N \cdot T - N)/(N \cdot T - N - 2) > 1$ , a negative  $\hat{\tau}^2$  indicates failure to reject the hypothesis that all  $\mu_i$ 's are equal.

Thus, we should assign a credibility of zero when  $\hat{\tau}^2$  is negative.

One additional point should be made. The derivation of these estimators requires that the loss ratios for a given class are independent from one year to the next. Most ratemaking procedures in use at this time use loss ratios at "present rates." If rates are revised yearly, all but the most recent year of experience is used in calculating the present rate. The premium, and hence the loss ratio, for the most recent year will be influenced by the experience of the prior years. Thus, the independence assumption is violated!

The effect of using premium at present rates is to understate our estimate of  $\tau^2$ .  $W$  is sharply reduced, while  $\hat{\Sigma}^2$  will not be significantly affected. An extreme case results when all years of the current review were used in making the present rates, and a credibility of one was used. In this case, all the  $X_{it}$ 's are equal to the expected loss ratio,  $W$  is equal to zero and  $\hat{\tau}^2$  is negative.

What to do about this problem is currently being debated by the Credibility Subcommittee. Some members feel that present rates should be used for estimating loss ratios, and the focus of the debate is on how to do this. In this

paper we do not use present rates. Instead we use the most recent rates which were not based on the current experience.

It should be noted that if  $X_{it}$  is a pure premium rather than a loss ratio, the  $X_{it}$ 's will be independent, and it is not necessary to refer to older rates.

In summary, we have presented a credibility formula whose parameters are derived entirely from available data, and we have stated the assumptions that are used in deriving this formula. As is often the case in actuarial science, the model associated with these assumptions is necessarily simpler than the real world. However, this formula is easy to use and can produce accurate results, as we shall now demonstrate.

#### 4. RATEMAKING WITH EMPIRICAL BAYESIAN CREDIBILITY

We now demonstrate how to use empirical Bayesian credibility in classification ratemaking.

##### *The Data*

Whenever the National Council files rates, it releases the raw data that underlie the rates. Recently, they began selling tapes containing loss and exposure data (Schedule Z), by class, derived from the Unit Statistical Plan. For this study, we obtained the tapes which correspond to the 1982 and 1983 rates for the state of Michigan.

The most recent rates which did not utilize any of the above data were those for the year 1979. Thus we calculate the premium by multiplying the payroll times the 1979 rate.

Below, we use the data on the first tape to calculate class relativities. Thus it is possible to make a direct comparison between the 1982 rates and the rates produced below. The tape which corresponds to the 1983 Michigan rates contained an additional year of data. We will use this additional year of data to compare the accuracy of the rates derived using the present ratemaking formula with those derived using empirical Bayesian credibility.

The losses were adjusted for law changes and loss development with factors taken from the 1982 Michigan rate filing. One technical point should be made here. The 1982 National Council rates do not reflect the modification due to (Michigan) Senate Bill 1044. This is appropriate since none of the experience reflects this bill and the adjustment was made outside the usual ratemaking formula.

Our purpose is to provide a direct comparison of ratemaking formulas, and so classes which presented special problems were deleted from this analysis. The special problems were of two kinds. First, many classes were absorbed into other classes between 1979 and 1982. It was felt that the 1979 rate for the new class could not be accurately estimated. Second, some classes contained disease elements which require special treatment. In practice, these problems must be dealt with. But that is beyond the scope of this paper.

Exhibit I shows the data used.

#### *Determining the Class Loss Ratios*

The empirical Bayesian credibility formula was applied to the data of Exhibit I with the following results.

$$\begin{aligned} N &= 319 \\ \hat{\Sigma}^2 &= 92374 \\ \hat{\tau}^2 &= 0.019237 \\ \hat{K} &= 4801900 \\ \hat{M} &= 0.5822 \end{aligned}$$

For each class  $i$ , the credibilities,  $\hat{Z}_i$ , and the estimates,  $\hat{\mu}_i$ , are given in Exhibit I.

#### *Distributing the Overall Rate Change*

Even a moderately large insurer is unlikely to have exposure in all classes for which it must have a rate. Thus most insurers must obtain data similar to that described above in order to make independent rates for all classes. However, a company does not need data in such fine detail to determine the overall rate change.

As noted above, the National Council uses financial aggregate premium and loss experience to determine the overall rate change. Individual companies operating in a competitive environment invariably will have their own way of deriving the overall rate level. It is not our purpose to describe methods of determining the overall rate change. Instead we will describe how a company might distribute the overall rate change to the individual classes.

The procedure described below will produce estimates,  $\hat{\mu}_i$ , of the loss ratio at 1979 rates for each class  $i$ . Since it is quite likely that an insurer's payroll in the various classes will have changed since 1979, a logical procedure for determining the final rates might proceed as follows.

Let  $L$  = Total loss provision for the insurer's current book of business at the proposed rate level,

$E_i$  = insurer's current payroll for class  $i$  and

$R_i$  = 1979 rate for class  $i$ .

We define the rate adjustment factor,  $A$ , as follows.

$$A = L / \left( \sum_i E_i \cdot R_i \cdot \hat{\mu}_i \right)$$

The loss provision in the rate for class  $i$  is then given by the expression  $R_i \cdot \hat{\mu}_i \cdot A$ . If the loss provision in the rate for class  $i$  is defined in this manner, the total loss provision for the new class rates on the current book of business will be equal to  $L$ .

It should be noted that the estimates,  $\hat{\mu}_i$ , are really being used to determine class relativities.

#### 5. TESTING CREDIBILITY FORMULAS

We shall now compare the accuracy of the rates produced by the empirical Bayesian credibility formula with those rates produced by the present ratemaking method.

##### *The Underwriting Test*

The accuracy of a ratemaking method can have a very important practical consequence. Suppose you are in an environment where some less accurate ratemaking method is being used. If you choose, or are required, to use the less accurate rates, you can use the more accurate rates to identify the better insureds. By writing these better insureds, you will have better than average underwriting results. Conversely, suppose you are able to use the rates indicated by the more accurate ratemaking method. You would then be charging a lower rate for the better insureds, and a higher rate for the worse insureds. You could then increase your writings for the better insureds and still make an adequate profit, while your competitors who use the other ratemaking method should write more of the worse insureds and make a less than adequate profit. A common phrase for this procedure is "skimming the cream."

Our first test will be based on this phenomenon, and will appropriately be called the "Underwriting Test." This test proceeds as follows. We first estimate the expected losses predicted by each formula for the test year. For each class,  $i$ , the expected losses are computed as follows.

*Present Method:*

$$\text{Expected Loss}_i = \text{Payroll}_i \cdot 1982 \text{ Rate}_i \cdot 0.769384$$

*Empirical Bayesian Credibility:*

$$\text{Expected Loss}_i = \text{Payroll}_i \cdot 1979 \text{ Rate}_i \cdot \hat{\mu}_i \cdot 1.053661$$

Since we are interested only in class relativities, we use the factors 0.769384 and 1.053661 to force the expected loss to sum to the total expected losses for the test year.

Next, we divide the classes into two groups. Group 1 consists of all classes for which the present ratemaking formula gives lower expected losses. Group 2 consists of all other classes.

For each group we then compare the ratio of actual losses for the test year to the expected losses predicted by both ratemaking formulas. The results are in the following table.

TABLE 1  
UNDERWRITING TEST

	<u>Group 1</u>	<u>Group 2</u>	<u>Total</u>
1. # Classes	162	157	319
2. Actual Loss	216906003	199032667	415938670
3. Exp. Loss (Pres. Mthd.)	208238132	207700538	415938670
4. Exp. Loss (E. B. Cred.)	220310030	195628640	415938670
5. (2)/(3)	1.042	0.958	1.000
6. (2)/(4)	0.985	1.017	1.000

Line 5 of Table 1 shows that by using the present ratemaking formula and underwriting in favor of the Group 2 classes, one expects a better than average profit. Line 6 of Table 1 shows that by using the rates produced by the empirical Bayesian credibility formula, one could charge less than the rates produced by the present formula for the Group 2 classes and still make an average profit. Competitors with the same overall rate level who use the present ratemaking formula may end up writing a greater concentration of Group 1 classes and make less than their anticipated profit.

Thus we conclude that the empirical Bayesian credibility formula produced more accurate rates for this data.

We now address the statistical significance of this result. Our test is similar to the "bootstrap" technique described by Diaconis and Efron [21]. For our test, we constructed 2000 groups of insureds in which the members of the group were selected at random with a probability of 0.5. The loss ratios for each group were calculated and then listed by percentiles. These percentiles are given in Table 2.

TABLE 2  
RANDOM LOSS RATIOS—  
PRESENT RATEMAKING  
METHOD

<u>Percentile</u>	<u>Loss Ratio</u>
.010	.939
.025	.949
.050	.957
.100	.965
.150	.971
.200	.976
.250	.980
.750	1.021
.800	1.027
.850	1.033
.900	1.041
.950	1.053
.975	1.064
.990	1.075

Looking at Table 2 we see that the Group 1 loss ratio for the present ratemaking method of 1.042 is near the 90<sup>th</sup> percentile of the random loss ratio distribution. Similarly, we see that the Group 2 loss ratio of .958 for the present ratemaking method is close to the fifth percentile of the random loss ratio distribution.

Now there are two types of errors that can be made. A Type I error occurs when one keeps the present method when the empirical Bayesian method is better. A Type II error occurs when one changes from the present method to the empirical Bayesian method when the two methods are equally accurate. Table 2 shows that the probability of making a Type II error is less than one in ten. The probability of making a Type II error (i.e. the significance level) that should be required in order to change methods depends upon the relative costs of the two types of errors.

A single insurance company operating in a competitive environment may miss a good opportunity to expand in some profitable classes if it makes a Type I error, but should lose very little by committing a Type II error. A one in ten chance of making a Type II error should be sufficient to justify adopting the empirical Bayesian method.

A Type II error can be very costly for a rating bureau which is making an industrywide filing in a noncompetitive environment. Should the error be discovered after such a filing, the cost of returning to the present method can be enormous in time, money, and embarrassment. In such cases a one in ten chance of making a Type II error may not be sufficient to justify changing methods, and additional tests should be made. However, it should be noted that the cost of a Type I error is not insignificant. Companies can use the empirical Bayesian method for underwriting. There could be availability problems for some classes.

The table of loss ratio distributions for the empirical Bayesian credibility formula is similar to Table 2. The loss ratios of .985 for Group 1 and 1.017 for Group 2 are well within the normal range of fluctuation.

#### *Mean Squared Error*

A natural test for a ratemaking method is to measure how close the expected loss comes to the actual loss for the next year. With this in mind we calculate the following statistic.

$$MSE = \sum_i P_i \cdot (A_i/E_i - 1)^2/N$$

Where  $A_i$  = actual loss for class  $i$

$E_i$  = expected loss for class  $i$

$P_i$  = 1979 rate for class  $i$  times the payroll for class  $i$

$N$  = number of classes (319).

We shall refer to the number  $P_i \cdot (A_i/E_i - 1)^2$  as the squared error for class  $i$  and we shall refer to  $MSE$  as the mean squared error.

The test statistics for the ratemaking methods considered above are given in the following table.

TABLE 3

	<u>MSE</u>
Empirical Bayesian Credibility	289651
Present Ratemaking Formula	298063

Here we see that the empirical Bayesian credibility formula produces the lower mean squared error.

To test if the differences between these mean squared errors are statistically significant we must consider the following.

1. The squared error for a class using one method is not independent of the squared error for the same class using another method.
2. The distribution of the squared errors is not normal.

A test that can work under these conditions is the Wilcoxon signed ranks test [22], which we now describe.

For a class  $i$ , let  $SE1_i$  be the squared error for the present ratemaking method and let  $SE2_i$  be the squared error for empirical Bayesian credibility. Let

$$DSE_i = SE1_i - SE2_i$$

$$R_i = \text{Rank}(|DSE_i|) \cdot \text{Sign}(DSE_i)$$

$$T = \sum_i R_i / \left( \text{Square root} \left( \sum_i R_i^2 \right) \right)$$

We want to test the hypothesis

$$H_0: E\{SE1_i\} = E\{SE2_i\}$$

against the alternative hypothesis

$$H_1: E\{SE1_i\} \neq E\{SE2_i\}.$$

For large  $N$ , we reject  $H_0$  at the level of significance  $\alpha$  if  $T$  lies below the  $(\alpha/2)^{\text{th}}$  or above the  $(1 - \alpha/2)^{\text{th}}$  percentile of the standard normal curve.

When comparing the *MSE* of the rates produced by the empirical Bayesian credibility formula with those produced by the present formula, we get



$T = .198$  which is at the 56<sup>th</sup> percentile of the standard normal distribution. Thus we cannot reject  $H_0$ . Thus we conclude the expected mean squared errors are not significantly different.

Of the two tests conducted, the author considers the underwriting test to be the most relevant, since it corresponds directly to actions an insurance company can take. However the mean squared error test corresponds more closely to the criteria under which the empirical Bayesian credibility formula was derived, with the main difference being the substitution of actual loss ratios for "true" (but unmeasurable) loss ratios. This substitution adds a great deal of volatility to the test.

## 6. CONCLUSION

This paper describes how an empirical Bayesian credibility formula can be used to determine class relativities for Workers' Compensation insurance. Tests which compared the accuracy of this method with the present ratemaking method showed that the empirical Bayesian credibility formula produced more accurate rates.

The level of significance of these tests was sufficient for use by individual companies in a competitive environment, but the author would stop short of recommending industrywide use of this method in a highly-regulated noncompetitive environment until further tests are made.

However, it should be pointed out that if the empirical Bayesian approach is even marginally more accurate than the present approach, its accuracy should increase over time. One of the features of the approach described above is that it had to use the 1979 rates which were derived by the present ratemaking formula. If this method were adopted for the 1985 rates, the rates calculated above could be used in place of the 1979 rates. Gradually, the rates will become even more accurate.

Another advantage to the empirical Bayesian approach is that it calculates an optimal result based on an explicit set of assumptions. By knowing how well the assumptions are met, one can better decide when to adjust the calculated results on a judgemental basis, or when to derive a new formula based on alternative assumptions.

This author doubts that the above approach will be the last word in credibility theory, but it is hoped that this paper has set a standard that proposals for alternative formulas will follow. This standard is that the predictions should be

tested on independent data. This standard is part of the scientific method and should be applied to actuarial science.

## 7. ACKNOWLEDGMENTS

The ratemaking method described in this paper is being used by my company. In developing this method I worked very closely with Burt Covitz. Burt's very detailed knowledge of Workers' Compensation ratemaking made this method much better than it might otherwise have been. Brad Alpert and Mike Kooken also contributed many valuable comments.

I have also profited tremendously by the very thorough work done by the staff of the ISO Credibility Subcommittee. ISO deserves to be commended for the resources committed to this subcommittee.

The research for this paper was supported by a grant from the Actuarial Education and Research Fund.

## 8. REFERENCES

- [1] Paul Samuelson, *Economics*, 11<sup>th</sup> Edition, McGraw-Hill, Manchester, Missouri, 1980, p. 431.
- [2] Roy Kallop, "A Current Look at Workers' Compensation Ratemaking," *PCAS LXII*, 1975, p. 62.
- [3] P. G. Hoel, *Introduction to Mathematical Statistics*, Fourth Edition, John Wiley and Sons, Inc., New York, New York, 1971.
- [4] Charles C. Hewitt, "Credibility for Severity," *PCAS LVII*, 1970, p. 148.
- [5] Allen L. Mayerson, "A Bayesian View of Credibility," *PCAS LI*, 1964, p. 65.
- [6] Hewitt, op. cit.
- [7] H. Buhlmann, "Experience Rating and Credibility," *ASTIN Bulletin IV*, Part III, 1967, p. 199.
- [8] Stephen W. Philbrick, "An Examination of Credibility Concepts," *PCAS LXVIII*, 1981, p. 195.

- [9] L. H. Longley-Cook, *An Introduction to Credibility Theory*, Casualty Actuarial Society, 1962.
- [10] H. Buhlmann and E. Straub, "Credibility for Loss Ratios," (Translated by C. E. Brooks), *ARCH*, 1972.
- [11] Insurance Services Office, *Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Procedures for Classification Ratemaking*, September, 1980.
- [12] Glenn Meyers and Nathaniel Schenker, "Parameter Uncertainty and the Collective Risk Model," *PCAS LXX*, 1983.
- [13] Glenn Meyers, "An Analysis of Experience Rating," (Submitted for Publication).
- [14] H. Buhlmann and E. Straub, op. cit.
- [15] Insurance Services Office, op. cit. p. 80.
- [16] Insurance Services Office, op. cit. p. 81.
- [17] H. Buhlmann and E. Straub, op. cit.
- [18] C. Morris and O. E. Van Slyke, "Empirical Bayes Methods for Pricing Insurance Classes," *Proceedings of the Business and Economics Section*, American Statistical Association, 1978.
- [19] Insurance Services Office, op. cit.
- [20] R. J. Freund and R. C. Littell, *SAS for Linear Models*, SAS Institute, 1980.
- [21] P. Diaconis and B. Efron, "Computer Intensive Methods in Statistics," *Scientific American*, May, 1983.
- [22] W. J. Conover, *Practical Nonparametric Statistics*, Second Edition, John Wiley and Sons, Inc., New York, New York, 1980. p. 280.

## 9. NOTES ON EXHIBIT I

## Exhibit I—Individual Classification Data and Results

## List of Variables

<b>CLASS</b>	— NCCI class code
<i>PI1</i>	— Policy year starting 4/78 payroll times <i>RATE79</i>
<i>PI2</i>	— Policy year starting 4/77 payroll times <i>RATE79</i>
<i>PI3</i>	— Policy year starting 4/76 payroll times <i>RATE79</i>
<i>XI1</i>	— Policy year starting 4/78 loss developed from first report to ultimate divided by <i>PI1</i>
<i>XI2</i>	— Policy year starting 4/77 loss developed from second report to ultimate divided by <i>PI2</i>
<i>XI3</i>	— Policy year starting 4/76 loss developed from third report to ultimate divided by <i>PI3</i>
<i>RATE79</i>	— NCCI rate in effect for 1979
<i>RATE82</i>	— NCCI rate in effect for 1982 (Before S.B. 1044)
<i>PAYROLL</i>	— Payroll for policy year starting 4/79
<i>ACTLOSS</i>	— Policy year starting 4/79 loss
<i>PI</i>	— $P_i$
<i>XI</i>	— $\bar{X}_i$
<i>ZI</i>	— $\bar{Z}_i$ (credibility for class <i>i</i> )
<i>UI</i>	— $\hat{\mu}_i$ (credibility estimate for class <i>i</i> )
<i>ELOSS</i>	— Expected loss for policy year starting 4/79 predicted using <i>UI</i> ( = $RATE79 * PAYROLL * UI * 1.053661$ )
<i>NCCIELOS</i>	— Expected loss for policy year starting 4/79 predicted using NCCI rates ( = $RATE82 * PAYROLL * 0.769384$ )







EXHIBIT I (continued)

Table with columns: CLASS, P11, P12, P13, P14, P15, P16, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P27, P28, P29, P30, P31, P32, P33, P34, P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P45, P46, P47, P48, P49, P50, P51, P52, P53, P54, P55, P56, P57, P58, P59, P60, P61, P62, P63, P64, P65, P66, P67, P68, P69, P70, P71, P72, P73, P74, P75, P76, P77, P78, P79, P80, P81, P82, P83, P84, P85, P86, P87, P88, P89, P90, P91, P92, P93, P94, P95, P96, P97, P98, P99, P100. Each column contains numerical data for a specific class.





EXHIBIT I (continued)

CLASS	PII	PIZ	PI3	XII	XI2	XI3	RATEI9	RATE82	PAYROLL	ACTLOSS	PI	XI	ZI	UI	ELOSS	MCRFIELDS
8833	4519887	7486166	7622172	0.580	0.528	0.746	1.62	1.33	2667172	2422120	18638325	0.625	0.805	0.617	2633069	2222525
8835	1526394	1499754	1499754	1.207	0.763	0.629	3.62	3.28	476618	1205124	4695396	0.866	0.499	0.724	1815073	1317573
8837	3262135	841591	841591	0.994	0.930	0.510	5.93	6.31	126289	827006	2497536	0.662	0.368	0.610	533789	466841
8868	5681087	8037452	8037452	0.265	0.661	0.502	0.94	0.26	1155613	625204	23762706	0.590	0.000	0.378	378248	301072
9011	3137223	3137223	3137223	0.550	0.769	0.658	6.20	5.87	655949	3823161	10576100	0.330	0.691	0.684	2933284	3568815
9033	64644	374776	374776	0.325	0.255	2.667	3.35	3.82	16566	156485	1450384	0.876	0.048	0.593	338859	46847
9040	2455879	4215752	4440302	0.581	0.872	1.045	5.12	5.55	396510	1987999	11113752	0.877	0.701	0.789	1753263	1917773
9052	4660359	4742673	4784389	0.567	0.635	0.774	6.19	4.77	698409	257158	14187421	0.666	0.468	0.645	2933682	3063357
9058	1938053	1688132	1060873	0.768	0.797	0.959	4.02	3.76	462289	1704957	4687570	0.822	0.499	0.702	137390	1871427
9060	2603563	1817406	1933050	0.607	0.395	0.635	3.99	2.84	664463	1387950	6478019	0.556	0.578	0.567	1583380	1863056
9061	1842885	1646881	1285497	0.858	0.536	0.593	3.98	2.80	538152	1284669	4575032	0.596	0.493	0.589	1162536	1150280
9063	462592	447239	453499	0.447	0.496	0.569	2.91	3.11	285628	233017	8338298	0.502	0.624	0.564	1047274	1527440
9071	3820742	1718237	1293910	0.298	0.807	0.846	5.95	5.21	763584	2535853	35248700	0.370	0.472	0.746	1847270	1574710
9102	1662642	118571	1067543	0.970	0.340	0.294	4.33	2.99	183540	4402062	33648700	0.663	0.814	0.533	5950584	6198270
9103	46305	74714	65678	0.292	0.613	0.467	5.35	4.26	16365	6552	1866637	0.562	0.066	0.581	411081	13349
9154	710004	63605	568681	1.037	0.507	0.152	2.87	2.07	264885	267078	1915141	0.598	0.282	0.587	470137	433417
9156	78026	56303	44321	0.169	0.022	1.659	0.95	1.03	84759	723697	1826650	0.528	0.066	0.580	492189	67583
9170	474569	58332	540464	1.255	0.420	2.262	40.32	36.27	19701	723697	1553334	0.620	0.262	0.592	243259	264815
9178	156237	12295	183758	0.353	0.220	0.132	8.16	5.48	12558	37990	4829970	0.163	0.097	0.542	58493	55168
9179	666833	65097	668938	0.552	0.458	0.425	20.13	15.10	36808	227424	2002849	0.479	0.301	0.351	407014	594580
9180	477833	52991	408398	0.132	0.792	0.700	12.32	9.73	29940	59428	1310942	0.372	0.629	0.363	233914	21878
9182	50514	32395	26316	0.261	0.266	0.278	6.78	2.70	1582	67066	80953	0.369	0.182	0.497	111736	9046
9200	7151378	563133	704966	0.733	0.543	0.269	6.33	5.66	110169	484504	2155347	0.589	0.316	0.584	443068	678658
9202	403352	736133	460426	1.041	0.409	0.469	6.33	5.37	80023	931210	1617670	0.662	0.259	0.403	320255	338446
9403	5951168	5701983	5470534	0.593	0.768	0.413	17.59	14.32	300128	3556465	1123666	0.594	0.783	0.591	3288505	3337155
9410	4074116	4801133	4587642	0.852	0.490	0.534	7.35	5.38	319375	2145720	13462831	0.551	0.740	0.559	1388558	1600005
9419	1780055	1857359	1592850	0.682	0.246	0.333	5.10	3.05	348285	1093720	5180480	0.523	0.524	0.499	933468	800911
9521	988453	1253359	1099400	0.528	0.713	1.222	4.78	3.76	269271	1002330	3724233	0.794	0.442	0.676	933468	1216694
9522	9884738	967683	83211	0.481	0.643	0.396	4.82	3.13	183846	582485	2789633	0.512	0.373	0.356	454561	443236
9545	55891	31234	44175	0.173	0.130	1.306	2.90	2.10	2672	21388	113259	0.488	0.290	0.262	454561	443236
9548	378739	378739	32267	0.260	0.961	1.430	11.19	10.28	31828	207857	993519	0.366	0.178	0.433	237587	260132
9549	163332	156500	122235	0.303	1.02	0.331	5.80	16.59	26567	56806	423918	0.218	0.090	0.350	80228	91905
9556	678594	581220	514081	0.383	0.427	0.662	1.94	0.71	688585	330493	1775830	0.873	0.277	0.552	416441	393868
9600	11643	2750	1592	0.000	1.738	0.000	1.70	2.01	1108	143	5985	0.199	0.011	0.364	1161	1687
9620	573137	493516	451331	0.354	0.393	0.298	2.34	1.63	256258	606673	1517983	0.312	0.247	0.315	325704	330334