MERIT RATING FOR DOCTOR PROFESSIONAL LIABILITY INSURANCE

ROBERT J. FINGER

Abstract

Merit rating is the use of the insured's actual claim experience to predict future claim experience. This paper discusses merit rating for professional liability insurance for both individual doctors and group practices. The paper presents several different theoretical formulations for merit rating. Credibilities are stated in terms of the parameters of the risk process. The paper discusses several methods of estimating the key parameters, along with sample data. Finally, the paper discusses several practical considerations in the design of a merit rating formula.

1. INTRODUCTION

The use of an insured's past claim experience for prospective premium determination can variously be called experience rating or *merit rating*. Merit rating is common for workers' compensation and commercial liability coverages. Merit rating for individual insureds is less common, although "claim-free discounts" or accident surcharges for personal automobile insurance are widely used. Several insurers now use merit rating for doctor professional liability insurance.

Section 2 of this paper provides a general statement of the merit rating problem. Section 3 presents the mathematical formulation of the risk process. It also discusses alternative merit rating formulations in terms of the parameters of the risk process. Section 4 provides several methods for estimating the required parameters. It applies these methods to actual data. Finally, Section 5 discusses various practical problems in implementing a merit rating program. The paper deals with two related situations: (1) Claim-free discounts and surcharges for individual doctors, and (2) merit rating for group practices.

2. GENERAL STATEMENT OF THE PROBLEM

We assume that there is some classification plan that will determine a premium for a given doctor (or group of doctors). The classification variables may include medical specialty, types of procedures, geography, and teaching or part-time status. For groups, there may also be schedule rating credits.

Why do we also want to use merit rating? Generally speaking, we want to use the insured's actual claim experience if (and only if) it is an efficient predictor of future claim costs. The insured's own claim experience may provide additional information that is not included in the other rating variables. Below, we give some reasons why the class rating variables may not have captured all of the relevant information. Using additional information may produce more accurate rates.

In a competitive environment, more accurate rates will generate greater profitability for the insurer. From the insured's point of view, more accurate rates are also fairer. Better doctors (in the sense of being less claims prone) will pay less and poorer doctors will pay more. From society's point of view, merit rating (and more accurate rating, generally) will provide an incentive for loss prevention.

Merit rating should be considered to be a complement to the classification plan (i.e., other rating variables). The more accurate the class plan, the less meaningful individual claim experience will be, and vice versa. Assume, for example, that the presence of a particular factor makes an insured 10% more expensive. If that variable is used in the classification plan, every insured with that factor will pay 10% more. If that variable is omitted, insureds with that factor who are merit rated will pay somewhat more than those without the factor, but most likely they will not pay 10% more. This follows from the concept that most insureds will receive less than 100% credibility.

Why do Individual Costs Differ?

Why would we expect doctors to have different loss costs? It is well recognized that different specialties have widely differing costs. This probably results from a variety of reasons. Certain specialties, such as surgeons, perform a higher percentage of procedures that can have devastating results, if done improperly. For certain specialties, such as psychiatrists, it may be very difficult to prove the causal connection between negligent practice and adverse results for the patient. For certain specialties, such as general practitioners with no surgery, the average patient is much healthier and any negligence is less likely to do damage. Thus, most insurers classify doctors by specialty. For physicians, most insurers also classify by the type or amount of surgery performed.

This classification plan does not cover all possible variations in costs among doctors in the same specialty. Costs may also vary for three general reasons: (1) Limitations in the class plan, (2) exposure, and (3) competence. Each will be discussed below.

Limitations in the Class Plan

Most class plans group specialties into about 10 different rate groups. In addition to specialty, the grouping may depend upon whether a doctor performs various procedures. The reason for this grouping is a lack of credibility for many specialties and procedures. That is, the number of insured doctors and the number of claims for many specialties and procedures are low. The volatility of claim experience for these low-volume categories makes it difficult to determine their cost. It is also difficult to determine how many of a certain type of procedure were performed during a given year. Doctors are usually classified by whether or not they perform a procedure, not on the number of procedures.

This classification scheme can result in significant cost variation within a given rate group. For example, Group 0 may have a rate relativity of 70%; Group 1, 100%; and Group 2, 150%. Within Group 0, there may be specialties that have relativities of 50%, 60%, 70%, and 80%. Within Group 1, there may be specialties with relativities of 90%, 100%, 110%, and 125%. In addition, the exposure to certain procedures may vary significantly. For example, the performance of procedure A may shift a doctor's classification from Group 1 to Group 2. Some doctors may perform 10 A's a year and some may perform 50 A's a year. A more exact classification plan might base the premium on the number of A procedures during the year.

The classification plan also may not consider other cost variations. Costs vary significantly from state to state. Some of this is due to differences in statutory or case law. Some of the difference may also be due to differences in the liberality of juries, the quality of the plaintiff's bar, and the claims consciousness of patients. These latter differences may exist within a state. In particular, there may be differences between urban, suburban, and rural areas.

Exposure

There may also be cost differences among doctors related to differences in exposure. For example, some doctors may treat more patients or may engage in more high-risk procedures. In addition, the type of patient may be different. Some doctors may have richer or poorer clients, who may have higher or lower damages, should negligence occur. Some doctors may also accept higher-risk patients, which could affect both the frequency and severity of loss costs.

Competence

Finally, doctors undoubtedly differ in competence, which has many causes. Training and experience differ. Doctors vary in their adherence to continuing education and changing practice standards. Doctors vary in their dexterity, judgment, attention to detail, bedside manner, and supervisory skills. The style of practice (e.g., number of patients, number of prescribed tests) may vary. Some doctors may have alcohol, drug, or other psychological problems.

Generalized Mathematical Structure

Now that we recognize that costs can vary significantly within the classification plan, how do we structure the merit rating plan? Virtually all merit rating plans use an adjustment to the class rate. In many lines, this is called a "modification factor." The adjustment could also be a credit or surcharge, which is expressed as a percentage of the class rate.

Modification Factor Formula

Virtually all merit rating plans calculate the modification factor according to the following generalized formula:

$$M = Z\frac{A}{E} + 1 - Z,$$

where

$M =$ the modification factor, which is multiplied against the class
rate;

$Z =$ the credibility factor;

$A =$ the insured's actual claim experience; and

$E =$ the average claim experience for the class.

In practice, virtually always the credibility is limited to values between and including 0 and 1. Thus M is a weighted average of the insured's relative experience (to the class average) and the class rate. (We could have written the right-hand term as $(1 - Z) \times 1$.)

We can express the same concept in terms of a discount or surcharge, as a percentage of the class rate. The adjustment to the class rate, as a factor of the class rate, can be calculated by subtracting 1 from M:

$$\text{Adjustment} = M - 1 = \frac{A - E}{E} Z.$$

When $M < 1$, the adjustment will be negative, or a discount from the class rate. When $A = 0$, the insured has no claims. The "claim-free" discount is thus Z, the credibility.

Indeed, this may often be the easiest way to measure credibility. If we have claim data for two experience periods, with a substantial number of claim-free insureds in the first period, the cost of these insureds in the second period relative to the average cost for all insureds in the second period is the empirical claim-free discount and the empirical credibility.

The formula for M is a linear function of the insured's actual claim experience. It would theoretically be possible for M to be some other type of function. However, other functions do not seem to have been used in actual practice. Perhaps the linear function is the most intuitively reasonable function. In addition, where a linear function might not be useful, the definition of A is modified. For example, it seems unreasonable in some cases to charge the entire amount of a large claim; very often, the maximum chargeable claim size is limited in some manner. An advantage of the linear formulation comes in the estimation and interpretation of Z.

Merit rating plans differ in defining A, in calculating E, and in determining Z. The usual process is to first define A, or what data are to be used for the insured's claim experience. Once this is done, E usually can be handled in a straightforward manner; it represents the class average claim experience for the given definition of A. The specification of Z can be done in at least three ways: (1) Ad hoc, (2) risk theory, and (3) direct estimation.

Ad Hoc Credibility

First, credibility can be established on an ad hoc basis. For example, we could decide that 100 expected claims represented "full" or 100% credibility, and partial credibility was the square root of the ratio of expected claims to 100. We might inject some actuarial or statistical theory into the selection of the full credibility standard. (See, e.g., Longley-Cook [5] or Venter [14].)

Risk Theory Credibility

Second, Z can be developed from risk theory. We can use the famous credibility formula:

$$Z = \frac{P}{P+K}\ , \tag{2.1}$$

where P is a measure of exposure and K can be determined from the following equation:

$$K = \frac{\sigma^2}{\tau^2}\ , \tag{2.2}$$

where σ^2 is defined as the "process variance" and τ^2 is defined as the "variance of the hypothetical means." The process variance is the variance we would expect for the class average insured's experience, given P units of exposure. The variance of the hypothetical means is the inherent variability of mean claim costs for the insureds within the given class, adjusted for P units of exposure. Depending on our definition for A, it may be possible to determine numerical equivalents for the process variance and the variance of the hypothetical means.

Direct Estimation of Credibility

Third, we can estimate Z statistically from actual data. Potentially, we could use any statistical estimation procedure. It happens, however, that the use of linear regression results in the same credibility formula and parameter explanation as the risk theory approach.

Although the risk theory and regression approaches are very similar, it should be realized that actual results may differ. The real world may differ from our theory or our theory may only approximate the real world. The theoretical approach allows us to apply knowledge from one context to another context. For example, measurement of the variance of the hypothetical means for one company, state, or line of business may be a useful input to another company, state, or line of business. The theoretical approach also allows us to generalize actual findings. For example, we may extrapolate three-year data to a four-year experience period. We should remember, however, that the real test of merit rating is how accurately it prices insureds in practice.

Alternative Forms for Modification Factor

There are several general considerations in the design of a merit rating plan. (See, e.g., Tiller [11].) First, it should be readily understood by insureds, agents, and company personnel. Second, it should be reasonably simple to administer. Third, it should not allow for manipulation by insureds. Finally, it should strike a balance between stability and responsiveness. On the last point, any formula can be adjusted to give greater or lesser weight (i.e., credibility) to the insured's own experience. If too much weight is given, rates may fluctuate too much from year to year. If too little weight is given, the pricing system may not be as accurate as possible and loss prevention incentives are reduced.

Definition of Actual Experience

The first decision in formulating a merit rating formula is the definition of A, the insured's actual claim experience. Choices involve the length of the experience period and whether to use counts or amounts. The length may be thought of as the number of years of experience, but could also include exposure from multiple locations or states. If the actual claim count is used, it could be defined as the reported count, the closed-paid count, or some definition of a non-nuisance claim. For example, a non-nuisance claim could be a settlement for more than \$5,000 (CP5). If amounts are used, there may be some limitation on the maximum chargeable claim; there is also an option of including or excluding allocated loss adjustment expense, loss development, and incurred but not reported (IBNR) claims.

In the National Council on Compensation Insurance (NCCI) revised Experience Rating Plan, A is defined in terms of loss amounts, usually for three policy years. A is divided into primary and excess losses, with the first \$5,000 of each loss being primary, and the remainder excess. There is also a per claim limit of 2.5 times the average cost per serious claim, a per occurrence limit of twice the per claim limit, and a limit on the total cost of diseases. Experience generally is pooled for all NCCI states and all entities with at least 50% common ownership. E, the expected loss, is divided into primary and excess portions. E must also be adjusted for loss development and the loss limitations.

The Insurance Services Office (ISO) has similar experience rating plans for general and automobile liability. A is limited to basic limits loss amounts. There is an additional limitation on the maximum claim size, based on premium size. A provision for IBNR, based on exposure, is added to A. E is adjusted for the loss limits and loss development.

Existing Plans for Doctors

Several insurers use merit rating for doctors. The typical plan offers an individual doctor a certain percentage discount for each claim-free year. Chargeable claims usually are limited to non-nuisance settlements (e.g., claim closed for more than $5,000). There is usually a maximum discount, which applies after five or six claim-free years. One insurer offers lower discounts for physicians than surgeons. A doctor loses the entire discount when a claim is charged; the discounts accumulate thereafter for each new claim-free year.

Rules may differ according to the insurer of the claim. For example, some insurers give credit for claim-free experience with other insurers. The experience period may be actual policy experience or it may be any settlements during a given period, regardless of the occurrence or reporting date.

Several insurers offer merit rating discounts to groups of doctors, based on the following generalized formula:

$$\text{Adjustment} = M - 1 = \frac{A - E}{JE + K},$$

where

$E =$ the expected claim count,

$A =$ the actual claim count,

$J =$ a constant (e.g., 2), and

$K =$ a constant (e.g., 1).

E is calculated from the number of insureds by rating class for the group; there is a separate claim frequency factor for each rating class.

Some Truisms

Finally, we consider some implications of merit rating. In workers' compensation there is the concept of the "off-balance" in the merit rating plan. That is, the average modification factor is not necessarily 1.0. The average collectible rate for a class will not necessarily be the same as the class manual rate. Thus the manual rate must be adjusted for off-balance. This concept is important for doctor professional liability insurance, particularly if we adopt a claim-free discount-only approach. With only discounts and no surcharges, the average collectible rate will be less than the manual rate.

Taking another perspective, it is necessary for those who do not receive the discounts to pay for the discounts. If some insureds pay less than the average cost, some must pay more. Even if we do not call it a surcharge, the difference between the claim-free discount and the manual rate is the cost of not qualifying for the claim-free discount. For example, the claim-free discount might be 25%. A doctor who loses the discount will pay an additional 33%. Whether we call this a surcharge or the manual rate, the cost of a claim is still 33%.

Although we will estimate credibilities in a later section of the paper, it is worthwhile to consider the tradeoffs between discounts of various sizes. Exhibit 1 shows the required manual rate increase, given discounts of various sizes (10%, 20%, 30%, 40%, and 50%). The manual rate increase is dependent upon the percentage of insureds receiving the discounts. For example, if 90% of insureds receive a discount of 10%, the manual rate must be increased 9.9%. In other words, 10% of insureds pay 109.9% of the average and 90%

pay 98.9% of the average. We give a discount of 1.1% to the 90% that are claim-free and require the other 10% to pay an additional 9.9%.

3. ACTUARIAL THEORY

As we have seen, the first step in formulating a merit rating plan is to define A, the insured's actual claim experience. Once that is done, usually it is straightforward to determine E, the average claim experience for the insured's class. The most complicated and difficult part is to determine Z, the credibility to attach to the insured's experience.

This section discusses various risk theory formulations for credibility. Although these formulations may not replicate the real world, they are useful in several ways. First, they provide a conceptual basis for understanding the statistical validity (i.e., credibility) of claim experience. Second, they provide a means to formulate credibilities when directly relevant claim experience is not available. Finally, they provide insight into the process of estimating credibilities.

In developing the following formulas, we will want to consider both claim counts and claim amounts. We also will want formulas for a single exposure period as well as multiple periods. There is no limit to the number and sophistication of formulas that can be developed; even so, we probably have included formulas that may be too difficult to test in practice.

The Basic Risk Process

We begin with a simple risk process and add various layers of complexity. We will develop formulas for variances. With few exceptions, the means are obvious and therefore omitted.

Assume that we have one doctor insured for one exposure unit (of time). We define N as a random variable for the number of claims for the period. We assume that N has a mean of λ. We assume that each claim has a claim size distribution S, with mean μ and coefficient of variation squared α. We also define T as the sum of individual claim amounts, or the total losses for that doctor for that exposure unit. If we assume that N and S are independent, we can calculate the variance of T from the moments of N and S.

$$\mathrm{Var}(T) = \mathrm{E}\,[N]\,\mathrm{Var}(S) + \mathrm{Var}(N)\,\mathrm{E}^2\,[S]\,.$$

We use the notation $\mathrm{E}[x]$ as the expected value of x. We previously defined α as $\mathrm{Var}\,(S)/\mathrm{E}^2\,[S]$. If we make the additional assumption that N is Poisson distributed, then $\mathrm{Var}\,(N) = \mathrm{E}\,[N] = \lambda$. Thus we have a fundamental risk theory formula:

$$\mathrm{Var}(T) = \lambda\mu^2\,(1 + \alpha)\,. \tag{3.1}$$

We can extend this formula to P exposure units. We assume that the same parameters apply to each exposure unit. Generally speaking, we can replace λ by $P\lambda$, if we assume that N is Poisson. Thus for P exposures, we have:

$$\mathrm{Var}(T) = P\lambda\mu^2\,(1 + \alpha)\,.$$

There are two important assumptions in this formulation: That the count and amount distributions are independent, and that the count distribution is Poisson. To the extent these are not true in practice, our use and interpretation of these formulas may be faulty. If we do not assume independence, we can still calculate the variances using covariance terms. This will be complicated, particularly when we make the formulas more complex. It seems reasonable in practice to assume independence, as long as we remove nuisance or closed-without-payment claims.

The Poisson assumption is very significant, particularly for the property that its mean equals its variance. The Poisson distribution arises from a process that satisfies three conditions: (1) events in two different time intervals are independent, (2) the number of events in an interval is dependent only on the length of the interval, and (3) the probability of more than one event occurring at the same time is zero. (See Beard [1], Chapter 2.) In practice, these conditions might be violated if there were some catastrophe (or contagion) or if an individual's claim frequency depended on its past history. As an example of the first case, we might have suits for breast implants or for the transmission of AIDS (acquired immune deficiency syndrome). As an example of the second case, we might have a plaintiff's attorney developing a series of suits against a practitioner, related to multiple incidents of unnecessary surgery or sexual misconduct with patients. For the most part, the Poisson assumption seems reasonable in practice, but we must be aware when it does not apply.

It would be possible to assume that N followed some other distribution with two parameters. The practical consequence of this, however, would be to add one more parameter that we would need to estimate. The interpretation of this parameter likely would overlap with the interpretation of other parameters, to be explained below. In addition, the estimation of this parameter might require data from an additional time period, which might be difficult to obtain.

Heterogeneity in the Insured Population

The above formulations assume that we know the parameters for the given doctor. We have calculated the "process variance." By the nature of merit rating, we assume that doctors will vary in their inherent claim costs. Thus we need to expand the formulation to add this heterogeneity. Conceivably, any of the above parameters could vary among the doctor population. We will assume that only the mean claim frequency varies among doctors; this should add sufficient complexity for practical purposes. We define a new random variable, χ, to have a mean of 1 and a variance of β. We will refer to β as the "structure variance." It is the (weighted average) variance of the insured population means (relative to the overall population mean). β probably varies from insurer to insurer. β also may change over time. For use in merit rating, β must be defined for the given insurer for the given experience period.

For any given doctor, the mean claim frequency is assumed to be $\lambda\chi$. We can incorporate these assumptions into our formulation by using a fundamental property of conditional probabilities:

$$\mathrm{Var}(N) = \mathrm{E}_\chi\,[\mathrm{Var}(N \mid \chi)] + \mathrm{Var}_\chi\,(\mathrm{E}\,[N \mid \chi])\,.$$

If we assume a Poisson process, we have $\mathrm{Var}(N \mid \chi) = \lambda\chi$. We can rewrite the last equation as:

$$\mathrm{Var}(N) = \mathrm{E}_\chi[\lambda\chi] + \mathrm{Var}_\chi(\lambda\chi)\,.$$

With the expectations taken over the variable χ, λ is a constant and can be taken outside of the operator. The variance of a scalar times a random variable is the scalar squared times the variance of the random variable. We previously defined $\mathrm{E}[\chi] = 1$ and $\mathrm{Var}(\chi) = \beta$. Thus we can rewrite the previous equation as:

$$\mathrm{Var}(N) = \lambda + \beta\lambda^2\,.$$

For P exposure units with the same parameters, we have:

$$\mathrm{Var}(N) = P\lambda + \beta(P\lambda)^2\,.$$

For the total amount, T, for a single exposure unit, we have:

$$\mathrm{Var}(T) = \mathrm{E}_\chi[\mathrm{Var}(T \mid \chi)] + \mathrm{Var}_\chi(\mathrm{E}[T \mid \chi])\,.$$

This can be written as:

$$\mathrm{Var}(T) = \mathrm{E}_\chi[\lambda\chi\mu^2(1+\alpha)] + \mathrm{Var}_\chi(\lambda\chi\mu)\,;$$

$$\mathrm{Var}(T) = \lambda\mu^2(1+\alpha) + \beta(\lambda\mu)^2\,. \tag{3.2}$$

For P exposure units with the same parameters, we have:

$$\mathrm{Var}(T) = P\lambda\mu^2(1+\alpha) + \beta(P\lambda\mu)^2\,.$$

Although we used the same notation, β, for the population heterogeneity for both counts and amounts, in reality there may be a different value in the two different contexts. For example, there may be differences in the inherent claim size distribution among insureds, as well as in claim frequency. Indeed, there may be a different numerical value for β, depending upon what claim data is used, such as reported count, CP5 count, or indemnity amounts limited to \$100,000. We will refer to β_C as the structure function, when the claim data is claim counts. We will refer to β_A as the structure variance, when the claim data is amounts.

For Equation 3.2, we note that the first quantity is the "process variance," or the variance given one exposure unit and known parameters, from Equation 3.1. The second quantity is the product of β, the variability in the insured population (given a mean of 1), and the square of $\lambda\mu$, which is the mean. This second quantity is the "variance of the hypothetical means." The $\lambda\mu$ term is a scalar that results from the variance calculation. Indeed, we can rewrite the first term, dividing by the square of the scalar, as:

$$\frac{(1+\alpha)}{\lambda} .$$

This quantity represents the process variance relative to the mean, just as β is the structure variance relative to the mean. We will use the term "relative variance" to be the ratio of a variance to the square of the mean. It is the coefficient of variation squared.

The Basic Credibility Formula

Using the fundamental formula for conditional probabilities, we can write Var (T) as:

$$\mathrm{Var}(T) = \mathrm{E}_\chi\,[\mathrm{Var}(T \mid \chi)] + \mathrm{Var}_\chi\,(\mathrm{E}\,[T \mid \chi]) .$$

This is the same form as:

$$\mathrm{Var}(T) = \sigma^2 + \tau^2 .$$

Here σ^2 is the average process variance and τ^2 is the variance of the means of the insured population. If we define τ^2 and σ^2 in terms of one exposure unit, our credibility Formula 2.1 becomes:

$$Z = \frac{\tau^2}{\sigma^2 + \tau^2} . \tag{3.3}$$

It is important to note that the denominator of the credibility formula is the total variance for the insured experience. Thus we have a general formula for credibility that conforms to our risk theory model of the claim process. For claim counts, we have $\sigma^2 = \lambda$ and $\tau^2 = \beta_C \lambda^2$. Dividing through by λ we have:

$$Z = \frac{\beta_C \lambda}{1 + \beta_C \lambda}. \tag{3.4}$$

If we divide through by $\beta_C \lambda$, we get the generalized formula, $1/(1 + K)$, with:

$$K = \frac{1}{\beta_C \lambda}.$$

For P exposure units, we substitute $P\lambda$ for λ above. This gives us an extra P in the τ^2 terms. By the same operations, we arrive at the generalized formula for $Z = P/(P + K)$, with the same K as above.

It will be useful to write the credibility in terms of the expected claim count, $E = P\lambda$. Thus we have:

$$Z = \frac{E}{E + K'}, \tag{3.5}$$

where $K' = 1/\beta_C$.

If A is defined in terms of amounts, then $\sigma^2 = \lambda\mu^2(1 + \alpha)$ and $\tau^2 = \beta_A(\lambda\mu)^2$. Dividing through the general formula for Z by $\lambda\mu^2$ yields:

$$Z = \frac{\beta_A \lambda}{(1 + \alpha) + \beta_A \lambda}.$$

Dividing this through by $\beta_A \lambda$ leads to the formula for K:

$$K = \frac{1 + \alpha}{\beta_A \lambda}.$$

We can also see that the scalar term for the mean will appear, squared, in both the σ^2 and τ^2 terms. These items will cancel in the credibility formula. We will be left with a formula for K that is the following ratio:

$$K = \frac{\text{(Relative) Process Variance}}{\text{(Relative) Structure Variance}}.$$

For counts, the numerator is $1/\lambda$ and the denominator is β_C. For amounts, the numerator is $(1 + \alpha)/\lambda$ and the denominator is β_A.

It also will be useful to analyze the total relative variance. We remember that the total variance is $\sigma^2 + \tau^2$ and the relative variance is calculated by dividing the variance by the square of the mean. For the above credibility formulation, for counts, we have the following formula:

$$\text{Total Relative Variance} = \frac{1}{\lambda} + \beta_C .$$

We know that the Poisson relative variance is $1/\lambda$. Thus the excess relative variance, for this formulation, is β_C.

Risk-Shifting

One of the limitations mentioned in connection with the Poisson assumption was the changing of an individual's mean costs over time. This can be handled formally by an adjustment to the credibility formula. This phenomenon has been called by various names, such as "parameter uncertainty" (see Meyers [10]) or "risk-shifting" (see Mahler [6], [7] and Venezian [13]). An interesting application is presented by Meyers [10] concerning the merit rating of Canadian automobile insurance.

In effect, the basic risk theory formulation breaks down when exposure is added for a given insured. Instead of credibility increasing approximately in proportion to P, in the general credibility formula the increase is significantly less. There is an intuitive explanation. Since the insured's mean costs may change over time, there is uncertainty that its historical mean may be the same as its future mean.

This phenomenon can be modeled in the same manner that we modeled heterogeneity among different insureds. The heterogeneity parameter, of course, should be different. Instead of reflecting the differences among the insured population, it reflects the differences for a given individual over time.

We define δ as the variance of the individual insured's mean costs over time. We should note that it may be difficult to differentiate between β and δ. Both parameters reflect the differences in individual insured experience: β reflects those differences between individuals in the same period, and δ reflects differences between the same individuals in different periods. Since we do not have the opportunity to observe different experience for the same individual in the same period, there may be some ambiguity in the measurement process. We should also note that δ may have different numerical values, depending on the definition of the claim experience.

The main difference in the mathematics from the previous formulation is that the process variance is different. Instead of being λ for counts, it now becomes:

$$\sigma^2 = \lambda + \delta\lambda^2 .$$

For amounts, the process variance is:

$$\sigma^2 = \lambda\mu^2 (1 + \alpha) + \delta (\lambda\mu)^2 .$$

The formula for credibility, $\tau^2/(\sigma^2 + \tau^2)$, for counts, becomes:

$$Z = \frac{\beta_C \lambda}{1 + \delta\lambda + \beta_C \lambda} .$$

The total relative variance is $1/\lambda + \delta + \beta_C$. The excess relative variance is $\delta + \beta_C$. Dividing through by $\beta_C \lambda$, we can rewrite the last equation as:

$$Z = \frac{1}{1 + \frac{\delta}{\beta_C} + \frac{1}{\beta_C \lambda}}. \tag{3.6}$$

If we let $K = 1/\beta_C \lambda$ and we define $J = 1 + \delta/\beta_C$, then we have a general credibility formula, $Z = 1/(1 * J + K)$. For P exposure units, we can derive the equation:

$$Z = \frac{P}{PJ + K}.$$

We can also state the credibility in terms of E, the expected claim count:

$$Z = \frac{E}{\mathrm{E}J + K'}, \tag{3.7}$$

where J has the same definition as above and $K' = 1/\beta_C$, as before in the basic credibility formulation, Equation 3.5.

For amounts, we derive the credibility formula:

$$Z = \frac{1}{1 + \frac{\delta}{\beta_A} + \frac{1+\alpha}{\beta_A \lambda}}.$$

This has the same form for J as for counts, $(1 + \delta/\beta_A)$, and the same K as for amounts in the basic credibility formulation.

We have the following changes from the basic formulation. The process variance is now larger, since there will be more variability in the individual insured's experience. The excess relative variance is the sum of δ and β. When we estimate β, we will have a smaller structure variance. Thus σ^2 is now larger and τ^2 is now smaller. The credibility will be reduced.

We should note that the maximum credibility is $1/J$. In effect, we are saying that, since the individual's mean cost may be different in the future than it was in the past, we may not be insuring the same risk and, hence, we will always give some credibility to the class average.

Heterogeneity within the Insured

The rationale for the next generalization in the credibility formula does not apply to individual doctor experience. It may be useful, however, in developing formulas for group experience. This generalization has been used by NCCI. As with risk-shifting, we have a situation where adding exposures does not yield as much credibility as if all exposures had the same underlying risk parameters.

In the first credibility formulation, we developed a parameter, β, which described the variance in the insured population. We now want to develop credibility for groups. If all of the doctors in the group were equally good or equally bad, we could apply the first credibility formulation, using P to represent the exposure for the number of doctors in the group. In all likelihood, however, the group will have some better doctors and some poorer doctors. Some of the underlying risk factors, such as geography, might apply to the entire group; other risk factors, such as training and experience, would be different for different members. If the composition of the group were entirely random with respect to the insured population, we could rate each doctor individually; there would be no additional statistical validity to the group experience, apart from the individual doctor experience.

We define γ as the variance of mean costs (adjusted by class) within a given group or insured. We expect that $0 < \gamma < \beta$. In other words, the variability within the group is not as large as the insured population, but it is not zero. As with β and δ, γ may have different numerical values for different formulations of claim experience, such as counts or amounts.

The variance of the insured population means is different than before. Here the "insured population" is groups with a degree of heterogeneity. Some part of the variance will be proportional to the number of exposures (i.e., each exposure has the same parameters, for which the variances are additive) and some part will be proportional to the square of the number of exposures. We can write this as:

$$\tau^2 = \lambda\gamma + (\beta - \gamma)\,\lambda^2 .$$

We know from the previous development that, for counts:

$$\sigma^2 = \lambda + \varepsilon\,\lambda^2 .$$

We also know that the total variance, ignoring the possibility of $\delta > 0$, is $\lambda + \beta_C \lambda^2$. From this we can solve for $\varepsilon = \gamma(\lambda - 1)/\lambda$. Thus we have:

$$\sigma^2 = \lambda + \gamma\,(\lambda - 1)\,\lambda .$$

Using the general formula for credibility and dividing by $\beta_C \lambda^2$, we have:

$$Z = \frac{\left(1 - \dfrac{\gamma}{\beta_C}\right) + \dfrac{\gamma}{\beta_C \lambda}}{1 + \dfrac{1}{\beta_C \lambda}} .$$

For P exposure units, we have:

$$Z = \frac{\left(1 - \dfrac{\gamma}{\beta_C}\right) P + \dfrac{\gamma}{\beta_C \lambda}}{P + \dfrac{1}{\beta_C \lambda}} .$$

In terms of the expected count, E, we have:

$$Z = \frac{\left(1 - \dfrac{\gamma}{\beta_C}\right) E + \dfrac{\gamma}{\beta_C \lambda}}{E + \dfrac{1}{\beta_C \lambda}} . \qquad (3.8)$$

We can write this in a more general form:

$$Z = \frac{(1 - I)\,E + I}{E + K'}\,, \tag{3.9}$$

where $I = \gamma/\beta_C$ and K' has the same form as the previous formulations for E.

The interpretation of this formula depends on the specific values for the given parameters. As we will see below, this formula may produce higher credibilities than the previous two formulations, when the expected claim count is low. Excepting this situation, however, we can relate this formula to the previous formulations. We see that the $(1 - I)$ term reduces the effectiveness of additional exposures. Since the exposures within a group are heterogeneous, we would not expect to generate as much credibility per additional exposure, compared to the situation where all exposures had the same parameters. We can also see that τ^2 is generally lower than it is in the other formulations, because we have incorporated some of the population heterogeneity into the process variance for the insured.

The NCCI credibility formulation includes both risk-shifting and insured heterogeneity. The credibility may be developed from the formulations for σ^2 and τ^2. As a practical matter, the sample data we used for this paper is not sufficient to separately estimate all of the required parameters.

4. PARAMETER ESTIMATION

There are several different approaches that we can take to estimate the appropriate credibility. We can estimate the credibility directly or we can estimate the credibility parameters. We can estimate credibility directly by using claim-free discount data or a regression method.

Direct estimation basically requires that we have data for the same insureds during at least two different experience periods. This is probably the best approach to estimating credibility, because our theoretical models may not always apply to the real world. We may also estimate credibility by estimating the parameters in the formulas that we developed above. This may be our only alternative if we do not have sufficient data. Even if we estimate credibilities directly, we may want to estimate the theoretical parameters, in order to gain more insight into the process.

Direct Estimation of Credibilities

We will define some generalized notation to simplify the estimation equations. Assume that we can measure the experience of Q insureds over two different experience periods. For each insured, i, we define x_i , the *relative* cost ratio for the first period. For example, if we have 10 claims for 100 insureds, the average claim frequency is 0.1. For an insured with one claim, $x_i = 10$. For an insured with no claims, $x_i = 0$. We define y_i as the *relative* cost ratio for the second period. We also define w_i as the weight that we will apply during the estimation process. We can think of w_i as being the relative exposure of that insured to the total group of insureds. Some of the following equations will have a special meaning where the sum of the w_i is 1.0.

We want the x_i to be defined in the same manner as A, the actual claim experience that we are using in the modification factor formula. We want to test the predictability of the actual experience. It is possible that different definitions of x_i will give similar values for certain parameters, such as β. For example, rating based on reported counts might produce the same value for β_C as rating based on CP5 counts. We would expect the level of credibility to be different, however, since the reported count frequency will be much higher than the CP5 frequency.

We can use *any* y_i data to test the validity of the modification factor. Since, ideally, we want to test the actual cost of insured experience, our preference is to use insured amounts for y_i. As we saw above, however, the variability in results likely will be much higher using amounts than counts. Using amounts may give too much weight to outliers and render the estimation process ineffective. The y_i using counts, however, may not be directly related to insurance costs. For example, the inherent claim size distribution might differ among insureds.

Thus, we want the x_i to reflect the definition of A and the y_i to reflect the actual costs of insurance. We can make substitutions, if we understand the limitations that this might produce.

The simplest estimate for Z is the claim-free discount. Our notation can be made simpler by grouping all insureds by their claim experience in the first period. x_0 would be the relative cost in the first period for insureds with no claims. x_1 would be the relative cost for insureds with one claim, etc. y_0 would be the second period relative cost for insureds with no claims in the first period. Similar definitions would follow for y_1, etc. The weights would represent the percentage of insureds with no claims, etc. in the first period.

The empirical claim-free discount is $1 - y_0$. This is the credibility that applies to this group of insureds. We have assumed that the credibility is the same for all insureds in the group. If this is not the case, the estimated credibility will be an "average" credibility for the individual in the group.

The stability of our estimate will depend upon how many insureds were claim-free in the first period, as well as how volatile the claim experience is in the second period. Note that there is no particular requirement for measuring y_i in the same manner as x_i. We could try several measures of y_i, such as pure premium and different count definitions. (The y_i may not be directly related to the cost of insurance, however.)

This claim-free discount formulation is somewhat limiting, however, in that we do not use the experience of non-claim-free insureds. We could expect to get a better estimate by using more information.

Least Squares Regression Formulation

A more generalized formulation uses the modification factor, M_i, to estimate the second period experience:

$$\hat{y}_i = Zx_i + (1 - Z) .$$

In effect, we want the most appropriate credibility, Z, to convert the insured's first period experience into a prospective rate for the second period. We can derive a mathematically appropriate Z by selecting some criteria to minimize the differences between the predicted experience (M_i) and the actual experience (y_i). Although it is not the only possible criterion, least squares minimization is commonly used to determine Z. Thus we have the following formulation:

$$\min_Z C = \sum_i w_i (\hat{y}_i - y_i)^2 ;$$

$$C = \sum_i w_i (Zx_i + 1 - Z - y_i)^2 .$$

We can solve for Z by taking the partial derivative of C with respect to Z and setting the result equal to 0.

$$\frac{\partial C}{\partial Z} = \sum_i 2w_i (Z(x_i - 1) + 1 - y_i)(x_i - 1) .$$

We can separate out the terms that have Z and those that do not.

$$\frac{\partial C}{\partial Z} = 2 \sum_i w_i Z (x_i - 1)^2 + w_i (1 - y_i)(x_i - 1) .$$

When we set this equal to zero, the 2 drops out. We can put all the Z terms on one side of the equation and the non-Z terms on the other side. Since Z is a constant, we wind up with a ratio for Z:

$$Z = \frac{\sum_i w_i (x_i - 1)(y_i - 1)}{\sum_i w_i (x_i - 1)^2} .$$

If the sum of the w_i is 1.0, the denominator is the total relative variance and the numerator is the relative variance of the means of the insured population, the structure variance. If the w_i are the exposures for both x_i and y_i, and the sum of the w_i is 1.0, then the formula simplifies to:

$$Z = \frac{\left(\sum_i w_i x_i y_i\right) - 1}{\left(\sum_i w_i x_i^2\right) - 1} .$$

We can also use this formula for grouped rather than individual insured data, but we must define the groups by the first period experience. For example, we might divide the data into 10 groups, the first having the lowest loss ratios in the first period, etc. This approach can remove the undue impact of outliers. Strictly speaking, Z will be optimal for the selected group means, not for every insured.

Figure 1 graphically depicts the regression process. It shows the prior relative frequencies (x_i), the subsequent relative frequencies (y_i), and the modification factors (M_i), which are the fit of the regression line between the prior and subsequent experiences. The estimate based on the claim-free discount is almost the same as the regression estimate; it can be different, in some cases, because the regression considers the experience of all of the insureds.

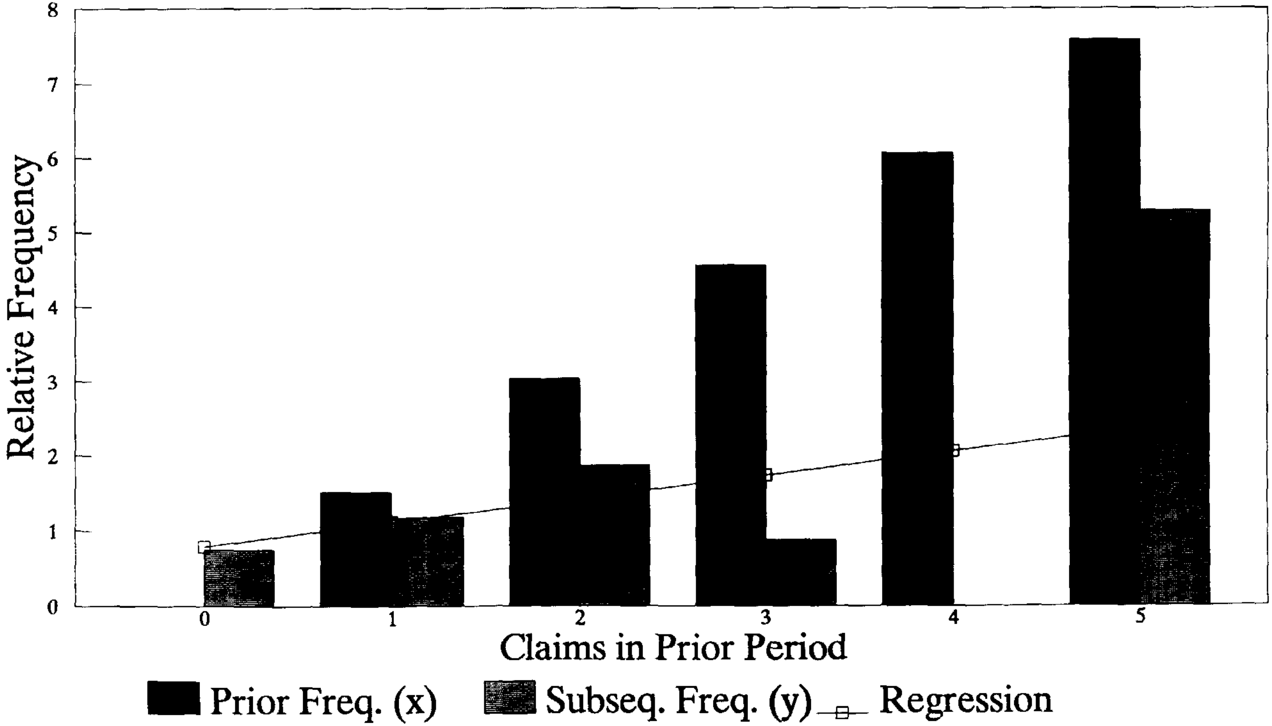
In certain cases, we may wish to pool data for which we know that the credibility is different for different insureds. This formulation would be:

$$\hat{y} = Z_i x_i + 1 - Z_i .$$

Since the Z_i vary for each insured, we cannot solve for a single value of Z. If we can formulate a reasonable function for Z_i, however, we can use the least squares approach to solve for the parameters of our Z_i function. Reasonable candidates for the credibility function can be developed from risk theory, as we showed in an earlier section. Given two periods of data, we would be limited to estimating one

FIGURE 1

RELATIVE CLAIM FREQUENCY



parameter. For example, we may assume that the appropriate credibility function is:

$$Z_i = \frac{\beta_C \lambda_i}{1 + \beta_C \lambda_i} . \tag{4.1}$$

where λ_i is the expected (mean) frequency for class i. We may use the regression approach to solve for β_C. In effect, we are determining the optimal β_C , if credibility does indeed follow the postulated function. If the selected function is not appropriate, we may not get a reasonable estimate for β_C. If the credibility function is complicated, we may not be able to calculate the optimal parameter from a simple equation. We might have to resort to numerical methods.

Estimation of Credibility Parameters

The parameters λ, α, and μ can be estimated from single-period experience. In fact, we do not even need individual insured experience to estimate them. (We do need individual claim experience to estimate α, but λ and μ may be readily available from aggregate data or other projections.) If we can somehow obtain estimates for β, δ, or γ and we also have confidence in the correct form for the credibility function, we do not need to obtain two periods of individual risk data to test the credibilities.

Estimates for the Structure Variance, β

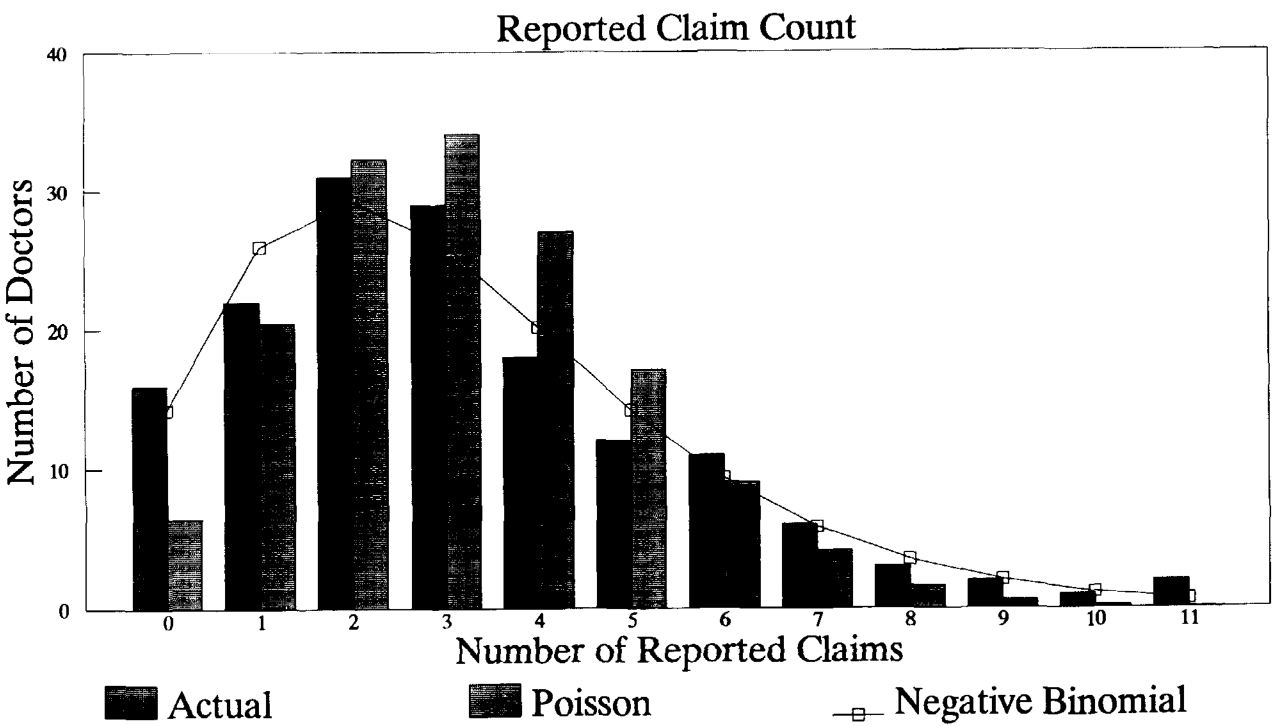
The simplest estimate for the structure variance comes from the basic properties of the Poisson distribution. Since we know that the mean and variance of the Poisson are the same, any "excess" variance in the data can be thought of as being the structure variance.

$$\beta_C = \frac{\mathrm{Var}(N) - \lambda}{\lambda^2} . \tag{4.2}$$

Figure 2 displays an example. It shows the actual number of doctors with a given number of claims. It also shows the theoretical number of doctors who would have had that many claims, had the

FIGURE 2

FREQUENCY DISTRIBUTIONS



distribution been Poisson. Under some generalized assumptions, incorporating the excess variance yields a negative binomial distribution, which is also shown. We see that the actual distribution is more dispersed than the Poisson assumption. There are far more doctors with no claims, and more doctors with only one claim, than the Poisson assumption would indicate. Of course, to balance out, there are also more doctors with large numbers of claims than the Poisson assumption would indicate.

The negative binomial provides a reasonably good fit to the data. It should be noted, however, that the excess variance method is greatly affected by the small number of insureds that will have very unusual experience. If we have a relatively limited sample, we would expect the excess variance estimates to be volatile.

Unfortunately, the structure variance may not be the only component of the excess variance. Other credibility formulations, such as risk-shifting and within-insured heterogeneity, also affect the excess variance. We can think of the excess variance as being a combination of all of these effects. Given a reliable estimate, the excess variance is probably an upper bound on the structure variance.

We obtained another estimate for the structure variance from the numerator in the regression approach, where the sum of the w_i is 1.0:

$$\hat{\beta} = \sum_i w_i (x_i - 1)(y_i - 1) .$$

If the w_i are the exposures, the formula simplifies to:

$$\hat{\beta} = \left(\sum_i w_i x_i y_i \right) - 1 .$$

This regression formulation probably is more reliable than the excess variance approach, because it is based on the predictability of actual data. This formula can be found in Woll [15] and can apply to any claim data (i.e., counts or amounts).

We can also apply this formula to grouped data, although we must group by the loss experience in the first period. We also would expect the grouping process to bias the estimate on the low side, since we are taking differences of group means. We could correct for this bias by multiplying by the ratio of the total relative variance for the individual insureds to the total relative variance of the groups.

Another estimator for the structure variance can be developed from the following relationship:

$$\beta = Z \frac{\mathrm{Var}(T)}{\mathrm{E}^2[T]} .$$

This can be used with a variety of inputs. The estimate for Z can come from claim-free discount data. The ratio on the right is the total relative variance. This can be calculated from one-period data. We can adjust the claim experience for all insureds by the mean experience and then calculate the variance over all insureds. This estimator is based on the general credibility formula, $Z = \tau^2/(\sigma^2 + \tau^2)$. It can be used for either count or amount data.

Another estimator is taken from Woll [15]. This was developed for count data where the structure function, χ, has a gamma distribution.

$$\hat{\beta} = \frac{y_1 - y_0}{y_0} . \tag{4.3}$$

Numerical Examples

We will present various numerical calculations, based on actual data. The data was developed from the experience of one insurer in one state, for insureds that were continuously insured for seven years on an occurrence form. The "prior" period consisted of the first five years and the "subsequent" period consisted of the last two years. The evaluation date was about four years after the inception of the last policy year.

For this insurer, most claims have been reported for the subsequent period, but many of these remain open. The large majority of claims from the "prior" period are closed. Data was available for the reported count, the closed-paid count, the CP5 count, and the basic limits amount, for both periods. Data was segregated for nine different class groups, based on the current classification plan by specialty. There are some rating variables that are not reflected in the class groupings.

Exhibit 2 shows numerical calculations for a number of the methods described above. This data includes the experience of 153 doctors in a particular rating group. For this exhibit, we have defined A, the actual claim experience, to be the number of CP5 claims in the five-year experience period. Ninety-one of the doctors (59.5%) had no CP5 claims in the first period. These doctors had 13 CP5 claims in the second period, for a frequency of 14.3%. The entire class had 29 claims in the second period, for a class frequency of 19.0%. The relative frequency for the claim-free doctors is 75.4%. Thus the claim-free discount, based on CP5 count, is 24.6%. (A claim-free discount can also be calculated for the other data items, such as reported count and pure premium.)

The CP5 frequency for the group is 0.660 and the CP5 variance is 0.969. The variance for a Poisson process also would be 0.660; thus the excess variance is 0.309. All of these numbers reflect the frequency of the actual data. For analysis purposes, it is easier to work with the "relative" variances, which are the actual variances divided by the square of the frequency. The total relative variance is 2.225. The Poisson relative variance is 1.515 (the reciprocal of the frequency). Thus, the excess relative variance is 0.710 ($=2.225-1.515$). We could also calculate the excess relative variance as the actual excess variance (0.309) divided by the frequency squared ($0.660 * 0.660$).

If we use the basic credibility formulation, β_C can be estimated from the excess relative variance, by Equation 4.2, as 0.710. This would imply a credibility of 0.319, from the formula $Z=\beta_C\lambda/(1+\beta_C\lambda)$. If we use the risk-shifting credibility formulation, the excess relative variance is the sum of β_C and δ_C. Thus, if we believe there is some risk-shifting, the excess variance method will overstate the estimate for β.

The regression method produces a credibility of 0.208. This estimate can be interpreted as the ratio of an estimate of β_C and the total relative variance, which is 2.225, as above. Based on the regression method, the estimate of β_C is thus 0.463 ($=0.208 * 2.225$). This might indicate that either: (1) δ_C is 0.247 ($=0.710-0.463$), or (2) the data is relatively unstable. Normally, we would think that the regression approach, which is based on two-period data, would produce better estimates for β and the credibility.

The claim-free discount data indicates a credibility of 0.246. This may imply a β_C of 0.548 ($=2.225 * 0.246$). We can also derive another estimate of β_C from the relative costs of claim-free and one-claim insureds in the second period, from Equation 4.3. This estimate is 0.556, as shown. As can be seen, the results for this class are relatively similar among the different methods above.

We also used first period reported count experience. We would expect the numerical amount of the credibilities to be different (because the frequency was different). The β_C estimates could be similar or different, depending upon whether the use of reported counts has the same predictability as the use of CP5 counts. (For example, does the fact of a CP5 claim imply a higher prospective cost than the fact of a reported claim?) For this data set, the β_C estimates were quite similar for both reported counts and CP5 counts.

We also used claim-free discount data based on reported counts and pure premiums. As we might expect from risk theory concepts, the pure premium data was more volatile.

In merit rating, we want to vary premiums based on differences in prospective costs. Ideally, we would measure the cost differences in terms of pure premiums. Due to the volatility of claim size data, however, estimates based on pure premiums will be much more volatile than estimates based on claim counts. It may be more efficient to estimate credibilities or parameters, such as β and δ, from claim count data. We can either use these parameter estimates directly, by assuming that there is no inherent variation in claim sizes among insureds within the given class, or we can use adjusted values.

We can think about the optimal estimation procedure by considering the regression approach. There, the x_i are best defined by the claim experience used for merit rating. For example, we may use CP5 counts. The y_i are best defined by actual insurance costs. Our structure function estimate for this situation could be given the following notation: β_{CA}, where the first subscript defines the prior period data and the second subscript defines the subsequent period data.

For some of the classes, the number of insureds was small or the actual claim experience was erratic. This raised dual questions: (1) how do we determine β for the smaller classes, and (2) does β vary by class?

Exhibit 3 shows the calculation of the excess relative variance by class for reported counts. Assuming the basic credibility formulation, the excess relative variance is an estimate of β. Several classes have β_C of about 0.6 or 0.7 and several are in the 0.2 to 0.35 range. This might indicate that the β_C vary by class. Class 6, however, has the lowest excess relative variance of 0.215 for reported counts. We saw in Exhibit 2 that its β_C for the CP5 count was about 0.5. Thus the variations by class may be due to random fluctuations in the data.

We can also estimate β by the other methods. Exhibit 4 estimates β using the claim-free discount method. For two classes, the subsequent claim experience for claim-free insureds was actually worse than the average. This would imply a negative value for β. We also note from Exhibit 4 that the claim-free discount based on CP5 counts is significantly different from the claim-free discount based on pure premiums, for several of the classes. Part of this probably is explained by the greater volatility of pure premium data. We also obtained varying β estimates by class from the regression approach.

In reviewing the individual calculations, it appears that much of the volatility is caused by the relatively low number of insureds and claims. We should also note that variance methods give exceptional weight to outliers. There may be a difference in β from class to class, but it does not appear to be statistically significant.

We also pooled the data, for all classes, for the regression and claim-free discount methods. We assumed that the credibility function was the same as Equation 4.1, with λ_i being the expected claim frequency for the class. For the claim-free data, for insureds grouped by CP5 in the first period, the estimate of β_{CC} was 0.54, based on CP5 counts in the second period, and β_{CA} was 0.59, based on pure premiums in the second period. For insureds grouped by reported count in the first period, β_{CC} was 0.54, based on CP5 counts in the second period, and β_{CA} was 0.36, based on pure premiums in the second period.

For the regression approach, for insureds grouped by CP5 in the first period, β_{CC} was 0.51, based on CP5 count in the second period. When insureds were grouped by the reported count in the first period, β_{CC} was 0.50, based on the reported count in the second period.

Estimates for δ and γ

We have mentioned that all three parameters, β, δ, and γ, arise in a similar manner, to explain additional variance beyond a Poisson process. The basic formulation for δ is a shifting of relative claim costs for the individual insured over time. With more years of data, it might be possible to estimate this parameter. The basic formulation for γ is heterogeneity among different doctors within the same insured group. We could estimate this parameter if we had credible data for at least several different size groups and if we assumed that the same heterogeneity applied to all size groups. In fact, the NCCI has used a similar approach to calibrate all of its credibility parameters. It divided risks into various size groups; it estimated optimal credibilities for the different groups; and it fitted these optimal credibilities to a credibility function.

We can use the above numerical example to see whether δ might be significant. If the risk-shifting formulation is correct, the total variance will include a provision for β and δ, as well as the usual

Poisson variance. The excess variance estimate should be the sum of β and δ. The numerator of the regression credibility estimate, however, should include only β. Thus we can compare the two estimates to see if the excess variance estimate is significantly larger. Exhibit 5 shows this comparison for the classes for which the individual estimates were satisfactory. In some cases the excess variance estimate is higher and in some cases it is lower! It does not appear that the excess variance estimate is consistently higher. In practical terms, this might imply that an individual doctor's inherent (relative) risk does not change appreciably over time.

Other Published Data

Two published papers, Ellis [2] and Venezian [12], give some estimates of credibility parameters. The Ellis data included the number of closed-paid claims against doctors in various specialties, for four years, 1980 through 1983, in New York State. It is not clear what the authors used for exposure, but it would appear to be licensed doctors. The authors published theoretical prospective mean frequencies for doctors, in a given specialty, that had various numbers of closed-paid claims within a five year experience period. Comparing the prospective frequencies, for doctors with no claims and all doctors, yields the five-year claim-free discount, or credibility, for the five-year experience.

Except for some minor differences, probably caused by slightly different methods of estimation, we can generate the same credibilities using the procedures outlined above. The Ellis method is equivalent to a credibility formula of $\beta\lambda/(1+\beta\lambda)$, where β is the excess relative variance and λ is the five-year mean frequency. We have estimated the excess relative variance from the claim count distribution given in the paper. The results are shown in Exhibit 6.

For most of the specialties, the excess relative variances are much higher than those estimated from the data set used in this paper. There are several reasons for this. First, it is not clear what exposure was used. If it was licensed doctors, which includes retired, part-time, and government-employed doctors, a substantial number of the doctors would have virtually no claim exposure; we would expect the excess variance to be higher than that for full-time doctors in private practice.

Second, the exposure does not differentiate among other class variables. An insurer's premiums could vary significantly within a given specialty, due to class relativities, geographical relativities, and other rating variables. It is interesting to note that the specialties that are more likely to be grouped into one insurance class, such as anesthesiology, general surgery, neurosurgery, obstetrics, and urology, have much lower excess variances.

Third, New York State could have more geographical variation in costs than the state our data was taken from. Fourth, some doctors are not insured voluntarily. These doctors may have an extreme number of claims, which would produce a much higher excess variance than an insured population. In any case, we might use this data as an upper bound on β.

The Venezian data was taken from the Pennsylvania Medical Professional Liability Catastrophe Loss Fund, which covers both excess losses (attachment points have varied over time) and late reported claims (over four years). Although this data came from insured doctors, the exposures were estimated by the authors. The excess relative variance was estimated from the data in the paper and is shown by specialty in Exhibit 6. With one exception, the excess variances are smaller than in Ellis. Most of the above comments apply to these estimates, as well.

5. PRACTICAL CONSIDERATIONS

This section will consider several practical considerations in the design of a merit rating plan. These include:

- Is it better to use counts or amounts?
- Is it better to use the reported count or the CP5 count?
- What is the best length of the experience period?

- Is the credibility different if we offer only discounts and have no surcharges?
- How do we calibrate the expected costs?
- What if we use non-optimal credibilities?
- How do we establish a formula for insured groups?

Counts or Amounts?

The NCCI and ISO use amounts, rather than counts, in their merit rating plans. The situation for doctor professional liability insurance, however, may call for a different approach. We can analyze the situation by reference to the formula for K, in the basic credibility formulation:

$$K = \frac{1+\alpha}{\beta_A \lambda}.$$

The K for counts is similar, but a 1 replaces the $(1+\alpha)$ in the numerator and β_C may be different from β_A.

For amounts, the K will be $(1+\alpha)$ times larger, if the β are the same. For one exposure unit, the credibility of claim amount experience will be only about $1/(1+\alpha)$ times as much. To the extent an individual's experience is relatively better or worse than the average, it will receive credit for only about $1/(1+\alpha)$ of that difference. The claim-free discount also will be only about $1/(1+\alpha)$ as much.

It is likely that claim severity varies among insureds within the same class. If so, and if frequency and severity are not negatively correlated, we would expect the β to be larger for amounts than for counts. Most likely, however, the β will not increase by as much as $(1+\alpha)$. If we use indemnity amounts limited to \$100,000, $(1+\alpha)$ may be about 2 for doctors. For indemnity amounts limited to \$200,000, $(1+\alpha)$ may be about 2.5. We would expect that β_A for indemnity amounts would be only marginally higher than β_C for counts. Thus using indemnity amounts rather than counts would cut the credibility and the claim-free discounts about in half.

We could also use combined indemnity and allocated loss adjustment expense, limited to various amounts. The $(1 + \alpha)$ terms would be somewhat lower when allocated expenses are included. Credibilities would be much closer to those for indemnity only amounts than those for counts.

Which Count?

There are several choices for claim counts. We could use reported claims, closed with indemnity claims, closed with either indemnity or expense claims, or possibly some non-nuisance claim definition, such as CP5. We can analyze this situation by reference to the basic credibility formula, defined in terms of the expected count, E:

$$Z = \frac{E}{E + K} ,$$

where $K = 1/\beta_C$. We note that credibilities generally will be higher for higher expected counts. We saw from the sample data above that the β's for reported counts and CP5 counts tended to be about the same. This result might not be universally applicable, but we might conclude that the β's would not increase in the same proportion. Thus reported counts would generate more credibility and higher claim-free discounts. If the β's happened to be the same, the credibility for reported count experience might be three to five times higher, depending on the claim frequency for the class and the length of the experience period.

Using reported counts, however, may cause consumer relations problems. It is common for every surgeon in the operating theater to be named in a suit, even if only one is likely to be responsible. Most claims will be closed without a payment or for a nuisance-value payment. Even if more costly doctors are sued more often (which is the logical consequence of the β's being the same), it may be difficult to charge an individual doctor more, just for being named in a suit.

On occurrence policies, in particular, charging for reported claims may also deter or delay the reporting of claims. This could have adverse consequences for both the claim settlement process and the ratemaking process.

From a pricing perspective, using reported counts probably is preferred. Practical considerations, however, may favor a CP5 program.

What Should be the Length of the Experience Period?

Both the NCCI and ISO use a three-year experience period as a standard. Claim frequency for doctors, however, is quite low, particularly when using CP5 counts. Current doctor merit rating programs typically give a certain discount for each year of claim-free experience. This is a reasonable approach, although the discount percentages should vary by specialty. Recall that the basic credibility formula is

$$Z = \frac{\beta_C P\lambda}{1 + \beta_C P\lambda}$$

for counts, for P exposure units. For each additional year of claim-free experience, the credibility will increase about $\beta_C \lambda$. Assuming $\beta_C = 0.5$ and $\lambda = 0.02$ (for one year), the claim-free discount would be about 1% per year. After 10 years, the discount would be 9.1%. For a higher-rated specialty, where $\lambda = 0.1$, the first year discount would be about 4.8%, the second year, an additional 4.3%, the third, 3.9%, the fourth, 3.7%, and the fifth, 3.3%, for a total of 20%.

The above credibility formulation assumes that the doctor's relative cost remains the same over time; i.e., there is no risk-shifting. If there is risk-shifting, and the δ parameter is relatively high compared to β, the additional discounts for additional claim-free years will decline quickly.

Discount Only Plans

Current merit rating plans for individual doctors have claim-free discounts but no surcharges. What should the credibilities be for this type of program?

We can use the same regression formulation to select an optimal credibility. Let w_0 be the percentage of doctors with no claims in the first period and w_1 be the remaining doctors. The modification factors are $1 - Z$ and 1, respectively. Using these modification factors, however, will lead to an "off-balance." That is, the collectible premium will be less than the manual premium. The amount of the off-balance will be $w_0 Z$. The manual rates will be:

$$\hat{y}_0 = \frac{1 - Z}{1 - w_0 Z} \text{ , and}$$

$$\hat{y}_1 = \frac{1}{1 - w_0 Z} \ .$$

We can write the optimization function as:

$$\min_Z C = \sum_i w_i \, (\hat{y}_i - y_i)^2 \ .$$

Taking the partial derivative with respect to Z and setting it equal to zero, we obtain the optimal $Z = (1 - y_0)/(1 - y_0 \, w_0)$. This result can also be obtained in another manner. Since $y_0 \, w_0 + y_1 \, w_1 = 1$, it follows that $y_1 = (1 - y_0 \, w_0)/(1 - w_0)$. The above formula for Z makes the prospective rates proportional to the ratio of the actual second period experience, y_0/y_1.

The given credibility is optimal for the postulated pricing policy. It would be more accurate, however, to charge a higher premium for every additional claim in the experience period. The above pricing policy produces a single rate for all insureds with one or more claims. This rate will be relatively too high for the one-claim doctors and relatively too low for the more-than-one-claim doctors.

This can be demonstrated from another perspective. When there are only discounts, and no surcharges, the loss of the claim-free discount is essentially the surcharge for one or more claims. Recalling the general modification factor formula, and assuming that the average experience period frequency for the given class is λ, the appropriate amount to surcharge for each claim is:

$$\text{Surcharge} = \frac{Z}{\lambda}.$$

Given the basic credibility formula, with $Z = \beta\lambda/(1 + \beta\lambda)$, the surcharge becomes $\beta/(1 + \beta\lambda)$. If λ is relatively small, the surcharge will be approximately equal to β.

Calibrating the Expected Costs

Once we have defined the actual claim experience, A, we determine E, the expected claim experience, as the corresponding class average experience. If E is not calibrated to the class average, most likely we will generate an off-balance. (There also may be an off-balance due to other factors.) We briefly discuss some issues with respect to reported counts and CP5 counts.

First assume that A is defined as the reported count, for claims-made coverage, and that the insurer offers a certain fixed discount for each claim-free year. If claim frequency has changed over time, the optimal discount may be different for each year of experience. We may want to select an average frequency for the maximum number of years that credits are offered. We also may want to add an adjustment for the step of the insured policy, if we use the experience on non-mature years.

We may not have class frequencies or we may want to use our rate relativities. In this case, we should remove that part of the relativity that reflects differences in severities by class. We should also reflect other rating variables in the discounts. For example, if we give teaching doctors a 25% discount, logically their claim frequency should be about 75% of the class average and their credits should be 75% of regular doctors. The same adjustment would apply for territorial rate relativities.

We also may want to apply claim-free discounts to occurrence coverage. In this case, we should adjust for the reporting pattern of claims. Assume, for example, that 10% of claims are reported in the first year, 40% in the second year, 20% in the third year, and 10% in the fourth and fifth years. Thus the cumulative percentage of claims reported would be 10%, 50%, 70%, 80%, and 90%. We also assume that the average doctor in this class has an annual occurrence claim frequency, $\lambda = 0.20$, that has remained relatively constant for the past five years. The average doctor would have a reported claim frequency of 0.18 for the fifth prior year, 0.16 for the fourth prior year, and 0.14, 0.10, and 0.02, respectively. For the five-year experience period, the expected frequency is 0.60. If $\beta = 0.5$ and we use the basic credibility formulation, $Z = 23.1\%$ for the five years of experience. If we round off and simplify, we could give a 5% discount for each claim-free year. We should note, however, that after the first year the expected claim frequency is only 0.01 and the appropriate claim-free discount is only 1%. (The appropriate discounts for each successive year of claim-free experience would be 4.7%, 5.8%, 5.9%, and 5.7%.)

If we define the actual claim experience, A, in terms of non-nuisance claims, such as CP5, there is an additional problem in trying to match claim experience to exposure. Even on claims-made forms, the average claim may take three years or so to be settled. On occurrence forms, the average claim may take six years to be settled. One solution is to define A as being any CP5 claim closed within the last five years, regardless of policy period or occurrence date. This approach would be biased in favor of newer doctors, who would not have had as much chance to have had closed claims.

Non-Optimal Credibilities

For various reasons, we may design a plan that has non-optimal credibilities. For example, we may have the same discount per year for every class, even though we know that classes with higher frequencies should receive larger discounts (if their β's are the same). We may also use a discount only program.

With non-optimal credibilities, most likely there will be an off-balance. An off-balance can also arise if the book of business changes over time. (For example, those insureds that would have received stiff surcharges may move to a residual market program or another insurer.) A negative off-balance causes the class rate to be higher than the average class cost. This may cause problems in ratemaking and in analyzing claim experience. If off-balances are different by class, the ratemaking procedure for class relativities should adjust for these off-balances. Profitability analysis should focus on collectible premiums, rather than manual premiums.

Non-optimal credibilities imply an inaccuracy in pricing. This may place the insurer at a competitive disadvantage compared to an insurer that has more accurate pricing. An example may help to clarify this point.

Assume that the optimal credibility for claim-free insureds is 10%, that the insurer gives a 25% discount and no surcharges, that claim-free insureds constitute 80% of the class, that insureds with one claim constitute the other 20% of the class, and that all insureds have the same experience period. The insurer's off-balance would be 20% (80% of insureds receive a 25% discount), implying a manual rate of 125% $(1/(1-0.2))$ of the average cost. The claim-free insureds would pay 93.75% (0.75×1.25) of the average cost and the non-claim-free insureds would pay 125%.

The most accurate cost estimate for a claim-free doctor would be 90% of the manual rate. The off-balance would be 8% (80% times 10%) and the manual rate would be 108.7% $(1/(1-0.08))$ of the average cost. The claim-free doctor would pay 97.8% of the average cost $(0.9 \times 108.7\%)$ and others would pay 108.7%. The optimal competitor could insure all the one-claim doctors at a profit, while the given insurer would be left with all of the claim-free doctors, at a loss.

As a general rule, if claim-free discounts are higher than the optimal credibility, claim-free doctors will be under-priced and the non-claim-free insureds will be over-priced. The insurer will be vulnerable to price competition for the non-claim-free doctors. Another way of looking at this is as follows. When a doctor has a claim, he or she loses his or her claim-free discount and his or her premium increases. The additional premium is more than the insurer needs to profitably insure that doctor.

Group Formulations

Finally, we consider merit rating formulas for groups of doctors. To a large extent, the practical problems discussed above will also apply to groups. Given that the claim frequency may be much larger for groups, we may prefer a plan that looks more like the NCCI or ISO plans. We discuss the components of the merit rating formula, A, E, and Z, in turn.

The choices for the actual claim experience, A, include all of the possible choices for individual doctors plus several more. Since groups are likely to have several experience period claims, the claim-free discount approach may not be practical. Most likely we will use a fixed experience period of three, five, or more years. The credibility we can assign to the group's experience will increase for each additional year of experience. The amount of the increase will depend upon several factors, such as: Whether there is risk-shifting among individual insureds over time, whether the composition of the group changes over time, and the extent to which there is heterogeneity within the group.

If we use claim counts for A, we may want to define them in terms of occurrences. That is, more than one member of a group may be sued for a given incident; the statistical validity of this multiple-claim single incident is probably not much different than that for a single-claim single incident.

We may want to consider using loss amounts. The reduction in credibility that we saw above, for the variability in the claim size distribution, should be more than offset by the increased number of doctors within the average group. If we use loss amounts, we might want to consider a limit on the amount of a chargeable claim, as is done in the ISO plans. The limit could be determined so that the increase in the modification factor for a maximum claim might be a given percentage (e.g., 25%). Logically, this would reduce the credibility that could be given for the group's experience, since α would be lower for lower claim limits. An adjustment also would need to be made to the expected losses, E. Both of these adjustments could be determined from claim size distribution data.

The calibration of E depends upon the definition of A. If we use reported counts for occurrence policies for a five-year experience period, for example, we would need to adjust for the reporting pattern. The expected frequency might be calculated as the annual occurrence frequency times the number of years in the experience period times an adjustment for the reporting pattern (e.g., 60% in the above example). If A is defined in terms of loss amounts, we need to consider loss development and IBNR.

The determination of Z is more difficult, unless we have two-period claim experience for large numbers of groups of varying sizes. There are several approaches that can be taken. First, we could use the same K that we used for individual doctors. Most likely, this is not appropriate because all of the doctors within the group will not have the same relative cost. This approach would overstate credibilities, because the heterogeneity among groups is less than the heterogeneity among individuals. (Mathematically, the τ^2 for groups is lower than the τ^2 for individuals.)

Second, we could use the basic credibility formulation (e.g., Equation 3.4) and estimate the β from group experience. Since the groups (j) for which we have data most likely will have different claim frequencies (λ_j), we must use a generalized formula for Z, such as $Z_j = (\beta\lambda_j)/(1 + \beta\lambda_j)$. This approach has a few problems. If there is risk-shifting among individuals or a change in the group's composition over time, the appropriate credibility formula would have an additional term in the denominator, e.g., $\delta\lambda_j$. Thus our estimate for

β may not be entirely accurate. In addition, to the extent there is risk-shifting, the credibilities for very large groups should be less than those given by the basic credibility formulation. If we do not insure very many large groups and if there is reasonable homogeneity among the group, this approach may be a reasonable approximation to optimality.

Third, we could build in risk-shifting and insured heterogeneity. In order to measure the appropriate parameters, however, we would need additional data. This could be additional years of data for the same groups or a segmentation of group data by size. If we do not have the necessary data, we may make some educated guesses about the values of δ and γ.

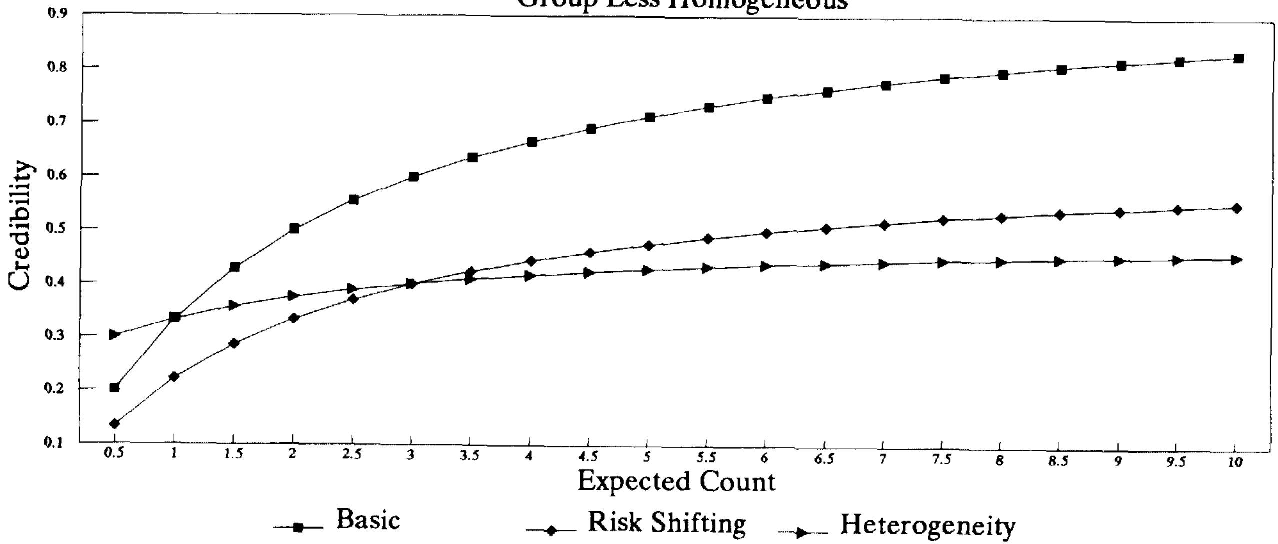
We can compare the results we get with the three different credibility formulations, Formulas 3.5, 3.7, and 3.9. We assumed that the excess variance was 0.5. For the first and third formulations, $\beta = 0.5$. For the second formulation, $\beta + \delta = 0.5$. We think there is a conceptual similarity between the δ parameter in the risk-shifting formulation and the γ parameter in the insured heterogeneity formulation. We think of risk-shifting as how different subsequent years of exposure are to each other. We think of insured heterogeneity as how different sub-exposures within the same experience are to each other.

We have prepared two graphs, Figures 3 and 4. Figure 3 shows the case where $\delta = 0.1$, which is relatively small compared to β. This would occur for groups that are relatively homogeneous. Figure 4 shows the case where $\delta = 0.167$, where the group is less homogeneous. We see that the credibility is always lower for the risk-shifting formulation. For less homogeneous groups, the credibility will be lower. We also see that the risk heterogeneity formulation generally produces lower, though similar, credibility to the risk-shifting formulation. For very low expected counts, the risk heterogeneity formulation may produce higher credibility than the simple formulation. Exhibit 7 gives the numerical credibilities for these two cases.

FIGURE 3

GROUP CREDIBILITIES FOR VARIOUS FORMULATIONS

Group Less Homogeneous

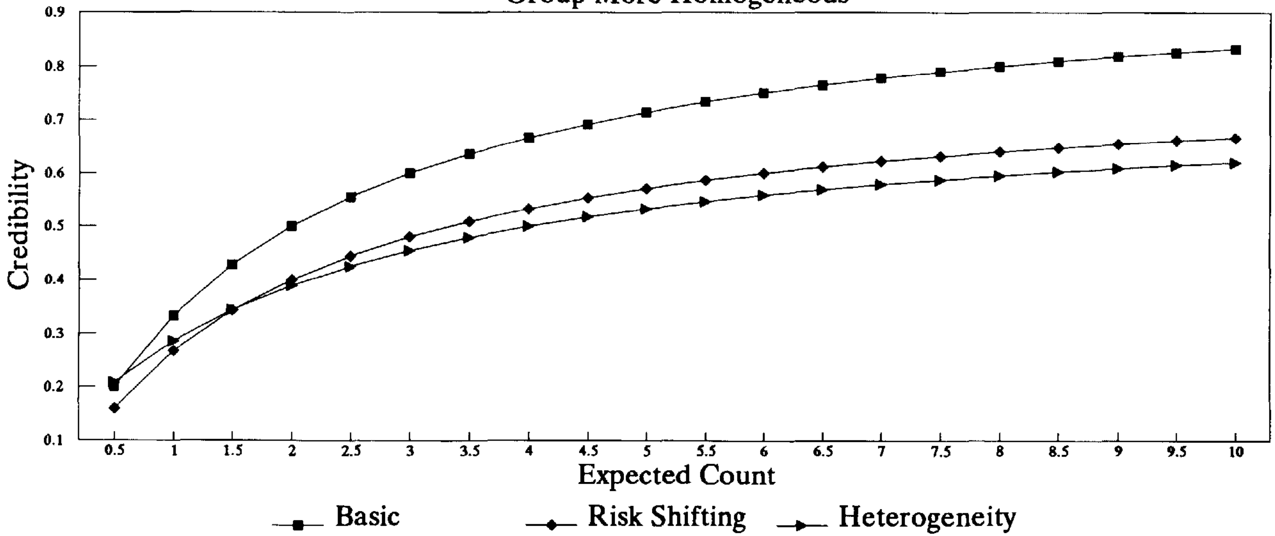


Beta = .4, Delta = .1, Gamma = .125.

FIGURE 4

GROUP CREDIBILITIES FOR VARIOUS FORMULATIONS

Group More Homogeneous



Beta = .333, Delta = .167, Gamma = .25.

6. SUMMARY

Merit rating is the use of the insured's actual claim experience to predict future losses. Merit rating modifies the otherwise applicable class rate. The modification depends on two factors: (1) how much better or worse the insured's experience is relative to the class average, and (2) how credible (i.e., statistically significant) the insured's experience is.

Merit rating formulas can differ in what claim experience is used. Variations include counts or amounts and different lengths of insured experience. There are several generic theoretical formulations for credibility that have been used in insurance pricing. Given sufficient actual data, the appropriate credibility can be estimated.

Merit rating is a complement to the rating plan. It will pick up statistically valid information that is not already reflected in other rating variables. The remainder of the rating structure must be considered in calibrating and applying the merit rating plan.

If the merit rating system creates a collectible premium "off-balance," class rates must be adjusted. If merit rating produces non-optimal discounts or surcharges, there will be inaccurate pricing. If claim-free discounts are too high, for example, those receiving the discounts will be relatively under-priced and those not receiving the discounts will be relatively over-priced.

The statistical validity of an insured's claim experience can be quantified by "credibility" and used in a merit rating formula. Many formulations for credibility are available. Under virtually all formulations, credibility will increase with: (1) The increasing expected claim frequency of the insured's actual experience (λ_i), and (2) the heterogeneity of the insured population, or structure variance, β, remaining after the application of all of the other rating variables. Credibility will decrease with: (1) Increasing variability in the claim size distribution, α; (2) changes in the insured's mean costs over time, or risk-shifting, δ; and (3) heterogeneity within the insured (e.g., with group practices), γ.

In practice, it is relatively easy to determine the expected claim frequency and the variability in the claim size distribution. The structure variance can be determined from single-period data (i.e., from the excess variance), but this requires the assumption that risk-shifting and within-insured heterogeneity are not significant. It is better to estimate the structure variance from two-period data. That is, we must know the relative costs of insureds, within the same rating class, in two different time periods. We would expect the structure variance to be different for different insurers (because they have different underwriting standards), for different states, and for different classes.

Risk-shifting and within-insured heterogeneity are important with respect to the merit rating of group practices. Since all doctors within the group will not be equally good or equally bad, credibility may not increase with additional exposure as it would for an individual doctor. For example, the credibility for one doctor's five-year experience is probably higher than the credibility of five different doctors' combined one-year experience. To measure these factors we need two-period or multi-period data for insured groups of several different sizes.

There are several practical conclusions that can be based on the general theoretical developments and the actual data presented above. Using claim count data will generate more credibility and, hence, larger discounts or surcharges than claim amounts. Using reported count data will generate more credibility than closed-paid count data, but this may cause consumer relations and other problems. Claim-free discounts seem to be a reasonable merit rating plan for individual doctors, subject to two limitations. The amount of the discount should vary with the class expected claim frequency and, generally, the amount should decline for each successive claim-free year.

REFERENCES

[1] Beard, R. E., Pentikänen, T., and Pesonen, E., *Risk Theory* (Third Edition), Chapman & Hall, New York, 1984.

[2] Ellis, R. P., Gallup, C. L., and McGuire, T. G., "Should Medical Professional Liability Insurance Be Experience Rated?" *Journal of Risk and Insurance,* Vol. 57, 1990, p. 66.

[3] Finger, Robert J., "Risk Classification," *Foundations of Casualty Actuarial Science* (First Edition), Casualty Actuarial Society, New York, 1990, Chapter 5, p. 231.

[4] Gillam, William R., "Parametrizing the Workers' Compensation Experience Rating Plan," *PCAS* LXXIX, 1992, p. 21.

[5] Longley-Cook, L. H., *An Introduction to Credibility Theory*, Casualty Actuarial Society, New York, 1962.

[6] Mahler, Howard C., Discussion of Meyers: "An Analysis of Experience Rating," *PCAS* LXXIV, 1987, p. 119.

[7] Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *PCAS* LXXVII, 1990, p. 225.

[8] Mahler, Howard C., "An Actuarial Analysis of the NCCI Revised Experience Rating Plan," *Casualty Actuarial Society Forum*, Winter 1991, p. 37.

[9] Meyers, Glenn G., and Schenker, Nathaniel, "Parameter Uncertainty in the Collective Risk Model," *PCAS* LXX, 1983, p. 111.

[10] Meyers, Glenn G., "An Analysis of Experience Rating," *PCAS* LXXII, 1985, p. 278.

[11] Tiller, Margaret W., "Individual Risk Rating," *Foundations of Casualty Actuarial Science* (First Edition), Casualty Actuarial Society, New York, 1990, Chapter 3, p. 91.

[12] Venezian, Emilio C., Nye, B. F., and Hofflander, A. E., "The Distribution of Claims for Professional Malpractice: Some Statistical and Public Policy Aspects," *Journal of Risk and Insurance*, Vol. 56, 1989, p. 686.

[13]Venezian, Emilio C., "The Distribution of Automobile Accidents—Are Relativities Stable Over Time?" *PCAS* LXXVII, 1990, p. 309.

[14]Venter, Gary G., "Credibility," *Foundations of Casualty Actuarial Science* (First Edition), Casualty Actuarial Society, New York, 1990, Chapter 7, p. 375.

[15]Woll, Richard G., "A Study of Risk Assessment," *PCAS* LXVI, 1979, p. 84.

EXHIBIT 1

REQUIRED MANUAL RATE INCREASES FOR GIVEN CLAIM-FREE DISCOUNTS

Percentage Claim-free	Discount (Z)				
(P_o)	10%	20%	30%	40%	50%
10%	1.0%	2.0%	3.1%	4.2%	5.3%
20	2.0	4.2	6.4	8.7	11.1
30	3.1	6.4	9.9	13.6	17.6
40	4.2	8.7	13.6	19.0	25.0
50	5.3	11.1	17.6	25.0	33.3
60	6.4	13.6	22.0	31.6	42.9
70	7.5	16.3	26.6	38.9	53.8
80	8.7	19.0	31.6	47.1	66.7
90	9.9	22.0	37.0	56.3	81.8

EXHIBIT 2
Part 1

PARAMETER ESTIMATION EXAMPLE

I. RAW DATA AND BASIC CALCULATIONS

Prior Period							Subsequent Period			
Count	Doctors	Percentage of Doctors	Claims	Extension	Relative Frequency	Relative Variance	Claims	Frequency	Relative Frequency	Extension
N	P	w	NP	wNN	x	wxx	q	q/P	y	wxy
0	91	59.5%	0	0.000	0.000	0.000	13	0.143	0.754	0.000
1	36	23.5	36	0.235	1.515	0.540	8	0.222	1.172	0.418
2	17	11.1	34	0.444	3.030	1.020	6	0.353	1.862	0.627
3	6	3.9	18	0.353	4.545	0.810	1	0.167	0.879	0.157
4	2	1.3	8	0.209	6.059	0.480	0	0.000	0.000	0.000
5	1	0.7	5	0.163	7.574	0.375	1	1.000	5.276	0.261
Total	153	100.0%	101	1.405		3.225	29			1.463
	(a)		(b)	(c)	$N/(1)$	(d)	(e)		$q/[P(2)]$	(f)
Frequency			0.660				0.190			
			$(1) = (b)/(a)$				$(2) = (e)/(a)$			

EXHIBIT 2
Part 2

PARAMETER ESTIMATION EXAMPLE

II. EXCESS VARIANCE METHOD

		Nominal			Relative to Mean	
		Source	Value		Source	Value
(3)	Frequency	(1)	0.660	(7)	By Definition	1.000
(4)	Total Variance	$(c)-(1)(1)$	0.969	(8)	$(d)-1$	2.225
(5)	Poisson Variance	= (3)	0.660	(9)	1/(1)	1.515
(6)	Excess Variance	(4) − (5)	0.309	(10)	(8) − (9)	**0.710** = $\hat{\beta}_{CC}$

III. REGRESSION METHOD

Credibility, Z	$[(f)-1]/(8)$	0.208
$\hat{\beta}_{CC}$	$(f)-1$	**0.463**

IV. CLAIM-FREE DISCOUNT METHOD

(11) Claim-Free Discount, Z	$1-y_0$	0.246
$\hat{\beta}_{CC}$	(11) (8)	**0.548**

V. OTHER METHOD—EQUATION (4.3)

$\hat{\beta}_{CC}$	$(y_1-y_0)/y_0$	**0.556**

EXHIBIT 3

EXCESS VARIANCE METHOD

Class	No. of Doctors (1)	No. of Reported Claims (2)	Frequency (3) = (2)/(1)	Total Relative Variance (4)	Poisson Relative Variance (5) = 1/(3)	Excess Relative Variance (6) = (4) − (5)
0	98	64	0.653	2.206	1.531	0.675
1	725	674	0.930	1.429	1.076	0.353
2	208	187	0.899	1.837	1.112	0.725
3	297	413	1.391	1.352	0.719	0.633
4	198	236	1.192	1.161	0.839	0.322
5	170	386	2.271	0.903	0.440	0.463
6	153	485	3.170	0.530	0.315	0.215
7	41	145	3.537	0.605	0.283	0.322
8	28	85	3.036	0.670	0.329	0.341

EXHIBIT 4

CLAIM-FREE DISCOUNT METHOD

Class	No. of Doctors (1)	No. Claim-Free (2)	Class CP5 Frequency (3)	Claim-Free Discount CP5 Count (4)	Total Relative Variance (5)	$\hat{\beta}_{CC}$ (6) = (4) (5)	Claim-Free Discount Pure Premium (7)	$\hat{\beta}_{CC}$ (8) = (7) (5)
0	98	88	0.102	-11.4%	8.800	-1.003	-11.0%	-0.968
1	725	624	0.154	3.7	6.860	0.254	3.5	0.240
2	208	172	0.183	12.1	5.050	0.611	3.7	0.187
3	297	233	0.285	4.1	5.971	0.245	44.4	2.651
4	198	155	0.261	-1.4	4.004	-0.056	-2.1	-0.084
5	170	105	0.547	30.6	2.322	0.711	31.3	0.727
6	153	91	0.660	24.6	2.225	0.547	16.1	0.358
7	41	22	0.829	33.4	1.696	0.566	20.6	0.349
8	28	17	0.464	58.8	1.817	1.068	52.0	0.945
Total	1,918	1,507						

EXHIBIT 5

IS THERE RISK-SHIFTING?

Class	Excess Relative Variance	Regression Estimate for β	Difference	Percentage Difference
1	0.353	0.318	0.035	9.9%
2	0.725	0.570	0.155	21.4
3	0.633	0.868	-0.235	-37.1
4	0.322	0.371	-0.049	-15.2
5	0.463	0.370	0.093	20.1
6	0.215	0.228	-0.013	-6.0
Sum	2.711	2.725	-0.014	-0.5%

Note: Based on reported counts.

EXHIBIT 6

OTHER DOCTOR EXPERIENCE

I. ELLIS, GALLUP, AND MCGUIRE

Specialty	Five-Year Claim-Free Discount	Excess Relative Variance	Five-Year Mean Frequency
Anesthesiology	3.4%	0.20	16.3%
Dermatology	28.4	4.04	9.2
Family Practice	17.6	2.88	7.1
General Surgery	20.2	0.90	35.2
Internal Medicine	24.1	3.87	8.3
Neurosurgery	30.5	1.07	42.8
Obstetrics/Gynecology	29.4	1.08	39.9
Ophthalmology	37.0	3.46	15.2
Orthopedic Surgery	52.6	4.22	26.0
Otolaryngology	38.2	2.64	24.5
Pediatrics	23.6	4.65	7.0
Plastic Surgery	59.6	6.78	34.2
Psychiatry	24.2	22.89	1.7
Radiology	21.0	2.92	9.1
Urology	19.2	1.22	15.9
All Other	10.0	5.22	2.5

II. VENEZIAN, NYE, AND HOFFLANDER

Specialty	Mean Frequency	Excess Relative Variance
Anesthesiology	7.5%	0.46
General Surgery	14.4	1.10
Internal Medicine	3.6	0.19
Neurosurgery	50.0	0.72
Obstetrics/Gynecology	18.7	0.62
Ophthalmic Surgery	3.0	5.34
Orthopedic Surgery	25.7	1.37

EXHIBIT 7
Part 1

Comparison of Different Group Credibility Formulas
Group More Homogeneous

$\beta = 0.400$
$\delta = 0.100$
$\gamma = 0.125$

Expected Count	Basic (1)	Risk-Shifting (2)	Heterogeneity (3)
0.5	20.0%	16.0%	20.8%
1.0	33.3	26.7	28.6
1.5	42.9	34.3	34.4
2.0	50.0	40.0	38.9
2.5	55.6	44.4	42.5
3.0	60.0	48.0	45.5
3.5	63.6	50.9	47.9
4.0	66.7	53.3	50.0
4.5	69.2	55.4	51.8
5.0	71.4	57.1	53.3
5.5	73.3	58.7	54.7
6.0	75.0	60.0	55.9
6.5	76.5	61.2	56.9
7.0	77.8	62.2	57.9
7.5	78.9	63.2	58.8
8.0	80.0	64.0	59.5
8.5	81.0	64.8	60.2
9.0	81.8	65.5	60.9
9.5	82.6	66.1	61.5
10.0	83.3	66.7	62.0

Notes: (1) $Z = E/(E + 2)$. Equation (3.5).
(2) $Z = E/(1.25E + 2.5)$. Equation (3.7)
(3) $Z = (0.75E + 0.25)/(E + 2)$. Equation (3.9).

EXHIBIT 7
Part 2

COMPARISON OF DIFFERENT GROUP CREDIBILITY FORMULAS
GROUP LESS HOMOGENEOUS

$\beta = 0.333$
$\delta = 0.167$
$\gamma = 0.250$

Expected Count	Basic (1)	Risk-Shifting (2)	Heterogeneity (3)
0.5	20.0%	13.3%	30.0%
1.0	33.3	22.2	33.3
1.5	42.9	28.5	35.7
2.0	50.0	33.3	37.5
2.5	55.6	37.0	38.9
3.0	60.0	40.0	40.0
3.5	63.6	42.4	40.9
4.0	66.7	44.4	41.7
4.5	69.2	46.1	42.3
5.0	71.4	47.6	42.9
5.5	73.3	48.8	43.3
6.0	75.0	50.0	43.8
6.5	76.5	50.9	44.1
7.0	77.8	51.8	44.4
7.5	78.9	52.6	44.7
8.0	80.0	53.3	45.0
8.5	81.0	53.9	45.2
9.0	81.8	54.5	45.5
9.5	82.6	55.0	45.7
10.0	83.3	55.5	45.8

Notes: (1) $Z = E/(E + 2)$. Equation (3.5).
(2) $Z = E/(1.5E + 3)$. Equation (3.7)
(3) $Z = (0.5E + 0.5)/(E + 2)$. Equation (3.9).