Experience Rating—Equity and Predictive Accuracy

by GARY VENTER, FCAS

Gary Venter, FCAS, is Vice President and Actuary at the National Council on Compensation Insurance. Mr. Venter, who heads the NCCI's Actuarial Research Division, is a member of the American Academy of Actuaries, the American Risk and Insurance Association, the International Actuarial Association, and ASTIN.

Workers compensation experience rating relates an insured's premium to its own loss experience, to the extent that this experience is a believable indicator of the insured's exposure to loss. For each experience-rated risk, the losses for the most recently available three-year period are compared to the losses that would be expected to arise in a business of its size, according to the manual classification rates, and a credit or debit modification is calculated. This credit or debit is applied to the manual premium for the policy period. The result is the "standard premium" for the insured.

The weight given to the insured's experience depends on the credibility of that experience. Credibility is a function of the size of the insured: the larger it is, the more weight will be given to its own experience; the smaller it is, the more reliance is placed on the manual rate. More mathematical interpretations of this intuitive relationship will be explored below.

FUNDAMENTAL QUESTIONS

Before the mathematics can be applied, however, a logical foundation must be established. What are the goals of the experience-rating process? How much of a charge for individual experience is equitable? At what point is the insured being asked to carry risk that is totally unpredictable? Even if complete answers to these issues are not available, addressing them will at least provide some perspective to the formulae used in experience rating.

Two of the primary goals of experience rating are safety incentive and predictive accuracy. By charging insureds something for accident occurrence, a financial incentive for safety is added to the other incentives that may exist. To the extent that the experience modification predicts the future loss experience of the insured, it allows the premium to be tailored more closely to each insured's own loss potential. That is the

equity contribution of experience rating.

The incentive value of experience rating cannot be used to define equity. It would seem that the more incentive for safety, the better; and thus for the incentive value, the more weight given to the insureds' own experience, the better. At some point, however, it would be unfair to insureds to charge back losses that were random and essentially unpredictable.

"By charging insureds something for accident occurrence, a financial incentive for safety is added . . ."

In terms of loss prediction, however, equity is easier to define. To the extent that the loss experience is indicative of true differences from the classification average, it appears equitable to charge for it; to the extent that the loss experience represents only a random fluctuation, charging for it would seem to require the insured to bear a risk that should really be spread among all the insureds.

To determine how much the experience of an insured indicates its loss potential, comparisons are made between the modification and the future experience it was designed to predict. Insureds with modifications—say, in the range 1.10 to 1.20—should produce, on the average, ratios of losses to manual premium 10 to 20 percent higher than the classification as a whole. If so, then on a standard premium basis, i.e., after the application of the modifications, these risks would have the same loss ratio as the classification.

From this perspective, the degree to which an insured should be charged for past loss experience is the degree to which that experience is predictive of future loss experience. This allocation charges each insured according to its identifiable loss potential, thereby differentiating among insureds.

From the viewpoint of the insurer, after experience rating, all insureds have the same expected profit potential, regardless of their past loss history. From the viewpoint of a detached economist, the market will charge each insured according to its identified loss potential, and this, in turn, produces the optimal and most equitable allocation of resources. In effect, equity is achieved by maximizing predictive accuracy. Thus, experience rating is an important step in satisfying the statutory requirement that rates not be unfairly discriminatory.

The problem now has enough specificity to impose a mathematical treatment. The question becomes: how much can an insured's past history be relied upon to predict future losses? Different experience-rating formulae could overstate or understate the proper degree of credibility and thus make the plan overly or insufficiently sensitive to fluctuations in loss experience.

How past experience can best be reflected in future cost estimates has long been pondered by casualty actuaries. Recent advances in statistical methods have allowed a larger group of academic researchers from related disciplines to contribute their insights. A recent study at NCCI reviewed the functioning of the experience-rating plan in light of this research and found that the plan worked well in general but that some fine tuning of parameter values was appropriate.

> **"... the degree to which an insured should be charged for past loss experience is the degree to which that experience is predictive of future loss experience."**

STATISTICAL APPROACH

To apply the statistical methods, it is first necessary to more precisely define the concept of predictive accuracy. Predictive accuracy measures the difference between the predicted losses and the losses that actually turn up. There are infinitely many methods for defining what is most accurate; the two most intuitively appealing are the expected squared error and the expected absolute error. The expected squared error is universally preferred because it leads to mathematically solvable equations, and also because it mitigates against large individual errors. Other better criteria may one day be found to be more appropriate, but a good deal of work would be necessary to apply them.

Given several years of observations of an insured's loss ratio, the expected squared error is minimized by a function statisticians refer to as the conditional expected value, which is simply the expected value of the next year's loss ratio given the observations. So far this function has not been widely used, however, because it requires additional statistical assumptions, is sometimes cumbersome to compute, and is not always readily understood.

These obstacles may be overcome in the not too distant future. However, at this time, the most commonly used approach to experience rating requires the experience modification to be a linear function of the losses. A linear function of x is one that can be expressed as a multiple of x plus a constant term, i.e., $f(x)=ax+b$. Subject to the linearity constraint, the formula that minimizes the expected squared error is sought.

Under some popular but overly simplified statistical assumptions, similar to the law of large numbers, the formula that does this is:

Modification = (A+K)/(E+K) **(Formula I)**

where A represents the total of the actual losses over the experience period, E is the expected losses from the classification rates, and K is the credibility constant.

Formula 1 is a linear function of the losses, A, and, no matter what the distribution of A is, minimizes the expected squared error, if K is chosen correctly. This formula has been used in workers compensation and other lines of insurance, but since about

> "... after experience rating, all insureds have the same expected profit potential ..."

1940 a different formula has been used, at least for workers compensation. This other formula is the multi-split formula, which divides both the actual and expected losses into primary and excess components. In the multi-split formula, each loss has

> "... equity is achieved by maximizing predictive accuracy."

a primary component, which is intended to be a relatively small dollar amount, reflecting loss frequency. Individual large claims will also have an excess component, which is the amount above the primary component. The excess component itself has a maximum cap, which limits the experience rating charge due to individual large claims. The expected losses underlying the manual rate are also divided into primary and excess components. Then the experience modification can be calculated by minimizing the expected squared error of the primary and excess losses separately.

The shift some 50 years ago to a multi-split plan was based on difficulties with the existing formulae. From a current statistical perspective, a primary-excess split can be understood in light of the constraint that the estimate be linear in the observations. The difference between the unconstrained optimal estimate (the conditional expected value) and the best linear estimator is small or zero for some distributions; but for highly skewed, heavy-tailed distributions it can be substantial. This is just the kind of distribution found for workers compensation loss ratios.

The typical response of a statistician faced with such a distribution is to transform it through some normalizing function, such as taking logarithms. Computing the primary value of the losses is analogous to this procedure. The excess losses may be thought to have much less predictive value than the primary losses, and formulae based only on primary losses have been suggested. The recent NCCI study suggests that, in fact, primary losses should have much more credibility, but that excess losses do have predictive value.

In fact, both the primary and excess

losses are less heavy-tailed than the total losses; this seems obvious for primary losses. For excess losses, by eliminating the first smaller portion, enough losses are eliminated to bring up the average value and to reduce the probability of a loss being a large multiple of the average value. This

> "... both the primary and excess losses are less heavy-tailed than the total losses ..."

makes the excess losses less heavy-tailed and thus more predictable than the total losses.

The experience-rating plan defines any loss below $2,000 as primary. For losses above $2,000 the primary value is determined by multiplying the total loss X by the factor:

Primary factor = 10,000/(X+8000)

(Formula 2)

This factor starts at 1 for X=2000 and gradually falls to 0 for large X. The primary value of a loss is always greater for larger losses, however, except for rounding. The maximum possible primary value is $10,000.

To see what the experience-rating formula is for the multi-split plan, it is instructive to review the foundation of the no-split modification formula. Formula 1 arises from a credibility weighting of the actual losses A with the expected losses E, where the credibility weight is denoted as Z:

Modified expected losses = ZA + (1-Z)E

(Formula 3)

If Z is calculated as Z=E/(E+K), then Formula 3 implies that the modified expected losses can be written as the expected losses E times the modification from Formula 1.

For a split plan, Formula 3 gets more complicated. Using subscripts p and e for primary and excess losses, the modified expected losses are:

$$\text{M.e.l.} = Z_pA_p + (1-Z_p)E_p + Z_eA_e + (1-Z_e)E_e$$

(Formula 4)

This is just like Formula 3, except with a primary piece and an excess piece. The credibilities Z_p and Z_e can also be defined similarly, but now two credibility constants, B and K, are needed. Defining $Z_e=E/(E+K)$, $Z_p=E/(E+B)$, and introducing the convenient notation $W=(E+B)/(E+K)$ will yield, after some algebra, that the modified expected losses are E times the split plan modification below:

$$\text{Modification} = [A_p + WA_e + (1-W)E_e + B]/(E + B)$$

(Formula 5)

This is the modification formula currently used at NCCI, with the exception that B and K are not constants but vary by size of risk according to published tables. The theory that B and K should be constants assumes that larger insureds are more stable than small insureds, and that this increased stability follows a particular mathematical form related to the law of large numbers. In particular, it assumes that the variance of the Formula 5 loss ratio is inversely proportional to the firm's premium. In

practice, the relationship between large and small insureds is more complicated than this; the variance does not decrease that fast for larger insureds. The result is that B and K should vary by size of insured and, in fact, should increase for larger firms.

> **"[An] NCCI study suggests that . . . primary losses should have much more credibility, but that excess losses do have predictive values."**

Why large and small insureds display the patterns they do is not fully understood. Some recent advances in the mathematical theory of risk have produced models that roughly fit the observed patterns, but even these do not fully predict the data. According to these models, larger insureds will display more variation than the law of large numbers would predict, basically because all insureds are subject to variations due to changing conditions, and this does not decrease by size of firm. The result of this is that B and K should increase in size for larger firms, thereby decreasing the credibility given to those firms' loss experience compared to what the more simplistic models would predict.

While the above considerations lead to tables of B and W that vary by size of firm, the historical tables were devised long before such arguments had arisen. The B and W values were designed with two goals in mind: to prevent too large a swing for small insureds, and to provide self-rating (100% credibility) for the largest firms. The result was that B and K decreased for larger insureds, where the current theory and data suggest they should in fact increase.

In terms of the weight given to individual experience, as opposed to the manual rates, too much weight was given to large insureds, and too little to small insureds. The result is that small insureds with credit modifications came to be regarded as preferred business, and in fact they have been generally more profitable than those with debit modifications; in effect, the modification did not fully reflect the actual predictive power of the experience. By appropriately second-guessing the modification calculation, the market adjusted for its shortcomings.

For large insureds, on the other hand, too much credit has been given for good experience, leading to less profit potential than for the debit modified firms, because too much credibility was given, in order to achieve self-rating. For many medium sized firms—say, with $10,000 to $100,000 of workers compensation premium each year—the present plan has been found to work close to optimally.

In NCCI's recent review of experience rating values, a proposed modification formula was developed which incorporates increasing values of B and K, according to the current theoretical and empirical findings. The theory of changing conditions mentioned above would lead to values of B and K that increase as a linear function of the insured's expected losses. In fact, the best performance was obtained with values that increased in a non-linear pattern. This is consistent with a risk theoretic model that, in addition to incorporating changing conditions, considers a large firm may be composed of diverse operations with different exposure to loss. Under this model, small firms get relatively more credibility than the linear B and K would provide, and this is in fact what data from studies conducted by NCCI, among others, seems to support.

> "For large insureds . . . too much credit has been given for good experience . . ."

MANUAL LOSS RATIO*		
RISKS WITH	CURRENT PLAN	PROPOSED PLAN
LOWEST MODS	.75	.70
NEXT LOWEST	.79	.78
MIDDLE	.89	.91
NEXT HIGHEST	1.10	1.12
HIGHEST MODS	1.40	1.43

* Indexed to 1.00

Table 1
Identifying Risk Differences
Small Rated Risks ($2500–$5000)

The proposed rating formula is exactly Formula 5 above, i.e., the same as the current formula, with two differences. One difference is that the Table B and Table W values are different. In addition, a simplification to the primary-excess split is used. The primary value of a loss is defined by a single split formula, rather than the primary split Formula 2 above. What this means is that any loss below the split point—in this case $5,000—is primary, and that for any larger loss the primary value is $5,000. As before, the excess value is the total minus the primary, subject to a maximum cap. The reason for the single split was to simplify the operation of the plan. Doing so had no deleterious effect on performance.

Two tests of plan performance were made according to the general reasoning discussed earlier. First the ability of the plan to identify risk differences was tested. Insureds with lower modifications should have better experience in the next period to which the mod applies on a manual premium loss ratio basis, and the higher mods should produce higher loss ratios. Thus the greater the dispersion of the subsequent period manual loss ratios, the better the plan works. On a standard premium basis, however, the loss ratios should be less

dispersed and, ideally, all equal for a better working plan. This is the premium equity argument that asserts that the price for all insureds should reflect their loss exposure.

To perform these tests, insureds were divided into five groups by size of modification: the lowest modifications, the next lowest, etc. up to the highest. This grouping was done twice: once using the current formula and once using the proposed formula. These tests were performed separately for several different premium groups. The insureds in each modification size group are somewhat different for the current and proposed modification formulae because the modifications themelves are somewhat different under each formula. For each of the groups the manual and standard loss ratios were computed for both the current and proposed plans.

> "The real test of experience-rating equity . . . is how well the plan corrects for the differences it identifies"

	STANDARD LOSS RATIO*	
RISKS WITH	CURRENT PLAN	PROPOSED PLAN
LOWEST MODS	.83	.89
NEXT LOWEST	.85	.96
MIDDLE	.94	1.06
NEXT HIGHEST	1.11	1.09
HIGHEST MODS	1.17	.97

* Indexed to 1.00

Table 2
Correcting Risk Differences
Small Rated Risks ($2500–$5000)

The results of the test of the ability of the plan to identify risk differences is shown in Table 1 for the small experience-rated firms ($2,500–$5,000 in annual premium).

As discussed above, the better functioning experience-rating formula will increase the dispersion of these manual loss ratios, as it better identifies the firms likely to have good or poor experience in the next period. The table indicates that the proposed plan works a little better in this regard, at least for the smaller, experience-rated firms.

The real test of experience-rating equity, however, is how well the plan corrects for the differences it identifies. Table 2 illustrates this for the small firms shown in Table 1.

Here, the closer to unity the better, and the proposed plan seems to work better, although the current plan is a significant improvement over no experience rating. For the small firms the proposed plan is more responsive than the current plan, which brings both the credit and debit risks closer to unity after experience rating.

For insureds with premium between $10,000 and $100,000, the current

and proposed plans both work fairly well. Table 3 shows the results for a segment of these firms.

Here again, the lowest and highest modification groups are a little closer to unity after experience rating for the proposed plan.

For the largest firms, the current plan is somewhat overresponsive, as Table 4 reveals.

	STANDARD LOSS RATIO*	
RISKS WITH	CURRENT PLAN	PROPOSED PLAN
LOWEST MODS	.88	.96
NEXT LOWEST	1.05	1.06
MIDDLE	1.03	1.02
NEXT HIGHEST	1.02	1.04
HIGHEST MODS	.99	.95

* Indexed to 1.00

Table 3
Correcting Risk Differences
Medium Sized Risks ($10,000–$25,000)

	STANDARD LOSS RATIO*	
RISKS WITH	CURRENT PLAN	PROPOSED PLAN
LOWEST MODS	1.08	1.00
NEXT LOWEST	1.03	1.01
MIDDLE	1.01	.98
NEXT HIGHEST	.96	1.00
HIGHEST MODS	.99	1.02

* Indexed to 1.00

Table 4
Correcting Risk Differences
Large Risks (Above $100,000)

The current plan shows the worst experience for the credit risks, basically because the credits are somewhat too high. The debits are also too high, resulting in better loss ratios for the two highest modification groups. The proposed plan builds in a little less sensitivity to actual loss experience for these risks, resulting in more appropriate charges for all groups.

In conclusion, it has been argued that while the experience-rating plan serves an important incentive purpose, its incentive value cannot be used to define equity among insureds because, for incentive, more weight to the individual insured's experience is always better. However, the predictive ability of the experience-rating plan can be related to equity: both equity and predictive ability require that just the right balance of responsiveness be maintained, and it has been argued that the same degree of responsiveness is appropriate for each. The current experience-rating values are found to provide a very good balance in this regard; the proposed values work even better, especially for the largest and smallest rated firms and the highest and lowest modifications.