PARAMETRIZING THE WORKERS' COMPENSATION EXPERIENCE RATING PLAN

By William R. Gillam, FCAS
Assistant Vice President & Actuary

National Council on Compensation Insurance

PARAMETRIZING THE WORKERS' COMPENSATION EXPERIENCE RATING PLAN

By: William R. Gillam

AUTHOR

Mr. Gillam is Assistant Vice President and Actuary, Individual Risk Rating, at the National Council on Compensation Insurance. He leads a tightly knit corps of actuaries extracting statistics of impeccable veracity for use in research, publication, and consultation. He previously worked for reinsurance companies - Prudential and North American - and before that for ISO. He is a fellow of the CAS and possesses a B.A. and M.S. in mathematics from Wesleyan and Rutgers Universities respectively.

ABSTRACT

This paper describes the underlying assumptions and statistical evaluations used to develop the Revised Workers' Compensation Experience Rating Plan. This revision is being filed in all jurisdictions in which NCCI administers the Plan. In addition, it has been recommended to the independent jurisdictions.

The foremost characteristic of the new plan is that it was thoroughly and objectively tested to see that its performance was optimal. As described in the body of the paper, mathematical statistics provide justification for the mathematical form of the modification, but the coefficients were evaluated by extensive trial and error. The resulting plan will produce ratings which justify the adjective Standard to the individual risk premium.

OUTLINE
PARAMETRIZING THE WORKERS' COMPENSATION EXPERIENCE RATING PLAN

I. **Theoretical Justification**

A. Motivation
B. Derivation of Mod Formula
C. Variance Assumptions

II. **Estimation of Parameters**

A. Simple initial testing
   1. Grouped means for 5 size groups - c/f Dorweiler's necessary condition
   2. Off-balance by state
B. Quintiles test-c/f Dorweiler's sufficient condition
C. Estimation of plan parameters for hypothetical plans

III. **Plans tested**

A. Current rating values
B. Alternate formula optional rating value
   1. Perryman I and II
   2. Single or Multi-Split of individual losses
   3. State relativities

IV. **Selected Plan**

A. Selection
B. Related Changes to the rating plan - D-ratio computation
C. Later Refinements
   1. Minimum credibility constant
   2. SRP/loss limitation
D. Ad-hoc maximal values on certain mods

SECTION I-THEORETICAL JUSTIFICATION

A. Motivation

Experience Rating is a superb source of data for testing the theory of credibility. The statement that "credibility theory is the best foundation for the practice of experience rating" has been inverted to one which puts the cart before the horse because that is what motivates this initiative. Compensation Experience Rating is a large scale application of a theory and, as such, is an ongoing test of that theory.

This paper describes how researchers at NCCI used such testing to create an improved experience rating plan. The power of modern electronic data processing enabled us to reopen older experience rating files and recalculate mods as if a hypothetical plan had been in place. The plans we tested and how we measured their performance are described in this paper.

Our general strategy was to start with a formula based on sound theory, then use iterative testing to parametrize that formula. We used Least Squares, or Bayesian, credibility to develop an algebraic form for the modification formula. Certain assumptions about loss variance were necessary to this development. It is heartening that the parameters which worked best were consistent with our prejudices about the components of risk loss variance.

This section outlines the theoretical development of the split plan modification formula. It is based on a Bayesian view of the process of individual risk rating and certain assumptions used to simplify the analysis. The view that the individual risk has loss parameters that are knowable only within some ranges described by a structure function (the probability distribution of the parameters) is the risk theoretic basis of the process.

It is impossible to apply risk theoretical models to real world situations without making assumptions. Since we are using actual experience rating data to parametrize theoretical formulas, we hope that weaknesses of some individual assumptions will be offset by the power of the model. Models with the ability to fit a variety of processes not necessarily satisfying the underlying assumptions are called robust.

Credibility is a function of the components of loss variance. If we begin with a general expression for variance of the loss process, then use extensive tests of real data to evaluate parameters, it is reasonable to expect the resulting credibilities will work well in practice. We believe the risk theoretic basis assures that the general framework will not lead to an overspecified system. The framework is filled in by means of an approach similar to that of Bailey and Simon[1], but taken a step further from the calculation of a single credibility value to the optimization of a the functional expression for credibility.

The underlying analysis is simplified by taking most of the administrative features of the current Experience Rating Plan as fixed. This eliminates the need for multiple subscripts. You can refer to papers by Hewitt[2], Meyers[3], Mahler[4] and Venter[5] for more general theoretical background.

The split plan modification formula can be derived with no distributional assumptions and only one major simplifying assumption: that unconditional expected primary and excess losses are uncorrelated. This is a useful simplification, but probably one of the weaker ones as alluded to above.

Here, "split plan " is one in which individual losses are split by formula into two components, primary and excess, and assignment of separate credibilities to the totals of the respective loss components. It is interesting to note that in the course of evaluation of plan parameters, we found that a change in the primary excess split formula improved the performance of the plan. I'd like to think that the change served to put the data used for rating in a form better fitting the assumptions.

B. <u>Derivation of Mod Formulas</u>

Hypothesize a linear approximation to the posterior mean experience P_o + X_o (split primary/excess) given prior experience P_t and X_t

1) $$P_o + X_o = Y + Z_p P_t + Z_x X_t + e$$

where

P = Primary loss
X = Excess loss
Y = Constant to be determined
t = (past) time period
o = (future) time period
e = error

Z_p and Z_x will be called the respective primary and excess credibilities: these and Y are coefficients to be evaluated.

These time periods are fixed in the experience rating plan, so that time period t is the three most recently completed one year periods before the prospective single policy period, labeled o. For example, this will mean the experience of completed policies incepting in 1986, 1987 and 1988 will be used to rate a 1990 policy.

Solving this equation for the coefficients leading to minimum value of e^2 yields the following expressions.

2) $$Z_p = \frac{Var_s[E[P_t|S]]}{Var[P_t]}$$

Also,

$$Z_x = \frac{Var_s[E[X_t|S]]}{Var[X_t]}$$

And

3) $$Y = (1-Z_p)\ E\ [P_t] + (1-Z_x)\ E\ [X_t]$$

Where S denotes the particular risk and the subscript s denotes the a priori structure.

Equation (2) also has been written

$$Z_p = \frac{1}{1 + \frac{E_s[Var[P_t|S]]}{Var_s[E[P_t|S]]}}$$

using $Var[P_t] = E_s[Var[P_t|S]] + Var_s[E[P_t|S]]$.

These equations reduce our original equation (1) for P_o and X_o to a linear credibility estimate of the posterior means.

4) $$P_o + X_o = E[P_t] + E[X_t] + Z_p(P_t-E[P_t]) + Z_x(X_t-E[X_t])$$

In practice, we will take the loss functions as ratios to the a priori expected total loss, hence $E[P_t] + E[X_t] = 1$.

The rate modification factor is:

5) $$M = 1 + Z_p(P_t - E[P_t]) + Z_x(X_t - E[X_t])$$

C. Variance Assumptions

It will take some more assumptions to derive the form of the components of variance in the formulas for Z_p and Z_x.

1. First Level - Individual Risk Variance

In (4), we begin focusing on loss ratio functions P and X. These ratios have a variance which decreases as the size of risk increases. The particular ratios P_t and X_t are the emerged primary and excess actual losses of the individual risk divided by the unconditional expected total losses; the denominator is the exposure. The most simple assumption is that the large risk is essentially a combination of a large number of independent homogeneous units. This leads to a process variance of the loss ratio which is inversely proportional to exposure. We must also assume that as more time periods are added to the exposure it can be thought of as adding more independent units of exposure; process variance again decreases proportionately. Thirdly, it is usually assumed that the variance of the hypothetical means does not vary with exposure. With these assumptions, we can write

$$E_S[Var[P_t|S]] = \frac{a}{T}$$

where T represents the total expected losses, or exposure, of the individual risk and

$$Var_S[E[P_t|S]] = b$$

is constant.

Here b, the variance of ratios less than one, would be small relative to a, which is measured in exposure units.

Using (2),

$$Z_p = \frac{b}{a/T + b}$$

$$= \frac{bT}{a + bT}$$

$$= \frac{T}{T + a/b}$$

This is the familiar expression

6) $$Z_p = \frac{T}{T + K}, \text{ where K is constant.}$$

Similarly,

$$Z_x = \frac{T}{T + L}, \text{ where L is the excess credibility constant.}$$

This "compound fraction" form with T alone in the numerator, and K a ratio of components of variance, helps to simplify the algebra of the final mod.

2. Second Level - Individual Risk Variance

Several investigators have refuted this simple variance assumption, for instance Meyers[3] and Mahler[4]. These writers show that conditional variance does not decrease in inverse proportion to exposure. If we assume there is a small, non-diversifiable component of risk loss ratio variance averaging $c > 0$, we can write

$$E_s[Var[P_t|S]] = c + \frac{d}{T}$$

Using b again as the variance of the hypothetical means,

$$Z_p = \frac{b}{b+c+d/T}$$

$$= \frac{bT}{bT+cT+d}$$

$$= \frac{T}{T + \frac{cT + d}{b}}$$

7) $$Z_p = \frac{T}{T + K'}, \text{ where } K' = \frac{cT + d}{b} \text{ is a linear function of}$$

the exposure. Here, b and c are small and d, like a above, is larger. It can be seen that the limiting value of credibility for the largest risks is less than unity or b/b+c.

This form for K' and a similar one for L' were among possible formulas tested as described elsewhere in the paper. With K a linear function of the exposure, it was no surprise that it performed better than the constant coefficient K, and considerably better than the formula B value of the old plan. It was still not as good as alternate formulas described below. The data showed that K should not be constant, nor even a linear function of T, but rather should be a curve, increasing rapidly at first, but then decreasing in slope to a more nearly linear form for large values of T.

3. Third Level - Structure variance

The variance assumptions resulting in the formula for K at this level, were independently derived by H. Mahler[4] at a particularly propitious point in time for NCCI researchers, who at the time were investigating ad-hoc functional forms for K with the perceived desirable characteristics. (It should be noted that Mahler credits Hewitt[6] with observation of the underlying phenonmenon.)

In the formula used, it was assumed that the variance of risk conditional means was not constant across all risk sizes, but had a component inversely proportional to exposure. This would result if each larger risk is, in part, a random combination of non-homogeneous components. The effect is to flatten the variance of the hypothetical means as risk size increases. The form this takes is:

$$\mathrm{Var}_s[E[P_t|S]] = e + f/T$$

Retaining our second assumption about individual risk variance, we can write:

$$Z_p = \frac{e + f/T}{e + f/T + c + d/T}$$

$$= \frac{eT + f}{(e+c)T + d + f}$$

In what I am calling compound fraction form, this is

$$Z_p = \frac{T}{T + \frac{cT + d}{e + f/T}}$$

8) $$Z_p = \frac{T}{T + K''}, \text{ where } K'' = \frac{cT + d}{e + f/T}$$

A similar form would follow for L". Notice that d and f are relatively large compared to c and e. Since c is a small component of within (loss ratio) variance and e is a large component of between (loss ratio) variance, it is plausible that $c < e$.

For some reason, we thought it desirable to "compound" the fraction still further and put K" in a (quasi) linear form. Using * for either p or x, write

$$K_* = \frac{cT + d}{e + f/T}$$

Dividing through by e, we redefine c/e=C, d/e=D, f/e=F, so that

$$K_* = \frac{CT + D}{1 + F/T}$$

Then proceed as follows:

$$K_* = \frac{CT^2 + DT}{T + F} = \frac{CT^2 + CFT - CFT + DT}{T + F}$$

$$= \frac{CT(T+F) - CFT + DT}{T + F}$$

$$= CT + \frac{(D - CF)T}{T + F}$$

$$= CT + \frac{(D - CF)}{1 + F/T}$$

9) $$K_* = CT + \frac{H}{1 + F/T}, \quad \text{where } H = D - CF$$

the form of the equations for K_p and K_x finally used.

In this form, we found C to be a small, positive number, which is reasonable if you can convince yourself that c is smaller than e. H and F, likely to be large positive numbers, indeed are.

In a sense, the final parametrization selected is more general than the underlying variance assumptions. This is because we did not try to estimate the components of variance underlying the assumptions, but rather used direct testing to see what worked best. Thus, the only constraint on plan performance was the algebraic form of (10), not our ability to analyze variance.

With the above definition of $K'' = K_p$ and a similar one for $L'' = K_x$ underlying Z in the form $E/(E+K_*)$, the modification formula becomes,

10) $$M = 1 + \frac{A_p - E_p}{E + K_p} + \frac{A_x - E_x}{E + K_x}$$

where A and E are the actual and expected losses from the experience period and p denotes primary, x excess. This will become our basic proposed formula. For each of K_p and K_x, we have three coefficients to estimate.

SECTION II-ESTIMATION OF PARAMETERS

A. Initial Testing

The idea of evaluating individual risk credibility by a hindsight consideration of how well it worked has been around since Dorweiler[7] who discussed Workers Compensation. Bailey and Simon[1] also gave a simpler example of the procedure for automobile merit rating wherein they were able to estimate implicitly optimal credibilities.

In our study of experience rating, the criteria of "working best" was first measured by ability of the plan to satisfy Dorweiler's necessary criterion for correct credibilities: that credit risks and debit risks would be made equally desirable insureds in the prospective period. This property should obtain across all size categories. I will call this the naive test, which belies its great value to our early investigations.

The early tests begin with risks in the rating year 1980 data files, compute their modifications according to the formula to be tested, then see what 1980 loss experience actually emerged by reference to the 1980 loss data in the 1982 rating year files. The risks in each size group would be separated by their 1980 modifications so that risks with mods in the lower 50th percentile would be in one stratum and risks with mods in the upper 50th percentile would be in the other. Especially for the smaller size groups, many more risks have credit mods, so that the upper percentile contained a proportion of risks with small credits.

To test the credibilities, the canonical comparison would be of the subsequent loss ratios of the two strata - actual losses to manual premium on the one hand, and actual losses to modified premium on the other. The first ratios, actual to manual, should follow the predicted quality of the stratum - risks with credit mods should prove to have favorable loss ratios on the average and those with debits poor ratios, showing that the plan was indeed able to separate the wheat from the chaff. In order to see if the differences in predicted quality were correctly offset by the mod, the loss to modified premium ratios of the two groups would be equal to each other. It would be too much to expect that premium rates are correct in aggregate and that the two subsequent loss to modified premium ratios be equal to the permissible loss ratio.

Unfortunately, effective manual premium rates corresponding to the experience period reported in the U.S.P. are not retained in the files. What can be found are the ELR's or Expected Loss Rates by class. These are not the true loss costs underlying the rates, but estimates of emerged loss for three policy years as of a certain evaluation date. In this case, we observe that the ELR's used to estimate rating year 1980 expected losses were the ones used to compute the expected ratable losses in the experience period for the 1982 policy year. The ELR's are meant to be correct on the average for the losses on three policy years, including, in this case, 1978, 1979 and 1980. Each policy year is at a different maturity. ELR's are probably not correct for any single policy year, but should bear some reasonable relation to the rates effective in the latest year. Our assumption is that the ELR's will be uniformly redundant or inadequate over all insureds with the same rate. There might be some discrepancy owing to emergence patterns which vary by class, but I don't think they put any bias in the test.

The ratios used for testing the mod, then, are ratios of loss, not to premium, but to expected loss. The comparison is between the ratios of loss to manual expected loss for the strata (which should be low for the credit risks and high for the debit risks), and ratios of actual loss to expected losses adjusted by the modification, or modified loss ratios. Specifically, 1980 actual loss to 1980 expected loss, taken from the 1982 rating year files, should reflect the predicted quality difference, while application of the 1980 modification, from the 1980 rating year files, to 1980 actual and expected losses taken from the 1982 data files should make the ratios to modified expected loss converge.

It is reasonable to hope that the subsequent ratios to modified expected loss would be near each other, but unreasonable to require values near unity.

Using this simple test on the actual data, we were able to determine a need for credibility in the form much like the one finally selected, so that even the smallest ratable risks would have non-zero excess as well as primary credibilities. Credibility would at first increase rapidly with risk size, then increase at a slower rate, but never reach full credibility for even the largest risks. We were aware that credibility in a form like this had been proposed by Meyers[3]. This would correspond to second level variance assumptions described in Section 2.

Exhibit II-1 shows some early results of using this naive approach.

B. The Quintiles Test

As the testing of the plans progressed and more sophisticated actuarial theory was applied to the algebraic form of the credibility constants, it became apparent a more sophisticated test would be needed to measure the quality of alternate formulae. Dorweiler's sufficient criterion for correctness of the modification was that any a priori subdivision of risks made before the subsequent experience is known should enjoy nearly uniform emerged loss ratios to modified premium. It would be impossible to try all possible stratifications of risk quality; certainly any underwriting information such as rating schedules have long since been lost to even the most assiduous researcher. The approach we finally chose was very much like the one used by Dorweiler himself.

We want the mod to distinguish risks of inherently different quality and then correct accurately for those differences. Instead of good vs. bad as in the naive test, we grouped the risks into five percentile strata according to the value of their modifications. The lowest 20% of the values belonged to risks in the first quintile, the next 20% to the second and so on. This is our a priori subdivision. The subsequent aggregate unmodified loss ratios of the strata should reflect the quality difference recognized by the mod. Application of the modification to the expected losses in the denominator should cause the ratios to flatten across the strata.

To combine these two facts into one statistic took the realization that in the first case, the variation should be great and in the second it should be small. This led to the ratio of two sums of squared differences, the five squared deviations from the mean of the modified loss ratios, divided by the sum of squared deviations before modification. The lower the better. The statistic would pertain to the experience of each size group, so for a particular parametrized mod formula, several such values would be available for comparison. In most of the testing, coincidentally, five size groups were considered. This was not done to confuse someone trying to make light reading out of a rather large output of intermediate results, but it probably did.

Our test statistic should be compared to the "efficiency" standard used by Meyers[3]. I would characterize our statistic as the ratio of remaining structure variance to the prior structure variance. (We really do not capture the entire structure variance, either prior or posterior, but the ratio is valid). This is the component of variance we can hope to reduce by means of experience rating. Meyers looks at the proportion by which the total variance is reduced. I believe either statistic to be useful; ours is computationally simpler and has an undisputable best value of zero.

C. Estimation of Plan Parameters

The heart of the testing was the trial and error estimation of six parameters, three each for K_p and K_x, which would minimize the test statistic. Section V outlines the variants of the basic plan for which minimal values of the statistics were sought, and some discussion of the rationale for each.

There were a large number of trial and error evaluations done at each stage of development of suitable tests and forms of credibility. Point estimates of constant values of K_p and K_x appropriate for each size group were made; these were seen to follow the general pattern of the curve resulting from the third level of variance assumptions.

In another test, primary and excess credibilities were evaluated separately.

We found that both primary and excess credibilities from these tests were higher than in the tests when they were evaluated separately. This conclusion should be contrasted with Meyers[3] who concluded a best modification formula could be based on primary only losses. I believe his conclusion is correct in the special case of a well behaved severity distribution for all risks, which was the model he tested. I also believe our tests of real world data support the split formula with two part credibility we are using. The Compensation severity distribution is composed of many types of loss. That the distribution of losses by type varies from class to class and risk to risk is an essential component of Workers' Compensation ratemaking and rating.

Exhibit II-1

1981
Actual To Expected Loss Ratio Before and After Experience Rating 7 States Total

Current Experience Rating Formula

Risk Size	Quality Indication		Subsequent Period Loss Ratios	Subsequent Period Modified Loss Ratios
2,500- 5,000	50%	Best	0.75	0.80
	50%	Worst	1.12	1.05
5,000- 10,000	50%	Best	0.71	0.80
	50%	Worst	1.11	1.01
10,000- 25,000	50%	Best	0.79	0.92
	50%	Worst	1.12	0.96
25,000- 100,000	50%	Best	0.75	0.89
	50%	Worst	1.15	0.93
over 100,000	50%	Best	0.71	0.93
	50%	Worst	1.00	0.79
sum of differences			1.79	0.68

Proposed Experience Rating Formula

Risk Size	Quality Indication		Subsequent Period Loss Ratios	Subsequent Period Modified Loss Ratios
2,500 - 5,000	50%	Best	0.76	0.91
	50%	Worst	1.10	1.00
5,000 - 10,000	50%	Best	0.71	0.89
	50%	Worst	1.11	0.98
10,000 - 25,000	50%	Best	0.79	0.98
	50%	Worst	1.12	0.95
25,000 - 100,000	50%	Best	0.75	0.91
	50%	Worst	1.15	0.94
over 100,000	50%	Best	0.72	0.89
	50%	Worst	0.99	0.83
sum of differences			1.74	0.30

SECTION III-PLANS TESTED

A. Introduction

The first plan tested was a rather modest departure from the then current experience rating formula, herein also referred to as the "former" or even "former current" plan. We retained most of the elements of that plan, including D-ratios (which measure the primary component of expected loss by class), ELR's, the primary-excess split formula and state ratable loss limitation. This latter was 10% of the Self Rating Point (defined below) for single claim occurrences and 20% for multiple claim occurrences.

The former formula was derived through practical simplifications which made sense at the time of its development. It was partly these simplifications, however, that moved the plan away from whatever underlying credibility theory it may have had. This formula was written as follows:

11) $$M = \frac{A_p + WA_x + (1 - W)E_x + B}{E + B}$$

one fraction, with weighting value W and credibility ballast B, both linear functions of risk total expected losses. A denotes actual, E expected. Subscripts p and x denote primary and excess portions of loss respectively.

$$W = \begin{cases} 0 & \text{for} \quad E < 25{,}000 \\ \dfrac{E - 25{,}000}{SRP - 25{,}000} & \text{for} \quad 25{,}000 \le E \le SRP, \\ 1 & \text{for} \quad E > SRP \end{cases}$$

$$B = 20{,}000(1 - W)$$

SRP is the state self-rating point, 25 times the state average serious cost per case.

In the former multi-split formula, the primary portion of a loss L was L_p.

12) $$L_p = L \quad \text{for } L \leq 2{,}000$$
$$= L * \frac{10{,}000}{8{,}000 + L} \quad \text{for } L > 2{,}000$$

Somewhere else in the literature the reader can find rationalizations for this split formula, but a general characterization of primary losses representing frequency and excess losses reflecting severity seems adequate without any undeserved pretention.

As described above, the SRP was used in the limitation of ratable losses. Denoting the loss as so limited by L_r, the excess portion of a loss greater than $2,000 would be

13) $$L_x = L_r - Lp$$

where L_p is calculated as noted above.

The new modification formula was (10) from the credibility section, written in terms of actual, A, and expected, E, losses from the normal experience period used for rating. K denotes a credibility ballast.

10) $$M = 1 + \frac{A_p - E_p}{E + K_p} + \frac{A_x - E_x}{E + K_x}$$

That both primary and excess credibilities depend on total expected losses E is the assumption underlying formula (8) for the credibility ballast values K_p and K_x. The same assumption underlies the former formula which is Perryman's First Formula, as distinguished from our newly regenerated Perryman I.

Setting $B = K_p$ and $W = (E + K_p)/(E + K_x)$ into (11) results in algebraic equivalence of the two formulae, (10) and (11). This would later allow easy implementation of a new plan in the same form as the old. Throughout the testing used to evaluate parameters, we used the form (10) for the mod, and concentrated on finding best values of K_p and K_x. Of course, the values of K_p and K_x which worked best in all of our testing led to values of W and B quite unlike those above.

It was clear that differing benefit levels by state should be reflected in credibility constants K_p and K_x. The current formula used the SRP to effect a nominal difference in the W and B tables by state, but only really impacting the risks large enough to be near the self rating point. We wanted to use a more vital adjustment which would result in a true scaling by state, valid across all risk sizes. This was accomplished by insertion of a value G, measuring relative benefit levels by state, into the formulas for K_p and Kx. As such, quite early on we used (9) to make the following expression for K_p by state:

14) $K_p = CE + GH/(1 + GF/E)$

and a similar version for K_x.

The G value would not only account for benefit level, but would index credibility constants for inflation in average claim size. This can be easily seen in the following analysis. Assume inflation of 1 + i between times t and s. For example, let primary credibility at time t be given by

$$Z(t) = \frac{E}{E + K_p}$$

$$= \frac{E}{E + CE + GH/(1+ GF/E)}$$

With inflation, both E and G increase by the factor 1 + i; it can be seen that this factor will cancel everywhere in the formula for Z_s so that

$$Z(s) = Z(t)$$

The formula for G had yet to be determined and was one of the parameters varied to optimize the test statistic. For most of the testing, G was taken as a linear function of the existing SRP.

The SRP was also retained for use in limitation of ratable losses, just as in the current plan. There would be no self-rating, however, under any analytic plan.

B. Plans Tested

Ultimately, we tested four alternate plans besides the former, herein called the Current Multi-Split formula. For each plan other than the existing one, optimal values of the credibility parameters were chosen based on results of the testing. Selection of a final plan from among the four optimized alternatives considered not only the associated values of the test statistic, but also ease of understanding and ease of implementation.

Specifications for the plans:

1. Current Multi-split
2. Perryman I
   Multi-split
3. Perryman II
   Multi-Split
4. Perryman I
   Single Split
5. Perryman II
   Single Split

1. Current Multi-Split Experience Rating Plan

   The basic specifications for this plan are given in the introduction of this section. These include formulas for B, W, the SRP, the primary/excess split of actual losses, and the modification formula itself. It also includes calculation of the Expected Loss Rates (ELR) and D-ratios by class. The rating values of each insured are included in the experience rating files for each rating year, in particular 1981 through 1984, which were used in the testing.

2. Perryman I - Multi-split

This is our first basic alternative to the old plan as described in the introduction. It is formula (10), with the old formula (12) used to split actual losses into primary and excess components. Values such as ELR's and D-ratios could be carried over directly from the experience rating files, while K_p and K_x could be easily calculated from the elements of the files: that is, total expected losses of the risk, state identification of the risk (which would be used to fetch indexed SRP and G values), and three coefficients in each formula, selected by trial and error.

The name, Perryman I, was used to distinguish it from the other new variations tested. The genuine Perryman's First Formula resulted from a basic assumption that both primary and excess loss ratio variances are a function of risk total exposure, that is, risk total expected losses. This assumption underlies new formulas (5), (6) and (7) in Section I, as well as the Current Multi-Split formula (11).

3. Perryman II - Multi-split

This formula resulted from a different assumption about loss variance than Perryman I. It is only nominally related to Perryman's Second Formula, as noted below.

In the version tested, it was hypothesized that conditional primary loss variance was a function of risk expected primary loss and excess loss variance was a similar function of expected excess loss.

The formulas for credibilities took the following form.

15) $$Z_p = \frac{E_p}{E_p + K_p''}, \text{ where } K_p'' = CE + GH/(1+GF/E)$$

and E_p is the expected primary losses. Notice that K_p'' ought to be expressed in terms of E_p, not E. This, however, would have further complicated the formulas. It was decided that the selection of C, F and H as determined by performance would incorporate average D-ratios if appropriate and K" could be a function of total expected losses. The resulting credibility parameters could be put in tabular form according to expected primary or excess size of risk and state.

We used the D-ratio appropriate by risk, which I will call δ, to write

$$Z_p = \frac{E_p}{E_p + K_p''}$$

$$= \frac{\delta E}{\delta E + K_p''}$$

16) $$Z_p = \frac{E}{E + K_p''/\delta}$$

Similarly,

$$Z_x = \frac{E_x}{E_x + K_x''}$$

which yields

17) $$Z_x = \frac{E}{E+K_x''/(1-\delta)}$$

Testing this plan was again accomplished using values available from the experience rating files.

For sake of historical accuracy, the true Perryman's Second Formula actually resulted from unusual expressions for credibilities

$$Z_p = \frac{E}{E_p + K_p}$$

$$Z_x = WZ_p$$

He does not derive these expressions and takes great pains to excuse them for not working.[8]

4. <u>Perryman I - Single Split</u>

One of the key assumptions of the Revised Experience Rating formula is the non-correlation of primary and excess loss components. So long as primary losses had a severity component, NCCI researchers were unable to feel fully comfortable with a credibility based plan still using the old primary-excess split.

It is classically assumed that frequency and severity are independent, hence uncorrelated. This is probably not valid, but reasonable. It is less reasonable to assume primary and excess losses defined by the multi-split formula are uncorrelated. Thus, we at one time considered using a modification formula based strictly on frequency and severity. One problem with this is obtaining a valid claim count. How could we be sure that bruises and abrasions were recorded on a consistent basis by all carriers for all risks? Another problem would be educating administrators as well as insureds and insurers in the transition to a new formula.

Our choice was to use a single split of losses into primary and excess. Losses below the single threshold value would be primary, and the portion of a loss in excess of that value, if any, would be excess. We used $2,000 as the single split point, which was a relatively easy choice: this was the smallest size for which individual claims data was reported, so it would be the closest to a frequency/severity dichotomy we could obtain using available data. Expected losses for this plan were the same as on the rating files.

It would be more difficult, but by no means impossible, to modify D-ratios so that they would correspond to the new split formula. This was accomplished by a single factor applied to the multi-split D-ratios in the files which would maintain the aggregate accuracy of the D-ratios. This is to say if the emerged actual to expected primary losses under the old formula had been .971, for example, D-ratios would be adjusted by a value ADJ to maintain the .971 ratio:

$$.971 = \frac{\Sigma \text{ Emerged single split Primary Loss}}{\Sigma \text{ ADJ} * \text{Old D-ratio} * \text{Expected Ratable Loss}}$$

$$\text{ADJ.} = \frac{\Sigma \text{ Emerged single split Primary Loss}}{\Sigma\ .971 * \text{Old D-ratio} * \text{Expected Ratable Loss}}$$

where the summation is over all risks. With D-ratios so adjusted, the formula was tested with K_p and K_e in the now established form (10), until optimal values for the six coefficients were obtained.

5. Perryman II - Single Split

The last plan tested utilized both of the two major variations from the basic Plan 2, Perryman I - Multi-split. Plan 5 used a single primary - excess split of losses, with the credibility formulas from Plan 3. This was the fully equipped model as compared to the other economy versions, so the question would be if there was enough improvement in performance to justify the additional cost and more difficult handling.

The adjustment to D-ratios alluded to above was also necessary in testing this formula.

C. Summary

Test statistics for the five plans tested may be seen in Exhibit III-1. The tests were performed for ratings effective on 1981 as well as 1980 policies. The footnoted parametrizations were those that produced the best statistics for each plan. Of course, "best" is subjective in that no single set of coefficients produced a lowest value for all ten evaluations (five size groups and two years). Still, the pattern for all was that the smaller sizes deserved much more credibility and the larger much less than under the current plan.

Page 2 of Exhibit III-1 shows details of the calculation of the test statistic for Size Group Two in 1980. It shows the ratios by quintile strata, actual to expected and actual to modified expected, on which the squared deviations used in the test statistic are based.

The associated credibilities are graphed in Exhibit III-2. It can be seen that the four optimized plans show the consistent pattern just mentioned. They also bear a fairly logical relation to each other. In particular, credibilities seem to increase substantially in the passage from a multi-split to a single split formula. This may be due to better satisfaction of our assumption that primary and excess losses are uncorrelated which is presumably better satisfied by the single split plan.

By contrast, the use of the Perryman II equation in place of Perryman I does not seem to increase average credibilities much. There is of course a slight improvement in the distribution of the credibility assigned to the individual risk, as reflected in the test statistic. As described in Section IV, the evaluation of this and all plans weighed the benefit of increasing the accuracy of the plan with the cost of increased complexity in application.

SUMMARY STATISTICS

FORMULA	1980 SIZE GROUP ONE	1980 SIZE GROUP TWO	1980 SIZE GROUP THREE	1980 SIZE GROUP FOUR	1980 SIZE GROUP FIVE	1981 SIZE GROUP ONE	1981 SIZE GROUP TWO	1981 SIZE GROUP THREE	1981 SIZE GROUP FOUR	1981 SIZE GROUP FIVE
CURRENT PLAN	.3277	.2236	.0918	.0228	.0293	.3230	.2361	.1116	.0453	.2187
PERRYMAN I MULTI SPLIT	.1978	.1248	.0994	.0148	.0012	.2664	.1674	.0930	.0380	.0831
PERRYMAN II MULTI SPLIT	.1632	.1058	.0976	.0112	.0033	.1809	.1333	.0985	.0414	.0980
PERRYMAN I SINGLE SPLIT	.0852	.0519	.0459	.0169	.0042	.1140	.0838	.0688	.0331	.0782
PERRYMAN II SINGLE SPLIT	.0803	.0366	.0380	.0091	.0075	.0785	.0735	.0583	.0312	.0187

PERRYMAN I
MULTI SPLIT

KP = (.067)T + 17000G/(1+3100G/T)
KX = (.600)T + 560000G/(1+5000G/T)
G = SRP/570000

PERRYMAN II
MULTI SPLIT

KP = (.068)T + 6900G/(1+1600G/T)
KX = (.670)T + 260000G/(1+5500G/T)
G = 1

PERRYMAN I
SINGLE SPLIT

KP = (.098)T + 2500G/(1+700G/T)
KX = (.750)T + 200000G/(1+5100G/T)
G = SRP/2700000 + .85

PERRYMAN II
SINGLE SPLIT

KP = (.040)T + 850G
KX = (.600)T + 97000G/(1+2500G/T)
G = SRP/570000

NATIONAL COUNCIL ON COMPENSATION INSURANCE

1980
ACTUAL TO EXPECTED LOSS RATIO
BEFORE AND AFTER EXPERIENCE RATING
PERRYMAN I
SINGLE SPLIT
KP = (.098)T + 2500G/(1+700G/T)
KX = (.75)T + 200000G/(1+5100G/T)
G = SRP/2700000 + .85
15 STATES TOTAL

5,000 - 10,000

OBS	QUINTILE	BEFORE	SQUARED DEVIATION FROM MEAN	AFTER	SQUARED DEVIATION FROM MEAN
1	1	0.62	972	0.83	106
2	2	0.76	295	0.95	2
3	3	0.86	58	0.95	4
4	4	1.05	148	1.00	42
5	5	1.32	1532	0.92	2
6	MEAN TOTAL:	0.93	3005	0.93	156

TEST STATISTIC

$$0.0519 = \frac{156}{3005}$$

CURRENT EXPERIENCE RATING

OBS	QUINTILE	BEFORE	SQUARED DEVIATION FROM MEAN	AFTER	SQUARED DEVIATION FROM MEAN
1	1	0.68	623	0.79	192
2	2	0.70	532	0.78	214
3	3	0.87	37	0.93	0
4	4	1.06	173	1.04	112
5	5	1.30	1372	1.03	94
6	MEAN TOTAL:	0.93	2737	0.93	612

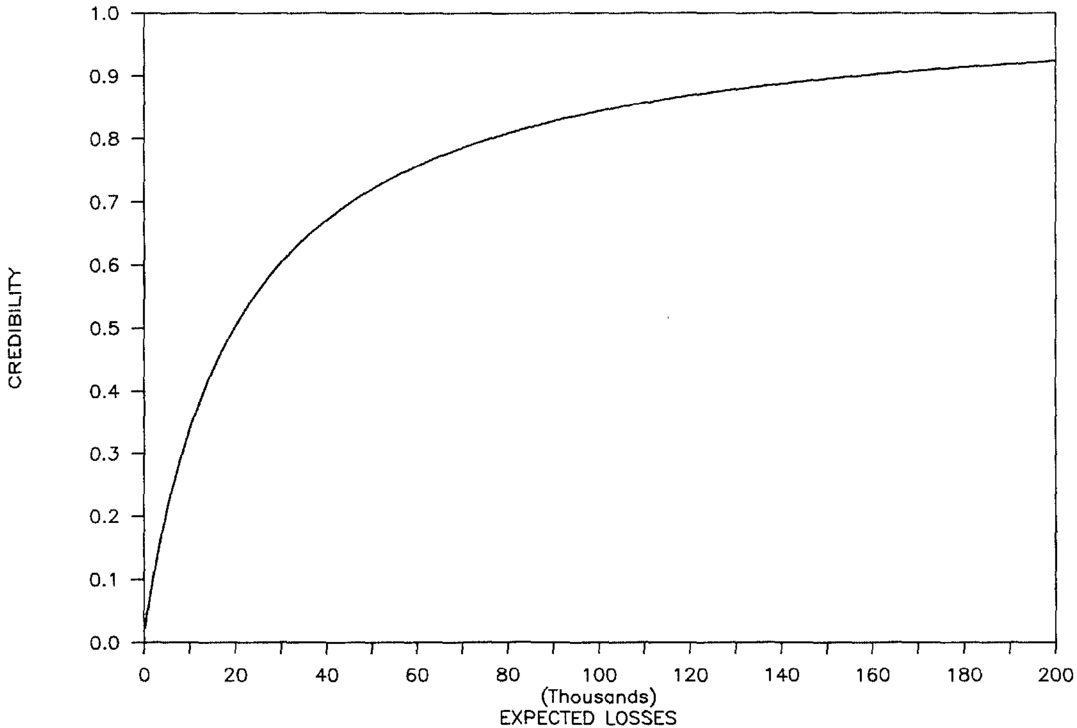
TEST STATISTIC

$$0.2236 = \frac{612}{2737}$$

Exhibit III-2, Page 1

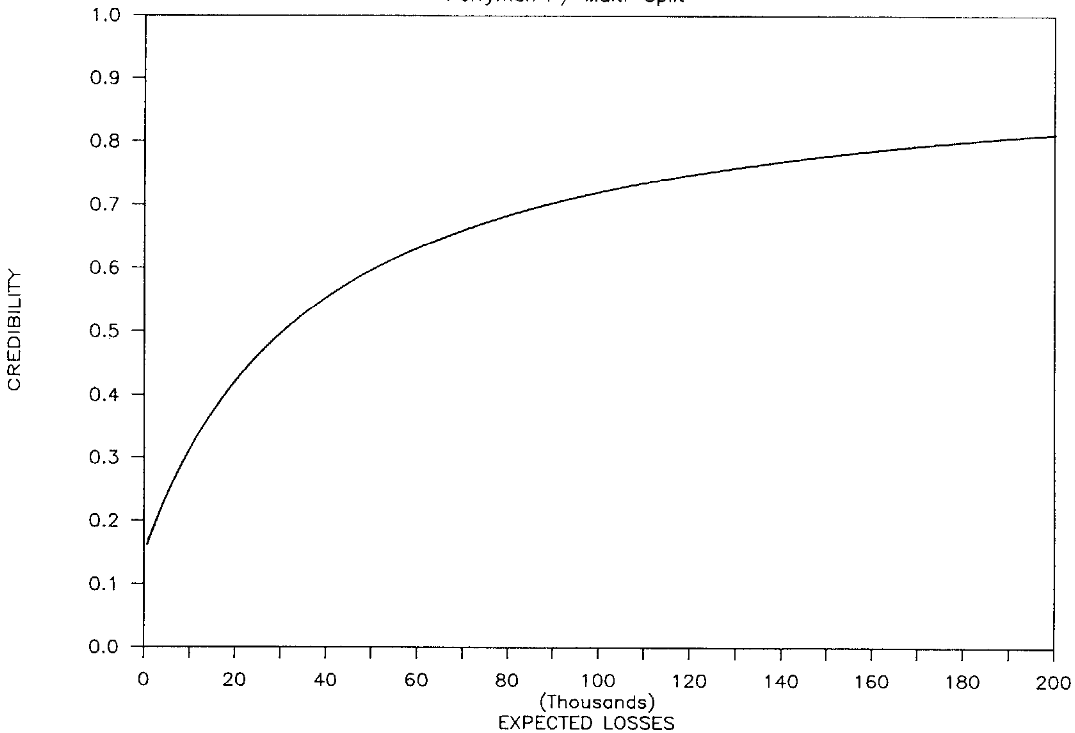
PRIMARY CREDIBILITY

Current Plan



PRIMARY CREDIBILITY

Perryman I / Multi-Split



PRIMARY CREDIBILITY

Perryman II / Multi-Split

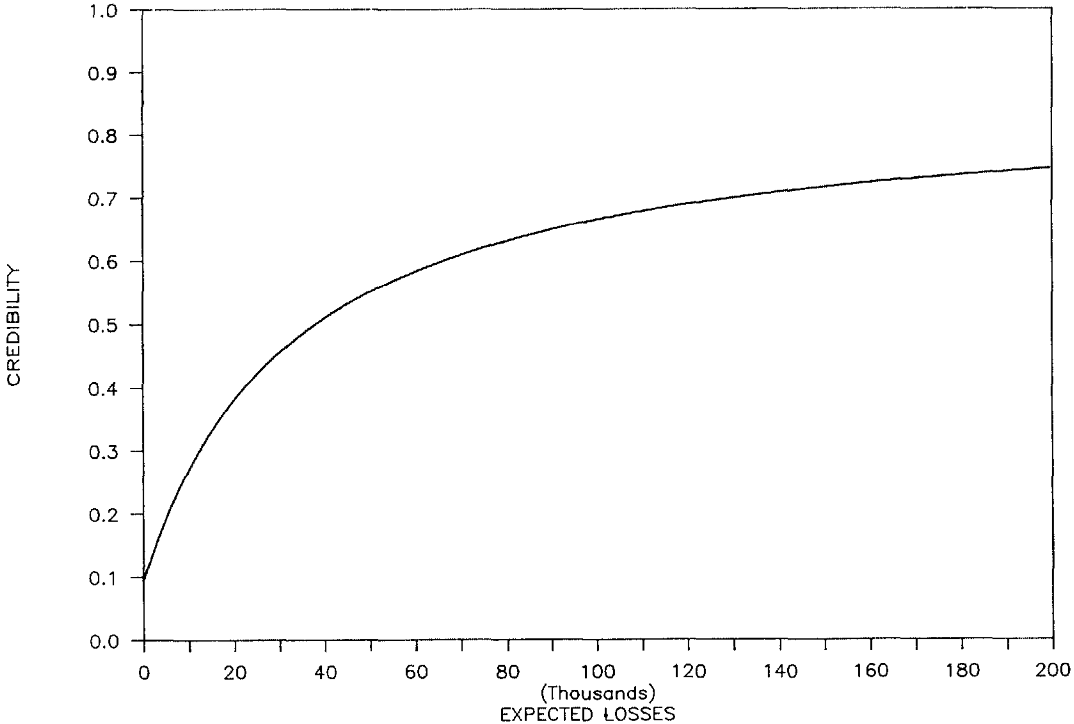
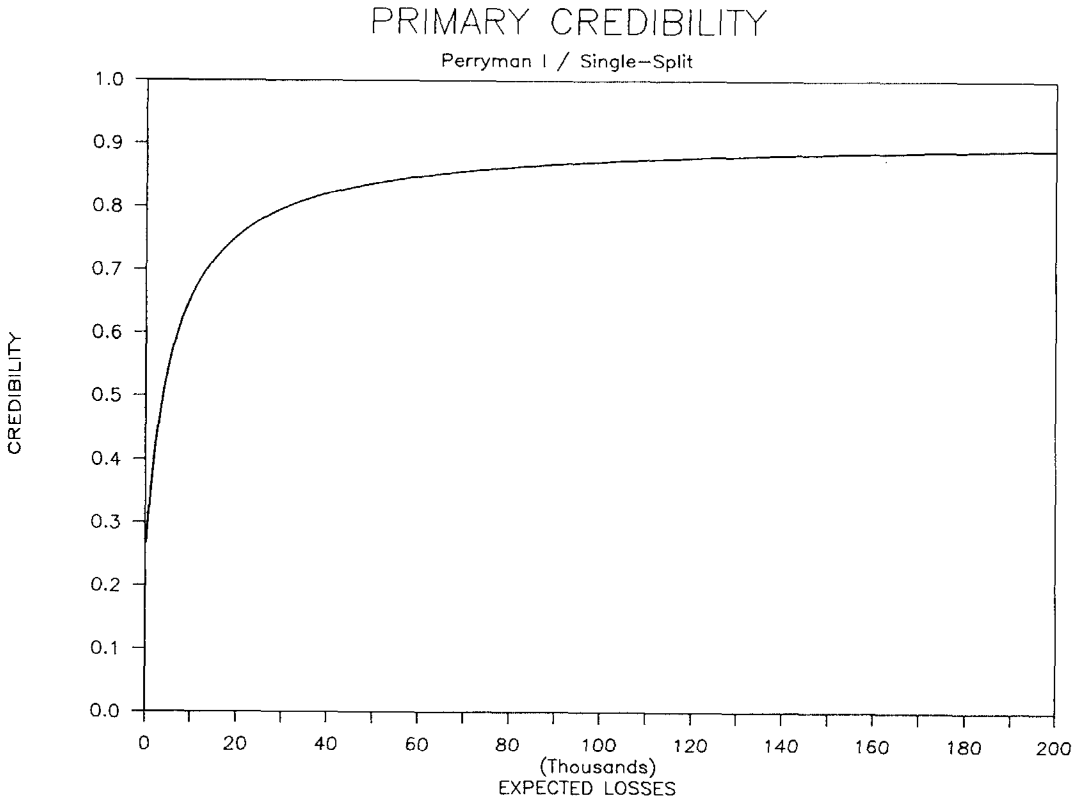


Exhibit III-2, Page 4



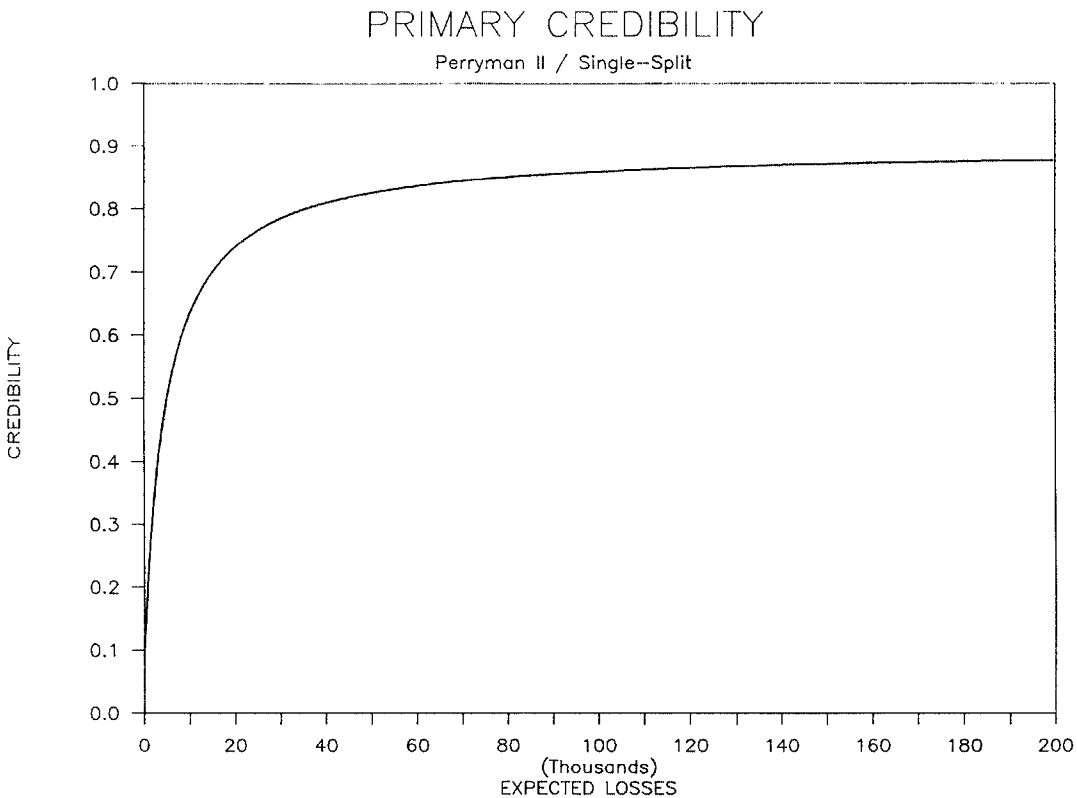


Exhibit III-2, Page 5

EXCESS CREDIBILITY

Current Plan

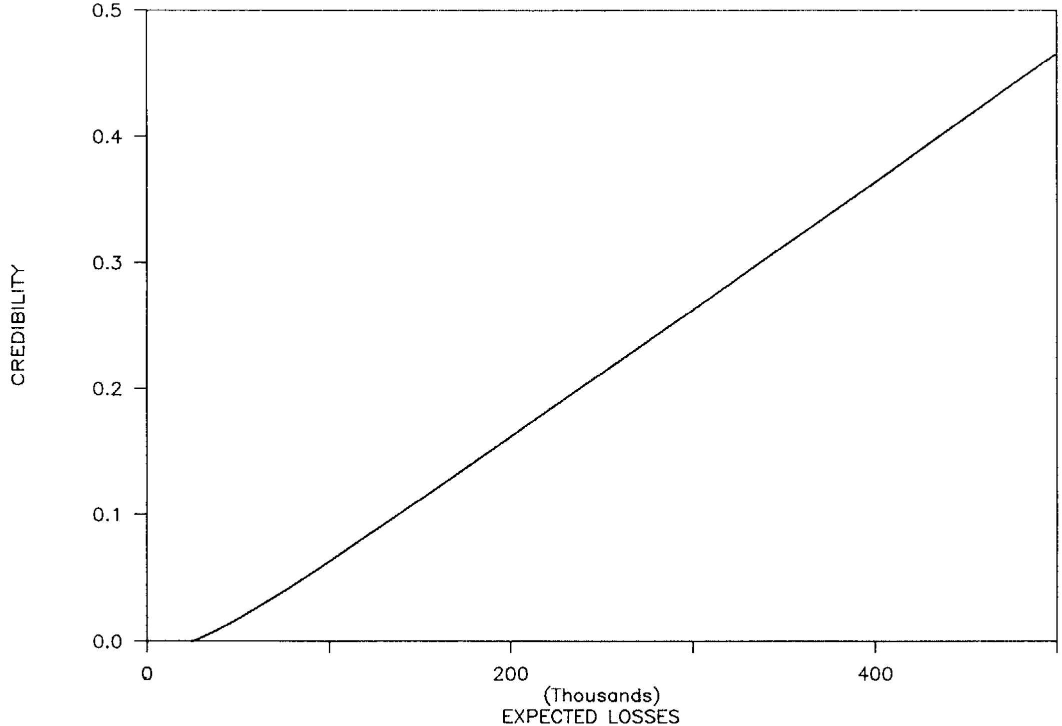
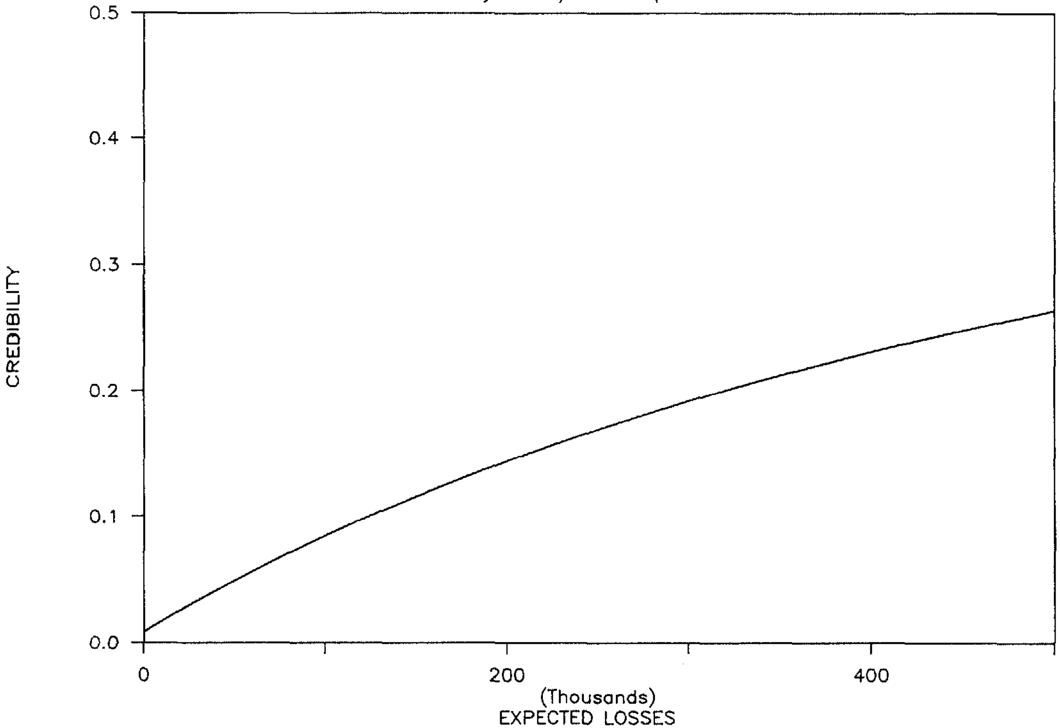


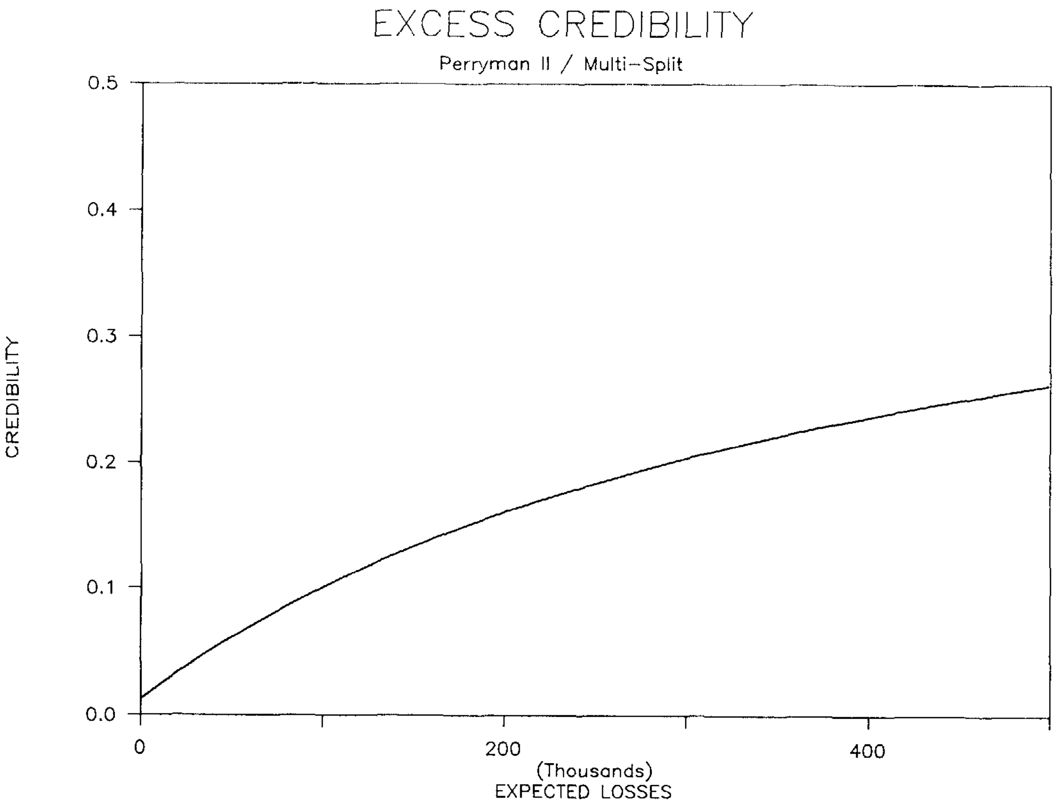
Exhibit III-2, Page 6

Exhibit III-2, Page 7

EXCESS CREDIBILITY

Perryman I / Multi-Split




EXCESS CREDIBILITY
Perryman II / Multi-Split
CREDIBILITY
0.5
0.4
0.3
0.2
0.1
0.0
0
200
400
(Thousands)
EXPECTED LOSSES

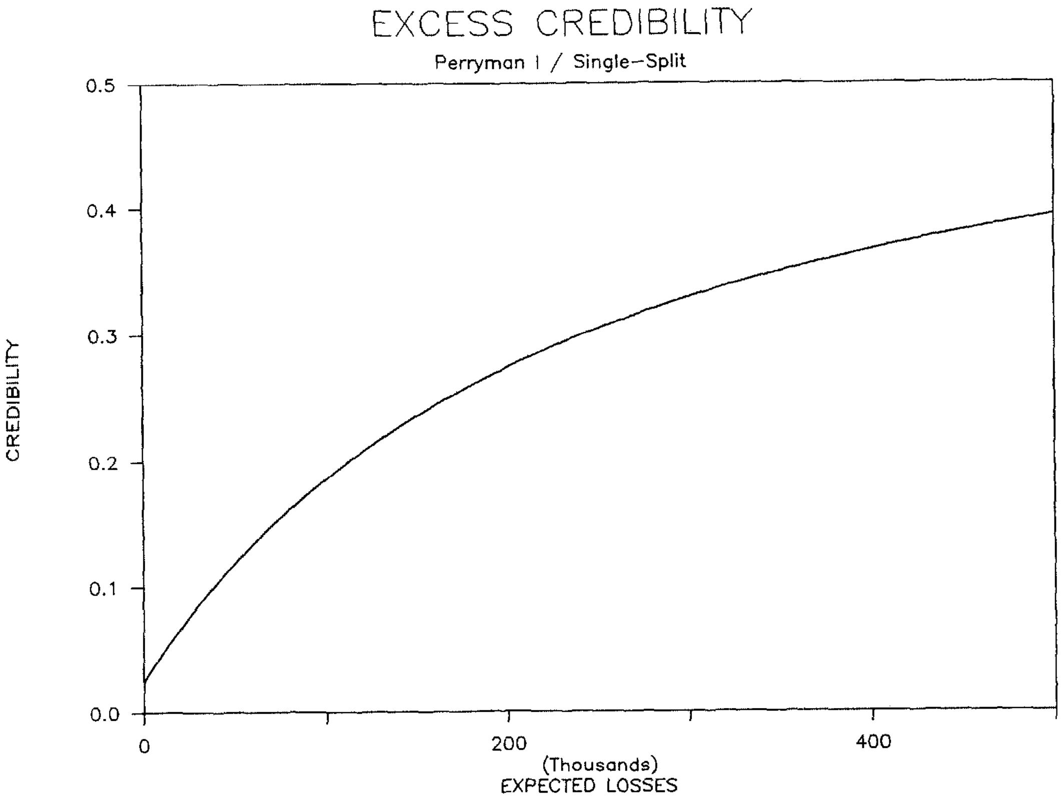
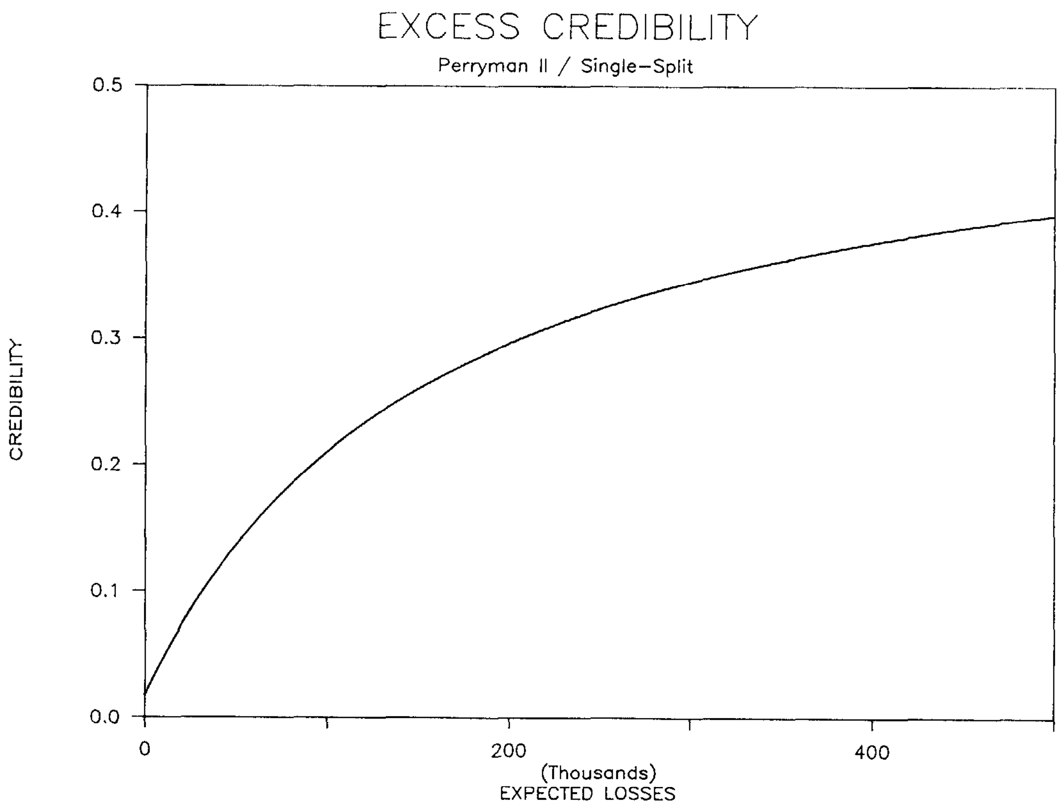


Exhibit III-2, Page 9


EXCESS CREDIBILITY
Perryman II / Single-Split
CREDIBILITY
0.5
0.4
0.3
0.2
0.1
0.0
0
200
400
(Thousands)
EXPECTED LOSSES

SECTION IV-SELECTION OF PLAN

A. Selection

Choice of a preferred experience rating formula was made primarily on the basis of performance, of course, but also on practical considerations. Ease of acceptance and implementation were among the considerations that unfortunately had to be made. Fortunately, this did not lead to any great compromise of actuarial principle.

The Individual Risk Rating Plans Subcommittee (IRRP) approved Perryman I Single Split as parametrized in Exhibit V-1. The performance was nearly as good as the Cadillac Perryman II Single Split, but the slight improvement offered by the latter did not appear to outweigh the effort necessary to make the more complex changes. The improvement offered by single split over multi-split was significant and the transition would not be as difficult.

B. Trending the Split Point

In the matter of a split point, the Subcommittee also recommended that appropriate trending be applied to the single split point used in the testing so that it would have the same relativity to the loss size distribution when actually applied. Trending was based on several years of change in the average cost per case, which led to the single split point of $5,000, used in the filing. This is a reasonable value, given that it will be 1990 or later before the revised plan is widely accepted, and a few more years after that before any study can be done to revise the point.

Still, several researchers thought we might be taking it too high; after all, we thought the single split worked well because of a resemblance to a frequency / severity model, and $5,000 leaves a lot of room for claims of different size. The rationale for such a high selected value was two-fold. First, there would be no automatic adjustment, and the $5,000 would be retained until a study could be made to determine the optimal value. Second, the resultant D-ratios for the single split point would not have to decrease dramatically from the ones in the old plan. Loss size trend would make primary losses increasingly resemble claim count

It should be observed that the decision not to index the split point was very consciously made. As a consequence, it can be expected that average D-ratios will decrease over time and primary credibilities should be monitored.

C. <u>Decreasing Swing</u>

Consider Exhibit IV-1. This shows the average change in modification, plan to plan, for small risks grouped by value of the then current modification. These risks comprised the smallest size group used in testing the original formula, as well as in this particular test. Small risks whose 1985 mod exceeded 1.20 could expect an average increase of 62 points in their mods! Even if this reflected correctly calibrated credibility, several of the committee representatives thought it would be unacceptable in the market. Some even doubted it was correct at all, despite the test evidence that credibilities were optimal for this size group. Of course, the tests worked on averages, and these were extreme cases, so it was possible to believe the tests, yet believe this was a problem that needed to be fixed.

The attack on this problem resulted in two changes to the plan: decreasing the SRP, as it affected limitations on ratable losses, and establishing minimum values for credibility constants Kp and Ke. It was decided to make the SRP a multiple of average cost per case by state (SACC), rather than a multiple of the average serious cost per case. Since the average cost per case was between $750 and 3,500 for most states in 1981, a multiple of A = 250 times SACC generally led to smaller SRP's than those effective at the time, usually close to $1,000,000. The limitation on ratable losses to 10% of the SRP would then lessen the impact of a single large loss on the mod.

Several minimum values of K_p and K_x were tested also, but our analysis quickly led to min K_x = 150,000 and min Kp = 7,500 which worked well in conjunction with the new loss limitations in the range above.

Exhibit IV-2 shows comparisons of the "swing" in mods for groups of risks by size with 1986 mod greater than 1.2, these computed for A = 250 or 300 with min K_p = 7,500 and min K_x = 150,000. In all cases, the swing was less than 25 points. These represent the end of an elimination process which considered many other possible values.

Changing the SRP formula also led us to reexamine calculation of the state scale factor G. The older G formula may be seen on Exhibit III-1, p.1 as G = .85 + SRP/2,700,000, where the SRP was the value from the former plan. The new formula, resulting from some trial and error, was G = SACC/1,000, which worked well with the modified plan.

D. Final Changes to the Plan

Independent tests of the new plan still showed the potential for large swings in the values of the mod for risks in the smaller size categories. After considerable discussion, the Individual Risk Rating Plans Subcommittee recommended one more change to the rating plan. Rather than tamper with credibility constants, loss limitations or split points, the Subcommittee decided to put absolute caps on the mods of smaller size risks. In this way, a small debit under the current formula could increase only a limited amount with change to the new formulas. The following table lists the limits by size.

Expected Loss Size	Maximum Modification
$0 < E < 5{,}000$	1.6
$5{,}000 \leq E < 10{,}000$	1.8
$10{,}000 \leq E < 15{,}000$	2.0

Exhibit IV-3 is similar to Exhibit IV-1 and shows that the "swing" problem is greatly reduced by this action.

There would be no transition program to phase this out, except the impact of inflation. As experience rating eligibility increased, fewer risks would enjoy a potential 1.60 cap. This action was practical in the sense of ease of implementation which was needed at the late date of the change. In addition, it was not without some actuarial justification. Several people believe that mods higher than the stated limits are probably not deserved, all statistical arguments notwithstanding. Actually, these test results only showed that mods were on average correct for the worst 20% of risks. Other testing (not shown) using higher percentiles than 80th showed the new formula could result in unreasonably high mods at least for the smallest risks. In addition, risks just below eligibility had a maximum modification of unity and there ought to be some continuity at the point of eligibility.

Finally, Exhibit IV-4 shows quintiles test statistics for the finalized plan. These compare reasonably well with statistics in Exhibit III-I.

1985

PERRYMAN I - SINGLE SPLIT

CHANGES IN AVERAGE MODIFICATIONS
FOR RISKS GROUPED BY VALUE OF CURRENT MOD

$2,500 TO $5,000 EXPECTED LOSSES DURING EXPERIENCE PERIOD

RANGE OF CURRENT MODS	NUMBER OF RISKS	AVERAGE CURRENT MOD	AVERAGE PROPOSED MOD	CHANGE
.80 TO .84	1	0.83	0.66	-0.17
.85 TO .89	32	0.89	0.70	-0.19
.90 TO .94	9,062	0.93	0.79	-0.14
.95 TO .99	9,600	0.96	0.87	-0.09
1.00	484	1.00	1.02	0.02
1.01 TO 1.05	1,637	1.03	1.13	0.10
1.06 TO 1.10	1,238	1.08	1.31	0.23
1.11 TO 1.15	944	1.13	1.42	0.29
1.16 TO 1.20	742	1.18	1.48	0.30
OVER 1.20	2,578	1.35	1.97	0.62
TOTALS	26,318	1.01	1.03	0.02

1986

PERRYMAN I SINGLE SPLIT, AS MODIFIED*

CHANGE IN CURRENT MOD TO PROPOSED MOD
RISK HAVING CURRENT MOD GREATER THEN 1.2

PLAN	OVER ALL	SIZE GROUP ONE	SIZE GROUP TWO	SIZE GROUP THREE	SIZE GROUP FOUR	SIZE GROUP FIVE
<u>SERA</u>						
A = 250	.06	.22	.21	.16	.09	- .05
A = 300	.06	.22	.21	.16	.09	- .05

*SRP = A x (AVERAGE COST PER CASE)

MINIMUM Kp = 7,500 AND Kx = 150,000

1985

CHANGES IN AVERAGE MODIFICATIONS
FOR RISKS GROUPED BY VALUE OF CURRENT MOD

$2,500 TO $5,000 EXPECTED LOSSES DURING EXPERIENCE PERIOD

RANGE OF CURRENT MODS	NUMBER OF RISKS	AVERAGE CURRENT MOD	AVERAGE PROPOSED MOD	CHANGE
.80 TO .84	18	0.83	0.79	-0.04
.85 TO .89	620	0.88	0.84	-0.04
.90 TO .94	10,951	0.93	0.88	-0.05
.95 TO .99	13,522	0.96	0.93	-0.03
1.00	682	1.00	1.02	0.02
1.01 TO 1.05	2,307	1.03	1.08	0.05
1.06 TO 1.10	1,675	1.08	1.20	0.12
1.11 TO 1.15	1,278	1.13	1.26	0.13
1.16 TO 1.20	1,066	1.18	1.31	0.13
OVER 1.20	3,576	1.35	1.49	0.14
TOTALS	35,695	1.01	1.02	0.01

With minimum Kp 7,500 and
minimum Kx = 150,000
SRP = 250 x SACC
G = SACC 1,000
MODS LIMITED

1981

SUMMARY STATISTICS - QUINTILES TEST

PLAN	SIZE GROUP ONE	SIZE GROUP TWO	SIZE GROUP THREE	SIZE GROUP FOUR	SIZE GROUP FIVE
SERA					
A = 250	.1800	.1183	.0546	.0127	.1588
A = 300	.1704	.1150	.0645	.0524	.3138

SRP = A x (AVERAGE COST PER CASE)
MINIMUM K_p = 7,500 AND K_x = 150,000

ACKNOWLEDGEMENT

Many people played a significant role in development of the Revised plan, including NCCI Management, its committees and a host of actuarial technicians.

Special credit is due to two people without whom there would probably be no new plan: Gary Venter, whose strong theoretical bias and ability to quickly assimilate the work of others is at the heart of the plan, and Tony Pipia, whose knowledge of NCCI experience rating data files and belief in purity of form underlay most of the testing. Although I am lucky enough to have been a part of this effort nearly from the beginning, my own role has become essential probably only in the implementation and documentation phases; this paper is a manifestation of the latter responsibility.

REFERENCES

[1]Leroy J. Simon and Robert A. Bailey, "An Actuarial Note on the Credibility of Experience of A Single Private Passenger Car," PCAS XLVI, 1959, p. 159.

[2]Charles C. Hewitt, Jr., "Credibility for Severity," PCAS LVII, 1970, p. 148.

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

[4]Howard Mahler, "Review of Meyers (3)," PCAS LXXIV, 1982, p. 119.

[5]Gary G. Venter, [CAS Textbook Chapter]

[6]Charles C. Hewitt, Jr., "Loss Ratio Distribution-A Model," PCAS LIV, 1967, p. 70.

[7]Paul Dorweiler, "A Survey of Risk Credibility in Experience Rating," PCAS XXI, 1934, p.1.

[8]Francis S. Perryman, "Experience Rating Plan Credibilities," PCAS XXIV, 1937, p.60.