PRICING PROPERTY AND CASUALTY INSURANCE PRODUCTS

1980 DISCUSSION PAPER PROGRAM

Presented May 11-14, 1980
Caribe Hilton, San Juan, Puerto Rico

CASUALTY ACTUARIAL SOCIETY

ORGANIZED 1914

PRICING PROPERTY AND LIABILITY INSURANCE PRODUCTS

These papers and reviews have been prepared in response to a call for papers by the Casualty Actuarial Society to provide discussion material for its Spring meeting, May 11-14, 1980, in San Juan, Puerto Rico.

A wide range of pricing topics is covered by these papers. It is hoped that they will further the knowledge of CAS members and others in this important area. In addition, we trust the efforts of the authors and reviewers will lead to further research into factors which influence the pricing of insurance products.

Committee on Continuing Education

NOTICE

The Casualty Actuarial Society is not responsible for statements or opinions expressed in the papers or reviews in this publication. These papers and reviews have not been reviewed by the CAS Committee on Review of Papers.

TABLE OF CONTENTS

TITLE: RELATIVITY PRICING THROUGH ANALYSIS OF VARIANCE

AUTHOR: Mr. Carl Chamberlain

Mr. Chamberlain is an Assistant Secretary with Insurance Company of North America. He received his BS degree in Math from the University of Pennsylvania and is a member of the Institute of Mathematical Statistics and the American Statistical Association. He is active on ISO and National Council committees.

REVIEWER: Mr. Daniel Goddard

Mr. Goddard is Associate Actuary for Industrial Indemnity Company. He holds a BA degree from Yale University and received his FCAS in 1977.

I Introduction

The ideas for this paper were the outgrowth of considerations of the Construction-Protection relativity question in Commercial Property (Fire) Insurance. A literature search on this subject in the Proceedings suggests Bailey's method from Bailey & Simon's paper "Two Studies in Automobile Insurance Ratemaking" (Proceedings Vol. XLV II).

Using that paper as a base line, the only problem existing in their work is the conceptual model. Their model was either additive or multiplicative. Because of the direct similarities between the basic data that we have to investigate and the mathematical statistics formulation of an Analysis of Variance problem, this paper will purpose a model which contains both additive and multiplicative (interaction) terms.

Section II will develop the mathematical two-way model with estimation solutions and present a statistical test for the inclusion/exclusion of the interaction term.

Section III will give numerical examples, one from Commercial Auto and one from Commercial Fire. These examples, while on small data sets calculated by hand, will show the ease and simplicity of the approach and the accuracy of the final result.

Section IV will extend the model to three-way classifications and suggest the possible extension to n-way modelling situations.

Section V, the summary, will bring together all the theoretical and practical considerations that will justify this model under most situations. The topic of credibility and some possible alternatives will be discussed.

II The Two-Way Model

In any relativity problem there are a few questions to be determined before any work can proceed. One of these questions is dimensions. Are we going to consider only two dimensions, or will we need three or more? In Commercial Auto, we could consider such things as age, sex, territory, type of vehicle, etc. all at the same time and use an n-way approach.

In Commercial Fire, the question was between Construction and Protection and a two-way classification was sufficient. Higher order models could be considered, although, at any time.

A second question, is that of model specification. Will the model be additive? multiplicative? or some combination of both? Presented here is the specification of a two-way additive and multiplicative model, for which, ultimately, the multiplicative term can be tested for statistical significance.

The object of this exercise is to estimate the relativities, r_{ij}, where $i = 1, . . . p$ (row effects - or α effects) and $j = 1, . . . q$ (column effects - or β effects) with the possibility of the interaction term, $\alpha\beta_{ij}$ in all cells. p is the number of levels of the row effect as q is the number of levels of the column effects. In the numerical example to follow, there are three Protection groups (q = 3) and three Construction groups (p = 3). The actual loss ratio data appears in Section III.

The basic model for r_{ij} appears below:

$$r_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ij} \qquad (1)$$

with the constraints

$$\sum_i n_{ij}\alpha_i = \sum_j n_{ij}\beta_j = \sum_i n_{ij}\alpha\beta_{ij} = \sum_j n_{ij}\alpha\beta_{ij} = 0$$

In an Analysis of Variance (ANOVA) context μ is the grand mean or some overall base line measure, α_i, β_j and $\alpha\beta_{ij}$ are the column, row and interaction effects, respectively and e_{ij} is the error term (the error term puts the equation in balance such that the estimates of the various terms while not exactly the relativities sought will vary by the "random" error term).

At this point, we shift away from the usual ANOVA perspective of trying to determine if the α, β, and $\alpha\beta$ terms significantly reduce the inherent variation or "explain" the modelling situation, to the estimation of the α, β and $\alpha\beta$ terms which are of more interest to us, at the moment.

As you can see, the constraints following Equation (1) have a term n_{ij} in them. This term is a weighting factor. I have used it as a surrogate for credibility in the examples to follow such that in the Auto data, n_{ij} is the appropriate Earned Car years, and in the Fire data, n_{ij} is the premium.

Continuing with the model, the least squares solution calls for the minimization of

$$Z = \sum_{ij} n_{ij} (r_{ij} - \mu - \alpha_i - \beta_j - \alpha\beta_{ij})^2 \qquad (2)$$

which is the minimization of the squared error with appropriate weights, i.e. from (1) $e^2_{ij} = (r_{ij} - \alpha_i - \beta_j - \mu - \alpha\beta_{ij})^2$

To find estimates of α, β, μ, and $\alpha\beta$, Z must be differentiated with respect to each term and set equal to zero. Therefore:

$$\frac{dZ}{d\mu} = 0 \text{ yields } \hat{\mu} = \frac{\sum_{ij} n_{ij} r_{ij}}{\sum_{ij} n_{ij}} \qquad (3)$$

$$\frac{dZ}{d\alpha} = 0 \text{ yields } \hat{\alpha}_i = \frac{\sum_{j} n_{ij} r_{ij}}{\sum_{j} n_{ij}} - \hat{\mu}$$

$$= \hat{A}_i - \hat{\mu} \qquad (4)$$

$$\frac{dZ}{d\beta} = 0 \text{ yields } \hat{\beta}_j = \frac{\sum_{i} n_{ij} r_{ij}}{\sum_{i} n_{ij}} - \hat{\mu}$$

$$= \hat{B}_j - \hat{\mu} \qquad (5)$$

$$\frac{dZ}{d\alpha\beta} = 0 \text{ yields } \hat{\alpha\beta}_{ij} = r_{ij} - \hat{A}_i - \hat{B}_j + \hat{\mu} \qquad (6)$$

Because there is only one observation per cell in this analysis, Equation (6) will give back r_{ij} as the estimate in all cases. This situation takes the modelling approach to an illogical conclusion, and we need to back off for a second to reconsider our position.

If we only consider, for the moment, the estimate given by

$$\hat{r}'_{ij} = \mu + \hat{\alpha}_i + \hat{\beta}_j \qquad (7)$$

we have a strictly additive model, and now we can consider the residual of this equation with r_{ij} from Equation (1). Substituting the estimates in Equation (1), and assuming a zero error, we get

$$\hat{r}'_{ij} - r_{ij} = \alpha\beta_{ij} \qquad (8)$$

Now define the residuals, AB_{ij}, as being equal to

$$AB_{ij} = \hat{r}'_{ij} - r_{ij} = \varepsilon_i \delta_j \qquad (9)$$

where ε_i and δ_j are the multiplicative column and row effects, respectively.

We can now solve this model, and then reformulate our original model to get final estimates for the r_{ij}'s. The new overall model will be

$$\hat{r}_{ij} = \hat{A}_i + \hat{B}_j - \hat{\mu} + \hat{\varepsilon}_i \hat{\delta}_j \qquad (10)$$

The least square solution proceeds as before for the model

$$AB_{ij} = \varepsilon_i \delta_j + e_{ij} \qquad (11)$$

Minimizing $Z = \sum_{ij} n_{ij}(AB_{ij} - \varepsilon_i\delta_j)^2$ and differentiating and setting equal to zero yields the following:

$$\hat{\varepsilon}_i = \frac{\sum_j n_{ij}\, AB_{ij}\delta_j}{\sum_j n_{ij}\delta_j^2} \tag{12}$$

$$\hat{\delta}_j = \frac{\sum_i n_{ij}\, AB_{ij}\varepsilon_i}{\sum_i n_{ij}\varepsilon_i^2} \tag{13}$$

Now we have a complete model for r_{ij} and the estimates are calculated from the data as follows

$$r_{ij} = \hat{A}_i + \hat{B}_j - \hat{\mu} + \hat{\varepsilon}_i\hat{\delta}_j \tag{14}$$

$$\mu = \frac{\sum_{ij} n_{ij}\, r_{ij}}{\sum_{ij} n_{ij}} \qquad \hat{A}_i = \frac{\sum_j n_{ij}\, r_{ij}}{\sum_j n_{ij}} \qquad \hat{B}_j = \frac{\sum_i n_{ij}\, r_{ij}}{\sum_i n_{ij}}$$

$$\hat{\varepsilon}_i = \frac{\sum_j n_{ij}\, AB_{ij}\delta_j}{\sum_j n_{ij}\delta_j^2} \qquad \hat{\delta}_j = \frac{\sum_i n_{ij}\, AB_{ij}\varepsilon_i}{\sum_i n_{ij}\varepsilon_i^2}$$

$$AB_{ij} = r_{ij} - \hat{A}_i - \hat{B}_j + \hat{\mu}$$

It will be shown in the examples that these estimates are always as good as the Baily & Simon approach and there are theoretical reasons to believe that a more general model, by definition, has to be a better approach. It is also true that under this modelling situation the following is true

$$\frac{\sum_i n_i \hat{A}_i}{\sum_i n_i} = 1$$

$$\frac{\sum_j n_j \hat{B}_j}{\sum_j n_j} = 1$$

which is a very nice intuitive result, i.e. the sum (weighted) of the marginal relativities is one.

A final consideration, is that it is possible to statistically test the significance of the interaction term and thereby in some cases reduce the complexity of the model. This testing, and the results of the testing, may not always be appropriate or useful but they are available.

The test, mathematically derived in the Appendix, is an "F-Test" with 1 and $\left[(p-1)\ (q-1)\ -1\right]$ degrees of freedom. Any standard statistical text on Analysis of Variance will clarify the above statement for the statistical novice, e.g. The Analysis of Variance, Henry Scheffé; John Wiley & Sons, Inc., 1959.

The test statistic follows:

$$F = \frac{\left[(p-1)(q-1) - 1\right]\left[\sum_{ij} n_{ij}\ \hat{\varepsilon}_i\hat{\delta}_j\ AB_{ij}\right]^2}{\left[\sum_{ij} n_{ij}\varepsilon_i{}^2\delta_j{}^2\right]\left[\sum_{ij} n_{ij}\ AB_{ij}{}^2\right] - \left[\sum_{ij} n_{ij}\varepsilon_i\delta_j\ AB_{ij}\right]^2} \qquad (15)$$

When the above value is larger than the table F, the interaction term is said to be significant at the α - level, where $(1 - \alpha)$ is the confidence (usually 95%; therefore $\alpha = .05$) of the test.

III Numerical Examples

The first numerical example is taken from Baily & Simon's "Two Studies in Automobile Insurance Ratemaking" (Proceedings Vol. XLVII) using Merit Rating and Class as the proposed discriminate variables on which relativities are sought. Using their data on page 15, the following matries are available for our exercise.

r_{ij}'s (Cell loss ratios ÷ total loss ratio)

Merit Rating Class

Class	A	X	Y	B
1	.786	1.016	1.115	1.358
5	1.071	1.079	1.410	1.642
3	1.212	1.285	1.450	1.885
2	1.269	1.747	1.519	1.784
4	2.050	2.192	2.412	2.853

n_{ij} (# of Earned Car Years)

Merit Rating Class

Class	A	X	Y	B
1	2758	131	164	274
5	64	4	5	9
3	247	16	20	38
2	131	7	10	22
4	157	18	21	57

With the above data, using Equation (14) and the subsequent definitions we will proceed with the calculation of the r_{ij}'s.

$$\hat{\mu} = \frac{\sum_{ij} n_{ij} \, r_{ij}}{\sum_{ij} n_{ij}} = 1.006$$

$$\hat{A}_1 = \frac{\sum_{j} n_{1j} \, r_{1j}}{\sum_{j} n_{1j}} = 0.8584$$

$$\hat{A}_2 = \frac{\sum_{j} n_{2j} \, r_{2j}}{\sum_{j} n_{2j}} = 1.1547$$

$$\hat{A}_3 = \frac{\sum_{j} n_{3j} \, r_{3j}}{\sum_{j} n_{3j}} = 1.3101$$

$$\hat{A}_4 = \frac{\sum_{j} n_{4j} \, r_{4j}}{\sum_{j} n_{4j}} = 1.4750$$

$$\hat{A}_5 = \frac{\sum_{j} n_{5j} \, r_{5j}}{\sum_{j} n_{5j}} = 2.2710$$

$$\hat{B}_1 = \frac{\sum_i n_{i1} r_{i1}}{\sum_i n_{i1}} = 0.9007$$

$$\hat{B}_2 = \frac{\sum_i n_{i2} r_{i2}}{\sum_i n_{i2}} = 1.1912$$

$$\hat{B}_3 = \frac{\sum_i n_{i3} r_{i3}}{\sum_i n_{i3}} = 1.2940$$

$$\hat{B}_4 = \frac{\sum_i n_{i4} r_{i4}}{\sum_i n_{i4}} = 1.6955$$

AB_{ij}

Class	A	X	Y	B
1	.0329	-.0276	-.0314	-.1899
5	.0216	-.2609	-.0327	-.2022
3	-.0028	-.2103	-.1481	-.1146
2	-.1007	.0868	-.2440	-.3805
4	-.1157	-.2642	-.1470	-.1075

Finally, using an iteration procedure,

$\hat{\varepsilon}_1 = 1.046$ $\qquad \hat{\delta}_1 = 0.0171$
$\hat{\varepsilon}_2 = 1.1858$ $\qquad \hat{\delta}_2 = -0.0506$
$\hat{\varepsilon}_3 = 0.7623$ $\qquad \hat{\delta}_3 = -0.0589$
$\hat{\varepsilon}_4 = 1.6931$ $\qquad \hat{\delta}_4 = -0.1862$
$\hat{\varepsilon}_5 = 0.5848$

And the final solutions look like

$\hat{r}_{ij}$

	A	X	Y	B	Class Marginals	$\hat{\varepsilon}_i$
1	.7710	.9907	1.0848	1.3531	0.8584	1.0460
5	1.0697	1.2799	1.3729	1.6234	1.1547	1.1858
3	1.2178	1.4567	1.5532	1.8577	1.3101	0.7623
2	1.3987	1.5745	1.6633	1.8492	1.4750	1.6931
4	2.1757	2.4266	2.5246	2.8516	2.2710	0.5848
Merit Rating Marginals	0.9007	1.1912	1.2940	1.6955	$\hat{\mu}$ = 1.006	
$\hat{\delta}$	0.0171	-0.0506	-0.0589	-0.1862		

Bailey & Simon showed four different models in their paper, using the Customary method (1), and straight multiplication (2), an additive model (3), and a scaled additive model (4). They used three measures to show the "goodness" of their models; Balance, Average Error, and χ^2. The results of the above three measures, with their definitions are listed below including a fourth measure $\hat{\sigma}^2$, the weighted square error, for Bailey & Simon's four methods and ANOVA.

$$\text{Balance} = \frac{\Sigma n_{ij}\,\hat{r}_{ij}}{\Sigma n_{ij}\,r_{ij}}$$

$$\text{Average Error} = \frac{\Sigma n_{ij}\,|\,r_{ij} - \hat{r}_{ij}\,|}{\Sigma n_{ij}\,r_{ij}}$$

$$\chi^2 = K\;\Sigma\left[\frac{n_{ij}(r_{ij} - \hat{r}_{ij})^2}{\hat{r}_{ij}}\right] \qquad \text{Where } K = \frac{1}{200}$$

$$\hat{\sigma}^2 = \frac{\Sigma n_{ij}\,(r_{ij} - \hat{r}_{ij})^2}{\Sigma n_{ij}}$$

	Method 1	Method 2	Method 3	Method 4	ANOVA
Balance	1.0103	1.0011	1.0006	0.9983	0.9999
Avg. Error	.0401	.0317	.0098	.0111	0.0256
χ^2	.1021	.0363	.0104	.0083	0.0276
$\hat{\sigma}^2$	.0136	.0030	.0009	.0006	0.0020

When one reviews the table, it is difficult to pick a clearly better model from Method 3, Method 4, and ANOVA. It appears that for this data base, an additive and a multiplicative model, Methods 3 and Method 4, respectively, fit the data equally well. Apart from this area of confusion is the ability of ANOVA to fit all types of data with equal accuracy. To further advocate this new model, is the ease of calculation of the estimates and the property stated earlier, that the weighted marginals sum to one,

$$\frac{\sum_i n_i \hat{A}_i}{\sum_i n_i} = 1$$

$$\frac{\sum_j n_j \hat{B}_j}{\sum_j n_j} = 1$$

A final compelling consideration is the interaction term. For this data set, the calculated F value is 22.41 which is significant at the 5% level. Therefore, statistically, an interaction term is appropriate and also intuitive. It is very easy to believe that there is most likely some interaction between Class and Merit rating.

The next example comes from Commercial Fire.

The data and estimates follow:

LOSS RATIOS

Construction	Protection 4-8	1,9	2,3	Total
1	0.569	0.477	0.558	
2,3	0.432	0.420	0.558	
4-6	0.445	0.463	0.369	
Total			0.474	

Relative Loss Ratios (r_{ij})

1.200	1.006	1.177
.911	.886	1.177
.939	.977	.778

Premiums (n_{ij}) (in millions of dollars)

320.4	54.2	37.5
677.0	62.7	100.2
194.3	17.1	40.9

Estimates: $\hat{\mu} = 1.000$ $\quad \hat{B}_1 = 0.993$
$\hat{A}_1 = 1.172$ $\quad \hat{B}_2 = 0.946$
$\hat{A}_2 = 0.941$ $\quad B_3 = 1.086$
$\hat{A}_3 = 0.915$

AB_{ij}

.035	-.112	-.081
-.023	-.001	.150
.031	.116	-.223

$\hat{\varepsilon}_1 = -3.4738$ $\qquad$ $\hat{\delta}_1 = -.0060$
$\hat{\varepsilon}_2 = 4.2795$ $\qquad$ $\hat{\delta}_2 = .0035$
$\hat{\varepsilon}_3 = -6.2127$ $\qquad$ $\hat{\delta}_3 = .0340$

$\hat{r}_{ij}$

	4-8	1,9	2,3	Construction Marginals	i
1	1.186	1.106	1.140	1.172	-3.4738
2,3	0.908	0.902	1.173	0.941	4.295
4-6	0.945	0.839	0.790	0.915	-6.2127
Protection Marginals	0.993	0.946	1.086	$\hat{\mu} = 1.000$	
$\hat{\delta}_j$	-.0060	.0035	.0340		

$F = 15.83$

The χ^2 for the above data is proportional to 1,016,570. Using a multiplicative model for the Bailey & Simon approach results in a χ^2 proportional to 5,994,534.

The F value ($F = 15.83$) is statistically significant even for this small data set. Therefore, the interaction term has remained, which is theoretically pleasing.

IV Three-Way Classification

The above two-way classification model can be extended to an n-way model, with perhaps reasonable complications in the mathematics. Presented here, are the assumptions and estimates for a reduced three-way Classification ANOVA model.

With a two-way model, the interaction term is fairly straight forward. In a three-way model there are available three, two-way interaction terms (pair-wise) and a single three-way interaction term in the complete specification of an ANOVA Model. Since our objective is to find "good" estimates of the relativities (r_{ijk}), we can imbed all of the different types of interactions into a single interaction term. This being the case, the model is as follows:

$$r_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta\gamma_{ijk} + e_{ijk} \tag{16}$$

with the constraints

$$\sum_i n_{ijk}\alpha_i = \sum_j n_{ijk}\beta_j = \sum_k n_{ijk}\gamma_k = \sum_i \alpha\beta_{ijk}\, n_{ijk} = \sum_j \alpha\beta\gamma_{ijk}\, n_{ijk} = \sum_k \alpha\beta\gamma_{ijk}\, n_{ijk} = 0$$

Using the least squares estimation technique here as in the two-way model, the following estimates result:

$$\hat{r}_{ijk} = \hat{A}_i + \hat{B}_j + \hat{C}_k - 2\hat{\mu} + \hat{\varepsilon}_i\hat{\delta}_j\hat{\phi}_k$$

$$\hat{A}_i = \frac{\sum_{jk} n_{ijk}\, r_{ijk}}{\sum_{jk} n_{ijk}} \qquad \hat{B}_j = \frac{\sum_{ik} n_{ijk}\, r_{ijk}}{\sum_{ik} n_{ijk}}$$

$$\hat{C}_k = \frac{\sum_{ij} n_{ijk}\, r_{ijk}}{\sum_{ij} n_{ijk}} \qquad \hat{\varepsilon}_i = \frac{\sum_{jk} n_{ijk}\, Z_{ijk}\, \delta_j\phi_k}{\sum_{jk} n_{ijk}\, \delta_j^2\, \phi_k^2}$$

$$\hat{\delta}_j = \frac{\sum_{ik} n_{ijk}\, Z_{ijk}\, \varepsilon_i\phi_k}{\sum_{ik} n_{ijk}\, \varepsilon_i^2\, \phi_k^2} \qquad \hat{\phi}_k = \frac{\sum_{ij} n_{ijk}\, Z_{ijk}\, \varepsilon_i\delta_j}{\sum_{ij} n_{ijk}\, Z_{ijk}\, \varepsilon_i^2\, \delta_j^2}$$

$$Z_{ijk} = r_{ijk} - \hat{A}_i - \hat{B}_j - \hat{C}_k + 2\mu$$

The above estimates, while not impossible, are more difficult to accomplish by hand than the two-way model, and the agony will increase with n, the number of dimensions. Conceptually an n-way model is a trivial extension from the above. I would advocate using only one interaction term for the same reasons as associated with the three-way model.

V Summary

The ANOVA model has shown in the Fire data its clear-cut superiority as measured by the χ^2 statistic of the Baily & Simon paper. With the Commercial Auto data, ANOVA was at least as good, but perhaps had a much better conceptual base.

In both examples, using the weights (n_{ij}'s) tends to mitigate credibility considerations, since the weights were chosen to be surrogate variables for the credibilities.

The ANOVA model is much more general, allowing for an n-way classification and with the test statistic a reasonable way to test the construction and formulation of the model.

The two-way model has been accepted by ISO, having been approved by the appropriate committees, and is currently being used by them for Construction-Protection relativities in Commercial Fire Insurance. It has been tested by an Ad-Hoc Subcommittee of ISO and, at least for Fire, has lived up to expectations.

APPENDIX

TEST OF SIGNIFANCE OF INTERACTION TERM

The basic model is

$$r_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ij} \qquad (1)$$

with $\sum_i n_{ij}\alpha_i = \sum_j n_{ij}\beta_j = \sum_i n_{ij}\alpha\beta_{ij} = \sum_j n_{ij}\alpha\beta_{ij} = 0$ as constraints.

We would like to make a statement about the significance of the interaction term ($\alpha\beta_{ij}$) that has been previously been estimated in the body of the paper. Therefore, let's try the following definition for $\alpha\beta_{ij}$:

$$\alpha\beta_{ij} = \lambda\epsilon_i\delta_j \qquad (2)$$

Under this formulation, we will attempt to test whether or not, statistically, that λ is equal to zero, thereby forcing our conclusion about the interaction term, $\alpha\beta_{ij}$.

Our ultimate objective is an F-test on the sum of squares of λ (SS_λ) which will be a comparison of that sum of squares to the remaining or balance sum of squares ($SS_{Balance}$), modified of course by their appropriate degrees of freedom (d.f.). These concepts and terms can easily be found in a standard statistical text dealing with Analysis of Variance.

Given Equation (2), we can now modify our model, with the appriate constraints of Eq. (1), to be

$$r_{ij} = \mu + \alpha_i + \beta_j + \lambda\varepsilon_i\delta_j + e_{ij} \tag{3}$$

To find a "least squares" estimate for λ, define Z to be the following

$$Z = \sum_{ij} n_{ij} (r_{ij} - \mu - \alpha_i - \beta_j - \lambda\varepsilon_i\delta_j)^2 \tag{4}$$

assuming α_i and β_j are known. We now differentiate Z and set it equal to zero and solve for λ.

$$\frac{dZ}{d\lambda} = -2 \sum_{ij} n_{ij} (r_{ij} - \mu - \alpha_i - \beta_j - \lambda\varepsilon_i\delta_j)(\varepsilon_i\delta_j) = 0$$

$$\hat{\lambda} = \frac{\sum_{ij} n_{ij}\varepsilon_i\delta_j (r_{ij} - \mu - \alpha_i - \beta_j)}{\sum_{ij} \varepsilon_i^2\delta_j^2 n_{ij}} \tag{5}$$

Now consider the sum of squares for the weighted interaction term, $\alpha\beta_{ij}$ using the definition from (2).

$$\sum_{ij} \alpha\beta_{ij}^2 n_{ij} = \lambda^2 \sum_{ij} \varepsilon_i^2\delta_j^2 n_{ij} \tag{6}$$

If we substitute our estimate of λ from (5) into (6) we have

$$\sum_{ij} n_{ij}\alpha\beta_{ij}^{2} = \frac{\left[\sum_{ij} n_{ij}c_i\delta_j (r_{ij} - \mu - \alpha_i - \beta_j)\right]^2}{\sum_{ij} n_{ij}c_i^2\delta_j^2} \tag{7}$$

Furthermore, since we do not know α_i, β_j, ε_i or δ_j, if we put their estimates into (7) we have

$$SS_\lambda = \sum_{ij} n_{ij}\alpha\beta_{ij}^{2} = \frac{\left[\sum_{ij} n_{ij}\hat{\varepsilon}_i\hat{\delta}_j AB_{ij}\right]^2}{\sum_{ij} n_{ij}\hat{\varepsilon}_i^2\hat{\delta}_j^2} \tag{8}$$

where the needed estimates are found in section II of the paper.

The final F statistic will take the form

$$\frac{SS_\lambda/\text{d.f.}}{SS_{\text{Balance}}/\text{d.f.}}$$

where the degrees of freedom (d.f.) for SS_λ are 1, and for SS_{Balance} is $\left[(p-1)\ (q-1) - 1\right]$. We need to define SS_{Balance} now.

What we are really doing here is looking at the residual sum of squares after fitting the additive terms and seeing if further, significant reduction can be made by fitting an interaction term. The residual sum of squares (SS_{Residual}) is

$$SS_{\text{Residual}} = \sum_{ij} n_{ij} (r_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2 \tag{9}$$

$$= \sum_{ij} n_{ij}AB_{ij}^{2}$$

If we reduce this by SS_λ we have the balance left unexplained or $SS_{Balance}$. Subtracting (8) from (9) yields

$$SS_{Balance} = \Sigma n_{ij} AB_{ij}^2 - \frac{\left[\sum_{ij} n_{ij} \hat{\varepsilon}_i \hat{\delta}_j AB_{ij}\right]^2}{\sum_{ij} n_{ij} \hat{\varepsilon}_i^2 \hat{\delta}_j^2}$$

$$= \frac{\left[\sum_{ij} n_{ij} \hat{\varepsilon}_i^2 \hat{\delta}_j^2\right]\left[\sum_{ij} n_{ij} AB_{ij}^2\right] - \left[\sum_{ij} n_{ij} \hat{\varepsilon}_i \hat{\delta}_j AB_{ij}\right]^2}{\Sigma n_{ij} \varepsilon_i^2 \delta_j^2}$$

Now dividing by the appropriate degrees of freedom and taking the ratio of SS_λ/d.f. to $SS_{Balance}$/d.f. gives us

$$F = \frac{\left[(p-1)\ (q-1) - 1\right] \left[\sum_{ij} n_{ij} \hat{\varepsilon}_i \hat{\delta}_j AB_{ij}\right]^2}{\left[\sum_{ij} n_{ij} AB_{ij}^2\right]\left[\sum_{ij} n_{ij} \hat{\varepsilon}_i^2 \hat{\delta}_j^2\right] - \left[\sum_{ij} n_{ij} \hat{\varepsilon}_i \hat{\delta}_j AB_{ij}\right]^2}$$

which has a standard F distribution and can be compared to any table of F-values in any standard statistical text.

RELATIVITY PRICING THROUGH ANALYSIS

OF VARIANCE

by CARL E. CHAMBERLAIN

Reviewed by DANIEL C. GODDARD

Mr. Chamberlain's paper is the first in years to address the problem of calculating class relativities for a two-way (or n-way) classification system. He proposes a new model that offers more flexibility (and more complexity) than previous ones. Essentially, his approach is to fit an additive model to the data, and then fit a multiplicative model to the residuals. As he explains, this approach is suggested by Analysis of Variance theory.

I have a few technical comments on Chamberlain's model. I also will discuss some problems with the model fitting approach in general, which suggest some areas for future research.

In fitting the additive model, Chamberlain minimizes

$$Z=\sum_{i,j} n_{ij}(r_{ij} - \hat{r}_{ij})^2$$

which is "the squared /absolute/ error with appropriate weights." For n_{ij}, he uses exposure in one case, and premium in another. Bailey and Simon, in "Two Studies in Automobile Insurance Ratemaking" (PCAS XLVII,1960) suggest minimizing

$$X^2 = \sum_{i,j} n_{ij}\hat{r}_{ij}\left(\frac{r_{ij} - \hat{r}_{ij}}{\hat{r}_{ij}}\right)^2$$

Here n_{ij} is exposure; so $n_{ij}\hat{r}_{ij}$ is proportional to the expected losses. Their formula amounts to the squared relative error, weighted by expected losses.

The choice of absolute or relative error is probably a toss-up. But I believe that expected losses are clearly the better choice for weights. Using just exposures can bias the results. In Chamberlain's auto example, Class 1A accounts for 66% of the exposure but only 52% of the losses.

Also, Bailey and Simon are minimizing Chi-squared, which is used to test how well the model fits. So, their approach will always do better than Chamberlain's on the Chi-squared test. Minimizing Z does have one practical advantage in the additive case, however. The equations can be solved explicitly, while minimizing X^2 requires an iterative approach.

An important purpose of the weights is to act as a "surrogate for credibility". Bailey and Simon explain why expected losses can be used to reflect the relative credibility of the squared error. This is based on what Insurance Services Office in their research calls classical credibility (characterized by a fixed standard for full credibility, and a square root formula for partial credibility).

However, I question whether this weighting really does replace credibility. Consider what happens if we apply any of the models - additive, multiplicative, or Chamberlain's - to a one-way classification system. Consider Bailey and Simon's equations (6) or (9), or Chamberlain's (14) for the case where j has only one value (i.e. the data has only one column). They all reduce to $\hat{r}_i = r_i$. There is no use of credibility left here. The accepted solution for the one-way problem is $\hat{r}_i = Zr_i + (1-Z)\ r$, where Z is the credibility, and r is the value for a larger group. So, our models for n-way class relativities do not work for the special case of n=1. It appears that the weighting by expected losses functions as credibility only to the extent that there are interactions between the dimensions. We are not really using the weighting as a surrogate for credibility; the surrogate is actually the structure of the model which defines the interactions we will consider. In other words, we decide on a model and data is "credible" to the extent that it fits the model.

To get a further sense of what is happening, let us look at the four criteria Bailey and Simon give for an acceptable set of relativities:

Criterion 1. It should reproduce the experience for each class and merit rating class and also the overall experience; i.e., be balanced for each class and in total.

Criterion 2. It should reflect the relative credibility of the various groups involved.

Criterion 3. It should provide a minimal amount of departure from the raw data for the maximum number of people.

Criterion 4. It should produce a rate for each subgroup of risks which is close enough to the experience so that the differences could reasonably be caused by chance.

In the one-way case, balance is taken care of with a balancing or "test correction" factor; it is not a consideration in calculating the relativities. In the two-way case, balance for each class is desirable to insure the model structure is reasonable, but it assumes each class is fully credible in total.

Where criterion 2 calls for reflecting the relative credibility of the groups, criterion 4 in effect calls for reflecting the absolute credibility of each class. These are familiar criteria in the one-way case; in fact, they are the only ones used.

Criterion 3 is a test of how well the model fits. It is irrelevant in the one way case, because we assume no model. So, as we go from the one-way case to the two-way, we add an important and fundamental assumption: there is some rational structure to the interactions between the classes. Do we need this assumption? Should we make it? Presumably, we are trying to get the best estimate of each r_{ij}. So why not calculate the classical credibility Z_{ij} for cell ij, and then

$$\widehat{r}_{ij} = Z_{ij} r_{ij} + (1 - Z_{ij}) r$$

There are at least three problems with this formula. First, we need a standard for full credibility. This has been discussed extensively elsewhere.

Second, what do we use for r? If we set $r=r_{\cdot\cdot}$ (using Bailey and Simon's notation) we are ignoring the information we have about other risks in row i or column j. We have reduced the problem to a one-way classification scheme. Another choice is to use some combination or average of $r_{i.}$ and $r_{.j}$. This is what the NCCI does with the national relativity and the pure premium on level in computing their class relativities.

Third, the classical credibility is independent of our choice of r. This is one of the chief arguments for Bayesian credibility. However, Bayesian credibility has only been developed for the case where each class is a member of only one group. That is, it is just for a one-way classification scheme. ISO has done considerable work on how to group classes where there are several different criteria that could be used. They have suggested using multi-dimensional scaling to reduce all the criteria to a one-way scheme.

It appears what we need is a multi-dimensional credibility theory (which I will leave to more mathematical actuaries than me to develop). Such a theory would solve ISO's grouping problem; it would give us the best estimate for each class relativity; and it would avoid having to guess at an appropriate structure.

There are practical problems with such an approach. A set of relativities so calculated would have to be published as a table, not as a few parameters and a formula. In some cases this is no hardship. In Chamberlain's property example, he used twelve parameters and a fairly complicated formula to fit a table of nine numbers. Other cases are not so easy. For example, multi-dimensional credibility would give us a different set of class factors for each auto rate territory. Apart from adding many pages to the rating manual, such a change would require major changes in most automated systems. So, we would probably want to select one or two sets of class factors that are "close enough" for most territories. In other words, we would fit a model to simplify the structure. Having already credibility weighted the data, we can attribute any residual error to the choice of the model. The present procedure cannot distinguish between errors due to the choice of model, and errors due to statistical fluctuations in the data.

This leads to a new perspective on Chamberlain's model. We start by assuming that there is a pattern to the relativities, and our estimates should reflect as much of this pattern as possible. So, we start by fitting a first-order additive model. We then test whether a second-order multiplicative model shows any significant remaining pattern in the residuals. In theory, we could go on fitting higher order models, until the F value is no longer significant. In practice, with the size of the data sets usually encountered, I would expect two stages to be sufficient. This procedure should extract the maximum possible pattern from the data, in the same sense that polynomial stepwise regression does for a time series. Just as with polynomial stepwise regression, Chamberlain's procedure does not necessarily give the simplest or most efficient model. And the fact that it detects a pattern is no assurance that the pattern is reasonable.

This brings me back to my main point. Our procedures for two-way relativities are based on a very different point of view than those for one-way relativities. I believe we need a multi-dimensional credibility theory to reconcile the two.

TITLE: EXPENSE ALLOCATION IN INSURANCE RATEMAKING

AUTHORS: Ms. Diana Childs

Ms. Childs is an Actuary with the Insurance Company of North America. She received her BS degree in Math from Penn State University. She received her FCAS in 1977, is a member of the American Academy of Actuaries and served the CAS on the Career Enhancement Committee.

Mr. Ross A. Currie

Mr. Currie is currently Associate Actuary with the Insurance Company of North America. He received his ACAS designation in 1978.

REVIEWER: Mr. David Klein

David Klein is currently Vice President and Director of Actuarial/Research Operations for The Hartford Insurance Group. He received his FCAS in 1974 and holds an MBA in Insurance from Temple University. David is on the CAS Finance Committee and served on the Examination Committee. Currently, Chairman of the ISO Actuarial Committee, he has served on a number of industry committees including the AIRAC Personal Lines Committee and the Executive Committee of the NIC/AIPSO.

Until the present time, the great majority of actuarial study and literature in the ratemaking area has revolved around analyzing and quantifying the loss component of the insurance rate. Actuaries have evolved an elaborate system in which losses are trended, developed and credibility weighted, and in which premiums are placed at current rates or at least current rate levels. At the same time, actuaries have virtually ignored the expense portion of the insurance rate, preferring to treat expenses as a constant percentage of premium. Current economic and political conditions are forcing a reevaluation of this simplistic approach towards expense allocation. Consumer groups have charged that current expense allocation procedures are discriminatory, and insurance companies are attempting to improve their pricing position through the development of rates which more accurately distribute the costs of doing business.

In this paper we shall take a look at the expense portion of the insurance rate. We shall examine the pros and cons of the traditional treatment of expenses and shall consider some alternate methodologies. Our focus will be on the personal lines; Automobile and Homeowners. At the same time, many of our conclusions and observations can be extended to other lines of insurance.

This paper does not attempt to answer all of the questions regarding expense allocation. Rather, its intention is to lay a general foundation upon which specific, detailed expense flattening procedures can be built.

Proportional Allocation vs. Expense Flattening

The traditional approach towards treating expenses in ratemaking is what we shall refer to as proportional allocation. Under this approach all underwriting expenses are considered to vary absolutely with the premium rate.

Given: R_n = the rate for a risk of a specific class = n

L_n = the underlying pure premium (including all loss expense)

E_n = the provision for underwriting expenses in R_n

Then: $R_n = L_n + E_n$ (1)

And: E_n/R_n is assumed to be constant for all n

We shall define expense flattening to be any allocation procedure in which some or all of the underwriting expense provision is considered to be independent of the final rate. Specifically, most proposed rate structures which incorporate expense flattening can be defined as follows:

$$R'_n = L_n + e_n + e \qquad (2)$$

e_n = the variable expense provision (i.e., e_n/R_n is a constant for all n)

e = a flat expense loading which is constant for all n

Of course this form is only a single representative in a wide spectrum of possibilities. There is no reason to assume that all expenses which do not vary by premium should be loaded as a flat charge by exposure. It is quite conceivable, for example, that many underwriting costs will vary by territory but remain constant for other classifications.

In their most complex form, the rates resulting from a flat allocation system would look something like this.

$$R'_n = L_n + e_n + e_1 + e_2 + \ldots + e_x + e \qquad (3)$$

$e_1, e_2, \ldots e_x$ = expense loadings which vary according to some identifiable characteristic

Each of the subscripted e's represents an expense component which may vary on a risk by risk basis. For example, if certain overhead costs were found to be twice as large in one territory as in another, those costs might be assigned to variable e_1 which would be defined as follows:

$e_1 = f(t) \times 0$

0 = the overhead loading

t = territory

$f(t_1) = 1$

$f(t_2) = 2$

Notice that, in theory at least, e_1, $e_2,\ldots e_x$ do not have to be categorized along the same lines as the pure premium. Different territory definitions may be employed and entirely different classes may be recognized. Even in cases where the various e's change in accordance with normal rating classifications, their relativities (defined as f(t) above) need not be identical to the pure premium relativities. In the above case, for example, the pure premium rate for territory 2 is not necessarily twice that of territory 1.

Practical considerations will, of course, limit the application of this very general formula. The introduction of brand new expense classifications would represent a data processing nightmare. While intuitive judgment may influence estimates of the relative cost of writing different classifications, precise quantification will often be difficult, if not impossible. In these cases the actuary may have to rely to a great extent on pure premium relativities or may be forced to ignore the existence of the differentials entirely. Legal restrictions will also be placed on the allocation of expense dollars and social implications will play as large a role as economic considerations in determining the exact allocation formula for a specific line of business.

It must always be remembered that regardless of the final formula chosen for loading expenses, we are reallocating, not reevaluating, our expense costs. Decreases in one insured's rate due to reallocation will be offset by increases elsewhere. Redefining the expense allocation procedure is not a remedy for the high costs of insurance. This may seem obvious to the actuary, but in the public forum it is often swept under the rug in the desire to lower rates for a specific classification. The public, loudly calling for a change in the ratemaking methodology, is almost certainly doing so with the misconceived idea that this change will save them money. In fact, the insurance companies will be collecting the same total expense dollars but will be requiring a few insureds to pay a significantly lesser amount while the large majority of insureds will pay a little more. There is no such thing as a free lunch, and in examining the reasons for adopting any expense allocation procedure it has to be kept in mind that changes which will benefit some groups will consequently penalize others.

Reasons for Expense Flattening

Prior to the mid-seventies, rate changes for the personal lines of insurance were relatively infrequent and represented modest increases to account for a modest inflation rate. While rates did vary by classification and territory, the overall level of the insurance premium represented a necessary but affordable item in the household budget. Pronounced differences of territory and classification rates did not exist and overall rate levels and increases were kept to a minimal level. In practical terms, varying the expense loading with premiums certainly simplified policy processing and ratemaking procedures, and as long as pure premium adjustments reflected inflation, the collected expense dollars also increased appropriately. Additionally, since most expenses did vary directly with premiums (commissions of 20-25% were not unheard of and represented the largest component of the expense loading), there seemed little point in devising a more complicated way of reflecting expenses in the premium dollar.

In recent years Homeowners insurance rates have remained at fairly stable levels; however, the unaffordability and lack of availability of private passenger automobile insurance has reached crisis proportions. With rates skyrocketing and consumerism in vogue, the social acceptability and equity of current ratemaking techniques have come under fire. Many aspects of the insurance mechanism are being questioned, whether it is rating by age, sex, marital status or geographical location, and the expense loading methodology is a ready target for change.

The reasons for questioning the current expense allocation procedure come under two guises; social acceptability and financial equity. The primary impetus for expense flattening has come from groups outside of the industry which maintain that it is not "just" or "fair" to assign different expense charges to risks merely because of expected loss differences. It is argued that the inequities inherent in a proportional allocation system have contributed significantly to the affordability crisis. It must be recognized, however, that the expense dollars currently subject to flattening represent a relatively small portion of the overall premium and their reallocation will not solve the affordability problem. In addition, although expense flattening will obviously benefit urban and youthful motor vehicle operators, it is a two edged sword which cuts the other way when applied to Homeowners insurance. The rich, surburban home owner will actually save money with the application of a flat expense costing technique, while the urban row home owner will be penalized. For this latter reason consumer groups understandably neglect to call for similar reforms in the pricing of

Homeowners insurance. Thus while the ostensible justification of expense flattening is a moral one, social activists actually advocate it only when it reduces the cost of insurance to certain selected economic groups. There is considerable danger in pricing an insurance product in response to social objectives, for unless there is some financial justification for revising pricing procedures great harm will be done to the industry and, ultimately, to the consumer.

Fortunately, expense flattening can be justified for financial reasons. From a pure equity standpoint, insurers would like rates to accurately reflect the costs of issuing a policy. If one vehicle's pure premium is three times that of another, does that also imply a threefold difference in incurred expenses? Matching expenses to policies as expenses are incurred provides not only a more accurate pricing mechanism but guards against the loss of collected expense dollars due to shifts in the mix of business, particularly within the territorial and classification distribution.

Expense Categories Subject to Flattening

The expenses associated with issuing and servicing an insurance policy can be segregated into loss adjustment expenses (those expenses incurred to investigate, litigate, and settle claims) and underwriting expenses (those expenses incurred while issuing the policy).

This paper is concerned with the allocation of underwriting, as opposed to loss adjustment expenses; however, loss adjustment expense is equally susceptible to a flattening procedure. Generally speaking, loss adjustment expense is considered to vary directly with dollars of loss. It seems obvious that claim count also influences loss expense cost. The possibility of loading claims expense into the rate as a composite factor of frequency and severity or other alternatives leaves a wide variety of possibilities which are open to future actuarial study.

While a good deal can be said for varying the loss expense loading as a function of loss, the opposite is true of the current rationale for including underwriting expenses in the rating structure as a function of pure premium.

Commisssions, the expense dollars paid to the agents for their efforts in underwriting, placing, issuing and servicing the policy, have historically been determined as a fixed percentage of the final premium. Interestingly, a modification of this approach has been avoided by proponents of expense flattening; however, the reasons for the omission may be related to those groups' unwillingness to oppose the various independent agents' associations and not belief in the equity of the current system. Differences in costs among territories (rural vs. urban) and in placing insurance for certain less desirable insureds certainly justify part of the commission differential inherent in the current rating structure. Still, the possibilities of flat commissions or a graduated scale of commission rates deserve further research to determine an appropriate cost accounting charge. It is the opinion of the authors that some form of flat charge plus percentage of premium provides more equity among insureds and also provides incentive to the agent to place difficult risks.

State premium taxes are levied against each company as a function of the direct premium writings for a given line and state. The cost is passed on to the consumer in the same manner as the charge is levied on the company. This procedure is the only way a company can ensure that it will collect exactly the dollars which the state will require as payment. While this charge amounts to an average of only 2-3% of premium, the expense allocation issue offers the opportunity for the states to study their procedures in assessing premium taxes on the insurance companies and in turn the insured.

Other Acquisition Expenses represent the insurance company's costs (ex commission) to issue a policy. Included in this area are advertising fees, computerized rating and policy issuance systems, postage and telephone charges, travel expenses, salaries, and other miscellaneous items. The General Expense category includes salaries, rents, equipment, boards, bureaus and association fees, and other overhead items in an insurance company's budget. Historically, they have averaged 10-12% of the premium dollar.

These two categories, Other Acquisition Expense and General Expense, are the most susceptible to an alternative form of expense allocation. The basic question regarding these items seems to be "Do any two risks with differing pure premiums also cost differing amounts from an administrative expense standpoint?" This question may be posed of two Homeowners policies - a $100,000 single home in an affluent suburban area and a $15,000 row home in an inner city area. Both of these policies utilize identical computer routines to rate the policy, require the same paper to be processed for policy issuance, and take up computer space to record the policy in the company's data system, yet the current premium charges reflect different amounts to pay for these items. The pricing of Automobile policies follows the same pattern, with the higher priced risks paying a large share of the company's expenses. An age 17 unmarried principal male operator with a rating factor of 3.50 is also paying 3 1/2 times the dollar amount of an over 30 male operator for the general expenses of an insurance company.

The answer to this question is both yes and no. Unfortunately, there is no clear-cut solution, and in fact each company must examine its own policy issuing systems, rating procedures, and other associated overhead expenses to determine which costs are variable and which are fixed.

Many of the areas influencing the cost of writing a policy are subject to judgment and intuition, and any company studying the expense flattening issue must compromise between accuracy and practicality. A recent study of expenses for Private Passenger Automobile Insurance completed by the Insurance Services Office concluded that 75% of the Other Acquisition and General Expenses and Miscellaneous Taxes are fixed while the remaining 25% are variable. It was further recommended that per car fixed expense loadings be developed by state and coverage. Companies can be guided by these conclusions but should study their own circumstances to determine the appropriateness of the application of this study to their individual situation.

While Profit and Contingencies is construed as an item of expense in the insurance rate, the proper allocation of profit to an insured is a difficult and complicated issue to resolve. A study of the concept of risk and its application to territory, classification, limit and other rating criteria is required before a proper determination of the apportionment of the profit and contingency charge can be made. The authors feel that such a study is beyond the scope of this paper and, in fact, is wide enough in scope to be the sole topic of a paper on the subject. With respect to this treatise, we will continue to treat the profit and contingency factor as a variable loading in the insurance rate.

Once those expense categories which will be subject to flattening have been selected, the actual flat expense charge is arrived at in a straightforward manner. The flat expense charge per exposure will be determined by dividing the total variable expenses now subject to flattening by the appropriate exposure base.

An Example

A company markets an insurance product which is priced in accordance with standard, proportionally allocated expenses. The business is segregated into two classes and the following data applies.

	Variable	Class 1	Class 2	Total
Exposures	X	5,000	5,000	10,000
Rate	R	\$50	\$150	
Premium	P	\$250,000	\$750,000	\$1,000,000
Permissible Loss Ratio	PLR			60.0%
Allowance for:				
Gen. Expense & Other Acq.	C1			13.3%
Remaining Underwriting Expense	C2			26.7%
Total Underwriting Expense	C			40.0%
Actual Loss & Loss Expense	L'			\$650,000

Given the above situation, the appropriate flat expense charge equals \$10 and is calculated as follows. (This example assumes that 75% of all General and Other Acquisition expense is subject to flattening.)

$$e = (.75 \times C1 \times P)/X = (.75 \times .133 \times \$1{,}000{,}000)/10{,}000 = \$10$$

If we are revising rates as well as incorporating flat expenses then some sort of expense trending might also be appropriate. For purposes of illustration, we will assume that e', the trended expense charge = $12.

At the same time a revised variable expense provision is calculated.

$$C' = (.25 \times C1) + C2 = 30\%$$

<u>Basic Ratemaking Techniques Using Flat Expense Allocation</u>

We have separated those components of the expense loading which will be assigned on a proportional basis from those which will be charged using some type of flattening procedure. The problem which now remains is to build our revised allocations into the ratemaking process.

In order to simplify our presentation, we shall assume that the rating formula used for flattening expenses follows the basic form:

$$R'_n = L_n + e_n + e \qquad (2)$$

Our formulas can, however, be adapted to accept the more complex form shown in formula (3).

Under the traditional approach of proportional allocation, the rates for a line of business are developed in accordance with formula (1). In the initial stage of a conversion to flat expenses, we wish to convert the individual classification rates to a formula (2) form without revising the underlying pure premiums or overall collected expenses. This is accomplished in a three step procedure.

1) Calculate the pure premium underlying present rates

2) Add the new flat expense provision

3) Load the total for the remaining variable expenses.

Using our previous example as an illustration, we develop new rates of \$57 and \$143 for classifications 1 and 2, respectively.

Pure Premium for class 1 = \$50 x .6 = \$30

Pure Premium for class 2 = \$150 x .6 = \$90

$R_1' = (\$30 + \$10)/.70 = \$57$

$R_2' = (\$90 + \$10)/.70 = \$143$

In general terms, the revised rate (R') for class n is calculated using the following formula:

$$R'_n = ((1-C) \times R_n + e)/(1-C') \qquad (4)$$

Note that $e/(1-C')$ is a constant. Therefore, if we so choose, we can present the revised rate in terms of a multiplier to the current rate plus a constant term.

$$R'_n = KR_n + h \tag{5}$$

$$K = (1-C)/(1-C') = .857 \tag{6}$$

$$h = e/(1-C') = \$14 \tag{7}$$

Essentially, formula (5) defines a rate which includes a provision for flat expenses as a combination of a loss rate and an expense rate, where each of these component parts includes a loading for variable expenses.

Once the rates have been modified to incorporate flat expenses, the next area of concern to the ratemaker is the calculation of rate level adjustments due to changing experience. Estimates of overall rate level need are calculated in a manner almost identical to that used when a proportional allocation system is in place. The familiar method of adjusting overall rate levels is:

I = LR/PLR = indicated rate level change

LR = experience loss ratio adjusted to current rate and prospective loss levels

PLR= permissible loss ratio

To accomodate rates which incorporate flat expenses, we replace the loss ratio term in the formula with a loss and flat expense ratio, and we replace the permissible loss ratio with a permissible loss and flat expense ratio.

$$I = ((L' + e'X)/P) / (1-C') \qquad (8)$$

L' = losses developed and adjusted to prospective levels

e' = trended flat expense dollars per risk

X = number of exposures

P = total premium at current rates

Again using our previous example:

L' = \$650,000

e' = \$12

X = 10,000

P = \$1,000,000

C' = .30

$$I = ((\$650{,}000 + \$120{,}000)/\$1{,}000{,}000) / .70 = 1.10$$

This calculated indication represents the necessary increase in rate level in order to maintain premium adequacy. Its primary use is one of measuring overall profitability and financial position, for unlike the indication developed under a proportional allocation system this modification cannot be applied directly to the individual classification rates. In order to maintain independence between the loss and expense portions of the final rates, two separate adjustments must be calculated.

$$M_L = \text{overall loss rate modification} = (L'/P_L)/(1-C') \qquad (9)$$

$$P_L = \text{premiums less fixed expenses} = P-hX$$

$$M_e = \text{overall expense rate modification} = e'/e$$

In our example:

$$P_L = \$1,000,000 - \$14\ (10,000) = \$860,000$$

$$M_L = (\$650,000/\$860,000)\ /.70 = 1.080$$

$$M_e = \$12/\$10 = 1.200$$

If revised rates are to be based on overall rather than classification loss experience then these factors can be applied directly to the individual loss and expense rates. If, however, class experience has some degree of credibility then the modification of the loss rates can be adjusted accordingly.

$$M_{L_n} = ((L'_n/P_{L_n})\ Z_n + (L'/P_L)(1-Z_n))/(1-C') \qquad (10)$$

Note that the individual class modifications must receive a subsequent adjustment to achieve the proper overall change.

This revised approach towards expense allocation will not affect credibility levels in Automobile or Homeowners insurance where credibility is based on claim counts and exposures. However, if flat expenses are introduced in a line where premium has been used as a credibility measure, then some revision in credibility values should be considered.

We have now developed a basic approach towards making rates when a system of expense flattening is used; however, before moving on it is necessary to briefly mention some of the practical considerations with which we must deal when using this system.

Separate loss premium and expense premium information must be available to the ratemaker. Accurate exposure data is also necessary in order to properly evaluate the magnitude of the flat expense loadings.

Flat expense costs will obviously be subject to inflation, and expense trending procedures must be developed. In many cases loss trend is being applied to expenses as an interim measure. It is obvious, though, that in most cases loss trend is not an appropriate measure of increasing expense costs. Automobile crash parts, liability judgments, and medical costs are certainly rising faster than general expenses. In the case of Homeowners insurance, trend factors are tied in to construction cost indices. The authors feel that if expenses are subject to the application of trend factors, then these factors should be developed using CPI-type wage and price indices which correspond, however roughly, with those costs which underlie an insurance operation.

The question of whether flat expenses should be allocated on a statewide, regional or countrywide basis must also be addressed. Costs associated with the operations of field offices will be influenced by local economic conditions, while EDP and Home Office operation expense could reasonably be assigned on a countrywide basis. At the moment, flat expenses by state are determined by applying existing, variable, loadings to state premiums. This methodology assumes that while certain expenses are flat within a state, on an interstate basis expenses continue to be a function of premium dollars. While this approach may not appear to be valid from a theoretical standpoint its use must be continued as long as state regulations vary with regards to expense flattening categories and procedures, for it is the only way to insure the collection of adequate expense dollars.

The implementation of an expense flattening procedure also introduces a new element when filing for rate changes which do not equal the required overall indication. It is often the case, for example, that for political or marketing reasons Automobile rate changes fall a great deal below indications. Should rate shortfalls be absorbed solely by the loss portion of the rate? Or should it fall on the expense rate as well? If these costs are to be shared it must be decided whether the division will be proportional or according to some other standard.

These and other general problems must be handled if a ratemaking system using flat expenses is to be effective. Of course, each line of business also has special considerations which must be addressed.

Expense Considerations in Automobile Insurance

Since the main thrust of expense flattening has been aimed at Private Passenger Automobile insurance, various methodologies have already been investigated, documented, and implemented in a few states. As we mentioned previously, the Insurance Services Office prepared a study of expenses concluding that 75% of company General Expenses and Miscellaneous Taxes, Licenses, and Fees are fixed. This result was incorporated into an expense flattening program implemented in rate revisions filed after January 1, 1979. The ISO has chosen to develop expense fees by coverage (based on the average expense loading currently in the rate) and by state, with the fees applicable on a per car basis.

Several questions arise from this proposal which each insurer should investigate. Specifically, each company needs to determine if the flat expense fee should apply per car or per policy and whether or not different charges are required by coverage. Should the same expense fees apply to renewal as well as new business and should any charge be made for mid-term endorsement activity?

Frequently companies issue one policy to insure multiple exposures and, while there is some additional expense incurred in the rating and processing of a multi-car risk, it is not likely to be proportional to that of a single car risk. A thorough investigation of the billing and policy issuance systems should provide sufficient information to determine the extent of the cost savings which results when writing a multi-car policy. In general, expense savings can be found in the areas of postage, paper, telephone and telegraph costs, and related processing expenses. Most companies issue a single policy for a multi-car risk, thereby reducing processing costs. At the same time, however, installment premium payment modes may be more prevalent with a multi-car risk, thus offsetting the savings obtained from the single policy issuance. Each company needs to determine if a significant cost difference exists in the issuance of a single vs. a multi-car policy and the feasibility of implementing a per policy charge within its systems capabilities.

Another aspect of ISO's expense flattening proposal requires comment. The ISO technique develops flat expense fees by coverage based on the variable loading currently in the coverage rate. As an example, consider a state in which the current average rates by coverage are:

Liability	$225
Comprehensive	$ 50
Collision	$125

If it is determined that 10% of the rate represents fixed expense, then flat expense fees of $22.50, $5.00 and $12.50 would apply to the respective coverages. A policyholder with a Liability-only policy would pay $22.50 in expenses; one with Liability and Comprehensive, $27.50; and one with the full complement of Liability and Physical Damage, $40. The actual expense differential between a Liability and Comprehensive policy and one which also includes Collision coverage is minimal since most of the rating information is already available in the data base. The largest expense is incurred in adding the first Physical Damage coverage. The above approach defeats the purpose of matching actual expenses to policies and instead allocates fixed expenses to coverage on the basis of pure premium. This is just as arbitrary as the current ratemaking procedure. A reasonable alternative is to determine a basic expense fee to be charged on the initial coverage written and a second, smaller fee if any secondary coverages apply.

An investigation of the costs to issue a new vs. a renewal policy or to add an endorsement will likewise determine if it is worthwhile to distinguish the expenses incurred by these transactions.

In all of these areas, the issue of expense flattening requires a complete re-examination of the costs to issue a policy and a re-thinking of the insurance industry's approach to charging for them.

Expense Considerations in Homeowners Insurance

The use of expense flattening in Homeowners insurance can create some problems if its application is not carefully planned. Three areas which should receive the actuary's attention are:

1) The impact of expense flattening on the pricing of policies which provide for automatic increases in face amount.
2) The assignment of expenses to Tenants insurance.
3) The pricing of endorsements.

In recent years insurance companies have attempted to offset the effects of inflation on Homeowners business by including what are referred to as "inflation guard" provisions in the standard policy. These provisions provide for a periodic, automatic increase in the policy face amount. This increase will, of course, result in a premium increase for the insured without the necessity of a rate revision. Under a ratemaking system which uses proportional allocation of expenses, this pricing mechanism will result in the collection of increased expense dollars along with the pure premium increase. This effect is lost. however, for any flattened expenses. As long as there is no revision in rates, insureds will continue to contribute the same flat expense premium regardless of the face value of the policy. The use of flat expenses will therefore necessitate an increase in the number and amount of rate level adjustments as inflation acts upon the flattened portion of the rate. To some extent the use of flat expenses will nullify the principle objective of an "inflation guard" system; an increase in collected premium without all of the problems inherent in filing and implementing a rate level change.

The expenses inherent in carrying a book of Tenants business are generally inseparable from those accompanying the Homeowners forms. The same processing and billing systems are used, and to a great extent the entire Tenants product is treated as another Homeowners form. This implies that the flat expense charge for a piece of Tenants business should be the same as that of a Homeowners policy. The implementation of identical charges creates a practical problem in that Tenants policyholders will often receive substantial rate increases as a result. This can be illustrated by looking at one company's experience for a single representative state.

	Homeowners	Tenants	Combined
1978 Earned Premium (000)	$4,778	$417	$5,194
1978 Earned Exposures	24,588	4,990	29,578
Average Premium	$194	$83	$176

If 10% of the total premium is subject to flattening, then the per policy charge based on combined experience equals $18. Thus, the new average premium for a Homeowners policy would equal $193 ($194 x .9 + $18) while the average Tenants policy would increase by 12% to $93. This Tenants increase would fall even more severely on renters with low contents values. In the above case, more than 50% of the exposures carried contents coverage of $8,000 or less at an average premium of $62. For these insureds, the average increase resulting from expense flattening exceeds 19%.

The use of identical expense charges for Homeowners and Tenants business is impractical from both social and marketing standpoints regardless of the financial equity of the system. A possible solution to flattening expenses for Tenants insurance is to adopt a separate charge even though Tenants expenses cannot be segregated from those of the Homeowners forms. In the above case, flat expense charges of $19 and $8 could be adopted for Homeowners and Tenants business, respectively.

The possibility of flattening expenses incurred when adding endorsements to a Homeowners policy must also be considered. As was true in the case of pricing a Tenants policy, the dollar impact of any change in the allocation of expenses is of as much concern as the equity of the pricing method. In many cases it will be concluded that the pricing of endorsements is best left unchanged due to the small costs involved, but this decision should be a conscious one which is made only after evaluating each particular situation with care.

Conclusion

Expense flattening is a subject which demands more attention than the actuarial profession has devoted to it. It continues to be a subject for public debate, and consumer pressure is rapidly forcing the adoption of legally mandated flattening procedures. Unfortunately, these procedures are often convoluted and lack firm statistical justification. Expense flattening is here to stay, and a continued lack of actuarial input will only insure the continued adoption of inconsistent, unjustifiable flattening schemes.

In this paper we have presented the basic concepts underlying the proper allocation of underwriting expenses, and we hope that it will open the door to further research in this area. Hopefully, by answering some of the questions and correcting the misconceptions which surround expense flattening we will serve both the industry and the public by helping to provide insurance products whose prices accurately reflect their associated costs.

Glossary of Variables

R_n = the rate for a risk of a specific class n

L_n = the underlying pure premium in R_n

E_n = the provision for underwriting expenses in R_n

R'_n = revised rate for class n after expense flattening

e_n = the variable expense provision in R'_n

e = the flat expense loading which is constant for all n

$e_1, e_2, \ldots e_x$ = expense loadings which vary according to some characteristic

O = overhead loading

t = territory

X = exposures

P = premium

PLR = permissible loss ratio

C1 = gen. expense and other acq. allowance

C2 = remaining underwriting expense allowance

C = total underwriting expense

L' = actual loss and loss expense

e' = trended flat expense charge

C' = revised variable expense charge

K = multiplier = (1 - C)/(1 - C')

h = constant = e/(1 - C')

Glossary of Variables (Cont)

I = indicated rate level change

LR = experience loss ratio at current rates and prospective loss levels

M_L = overall loss rate modification

P_L = premiums less fixed expenses

M_e = overall expense rate modification

M_{L_n} = loss rate modification for class n

L'_n = actual loss and loss expense for class n

P_{L_n} = premiums less fixed expenses for class n

Z_n = credibility assigned to class n

EXPENSE ALLOCATION IN INSURANCE RATEMAKING

BY DIANA CHILDS AND ROSS A. CURRIE

REVIEWED BY DAVID KLEIN

The authors are to be commended for their willingness to address as controversial a subject as expense allocation. Their approach provides one with a basic introduction to the subject. This reviewer, with a limited experience with the subject, feels that a few general comments are in order.

For sometime actuaries have recognized the necessity and appropriateness of expense flattening. In Workers' Compensation, the practice of expense graduation has been in place for many years. In the larger commercial lines, the issue of expense requirement has been implicitly or explicitly dealt with through large account programs such as IRPM's and Commission Contribution. It has been the staggering increases in Personal Lines premiums over the past five years which has brought the issue to the fore in this arena.

One major issue suggested by the author in their general discussion is the accuracy of the expense data upon which our fixed and variable allocations will be made. Actuaries familiar with the vagueness of New York Regulation 30, which forms the basis for much of property and casualty insurance accounting, will shudder when they think of the potential uses being made of data collected under those guidelines. Put another way, how good is the base we are allocating? As we move into an era where more refined treatment of expense provisions is required, we should not only focus on the redistribution of expenses, but also whether we must consider the costs have been properly assigned in the first place. Is Regulation 30 adequate in its definition? How good are company internal procedures? These issues must be resolved if expense allocation is to have any degree of credibility.

Another major issue surfaced by the authors is the role of critics of the insurance industrys expense allocation process. Industry critics tend to concentrate on the flattening of the higher premiums payers without proper recognition of the impact of this approach on those at the lower end of the spectrum. The authors appropriately reference to the "two edge sword" aspect of expense flattening, specifically when applied in Homeowners insurance and the adverse effect that it may have on the tenants forms, particularly, the lower valued forms. Unfortunately, the authors suggest an approach to deal with this "problem" which violates the basic principal of cost based pricing by artificially lowering the flat (affordable) expense charge for the tenants forms to make it socially acceptable. This is an apparent contradiction to an earlier section of the paper where they caution that "there is considerable danger in pricing an insurance product in response to social objectives."

The area of commission and taxes is one where this reviewer and the authors are in agreement. Together these two items represent an amount equal to, if not greater than, the expense dollars which would be flattened under the ISO approach described in the paper (page 52).

The points raised by the authors relative to the compensation of the producer force bear careful study by agents' associations if they are to avoid having the matter decided for them without their input. Regarding the flattening of premium taxes, the authors correctly point out that it is up to the states to take the lead in revising their procedures if anything is to be accomplished in these areas. Finally, this reviewer agrees with the authors that "the proper allocation of profit to an insured is a difficult and complicated issue" which "is wide enough in scope to be the sole topic of a paper" (page 46).

This reviewer was disappointed by the authors failure to discuss the impact of expense flattening on loss ratios. Non-actuaries rely on loss ratios to evaluate results and often actuaries have to use or explain loss ratio data. With the use of expense flattening, loss ratios will become less meaningful unless properly interpreted to reflect the impact of the revised expense allocation procedure. The authors failure to mention this point is hopefully a matter of oversight and not for lack of recognition of its importance.

The above comments are intended to be a general review of the major points of the paper. On a somewhat more technical basis, on page 47, formula (7) indicated a fixed expense portion which is loaded for variable expenses.

$$(7) \qquad h = \frac{e}{(1 - c')}$$

This implicitly increases the flattening and in the mind of this reviewer, is unnecessary. As a corollary, it makes the ratemaking process unnecessarily complicated.

An alternative procedure would be to make the denominator of equation (4), a function of only taxes, commission and profit, and to make variable expenses a function of pure loss. Under this suggested approach

$$R'_N = \frac{Rn \times (1 - C_2 - .25C_1)}{1 - C_2} + \frac{e}{1 - C_2}$$

This formula uses the factor $(1 - C_2 - .25\ C)$ to generate the loss, loss adjustment expense, and variable expense portion of the rate. The constant expense portion of the rate is shown separately, and both are loaded with those elements which are variable with <u>premium</u>. This makes the constant term independent of the variable portion of expenses.

Using the above suggested approach in the content of ratemaking simplifies the process. We may begin by expressing R'_N as follows:

$$R'_N = NR + KR$$

where

$$NR = \frac{R_N\ (1 - C_2 - .25C_1)}{(1 - C_2)}$$

$$KR = \frac{e}{1 - C_2}$$

NR can be thought of as the rate excluding the expense constant.

We may then proceed to redefine the expense elements in terms of NR and move along as in the past. When all revision calculations have been determined, the final rates are adjusted by the amount KR.

We define the PLR for NR as follows:

$$PLR' = \frac{R_N \times PLR}{NR}$$

$$= R_N \times \frac{(1 - C_2 - C_1)(1 - C_2)}{R_N \times (1 - C_2 - .25C_1)}$$

$$= \frac{(1 - C_2 - C_1)(1 - C_2)}{(1 - C_2 - .25C_1)}$$

The indicated change in NR can now be defined as follows:

$$I_{NR} = \frac{L'/P'}{PLR'}$$

where

L' = developed and trended loss and loss adjustment expense.

P' = total premium at current net ratio (NR).

Territory and/or class rates can then be developed as before and then adjusted for the revised expense constant:

$$KR' = \frac{e^1}{1 - C_2}$$

The method is conceptually sounder and procedurally easier to deal with than that suggested by the authors.

Other issues raised by the paper but not discussed, include items such as appropriate record keeping to reflect the collection, referral and cancellation of policy fees and allocation problems when the fee covers multiple lines of business on the same policy. This, along with the unlimited number of potential factors which could be considered for potential redistribution, strongly suggests that ratemaking and accounting concepts should be closely developed. This will permit responsiveness to the issues and ensure that the costs associated with improved equity objectives do not become overwhelming.

The authors point out that their paper tries to present "the basis concepts unlying the proper allocation of underwriting expenses." It presents the reader with a primer on some of the issues surrounding the subject and clearly demonstrates the potential for further work on the subject.

TITLE: IMPACTS OF STATE REGULATION ON THE MARKETING AND PRICING OF INDIVIDUAL HEALTH INSURANCE

AUTHOR: Mr. Charles Habeck

Mr. Habeck is a Consulting Actuary and Associate Member for Milliman and Robertson, Inc. specializing in health insurance. He received his FSA in 1977 and his FCA in 1979 and has been a member of the American Academy of Actuaries since 1972. Mr. Habeck has authored a number of articles on health insurance subjects for various publications.

REVIEWER: Mr. Robert Schuler

Mr. Schuler is Vice President for Blue Cross of Western Pennsylvania. Bob received his FCAS in 1967 and is a member of the American Academy of Actuaries. He also has an MBA degree from the University of Pittsburgh. Bob was on the CAS Education Committee from 1970 to 1978 serving as committee chairman for two years.

I. CURRENT REGULATORY ACTIVITY

State regulation of individual health insurance has increased greatly in recent years, both in scope and intensity. The need to comply with regulations has become the dominant objective in benefit design and pricing of individual health contracts. This shift away from the dominance of market forces results from the extension of regulation to almost every aspect of the development and marketing processes.

State regulation has grown so as to influence or mandate these items, among others:

The benefits that may or must be offered
The way that contract terms must be stated
The minimum "return" to policyholders
Permissible risk classes
Sales materials, product names, etc.

Mandated benefits. This facet of regulation is comprised of required, permitted, and prohibited benefits. The stress has been on required coverage of: treatment for alcoholism and drug abuse; expanded outpatient programs; expanded skilled nursing care to cover that care in other facilities; outpatient nervous and mental conditions; home health care visits; kidney dialysis and transplant surgery expenses; certain pregnancy benefits. In addition, there has been a general expansion of the definition of "physician." Now included are chiropracter, podiatrist, chiropodist, dentist, optometrist, osteopath, and psychologist, besides M.D., in general, any "licensed practitioner of the healing arts."

Statement of contract terms. This aspect of regulation involves, in addition to the already standardized uniform policy provisions, requirements as to structure, placement, type style, emphasis, and reading ease of the policy contract form. At last count, about two dozen states had enacted readability requirements or had begun to develop them. A model law has been developed, calling for a Flesch readability score of at least 40. Some states have adopted this standard, but at least one requires 50 and another is considering a score of 60 for its test.

Minimum loss ratios. Model rate filing guidelines have been developed, setting up a grid of loss ratio minimums, according to plan type and renewability provision. A number of states have revised their positions on this question within the last five years; prior to that most used a 50% loss ratio test, or had no test at all. The model law represents an effort to achieve greater uniformity among jurisdictions. The popular press has tended to make loss ratio comparisons a test of the suitability of an insurance plan, without regard to individual needs, resources, or method of sale of the product. Among its shortcomings, such an approach tends to ignore the distinction between group (wholesale) and individual (retail) marketing situations.

Classification of risks. Questions have been raised as to whether age, sex, or marital status may continue to be used as the bases for premium differentiation. Also, benefits may not vary for these classes. For instance, females must be offered the same disability benefits made available to males; pregnancy benefits must be offered to individually insured females, not just where both spouses are covered. Handicapped persons, including those with "stabilized" disabilities, must not be prevented from obtaining coverage at a reasonable cost. Maximum premiums for pooled uninsurable risks may be mandated by regulation.

Sales materials, product names, etc. The strictures applied here are, for the most part, not new. Outlines of coverage are required more often, sometimes with more demanding readability standards than those for the contract itself. Policies may be required to meet certain benefit standards in order to use particular labels. A few states, for instance, have prohibited the use of the product name "Medicare Supplement" unless certain benefit levels are provided, at least the same kinds of benefits covered by Medicare.

Regulation is steadily expanding in scope. Claim settlement practices have been the object of regulatory scrutiny in the past, but now there is increased concern for medical record privacy, at the time of issue also. Disclosure of underwriting procedures involving the Medical Information Bureau must be revealed to the applicant.

The general effects of increased state regulatory activity have been mixed. It is difficult to challenge the goals of this regulation. In practice, however, it produces considerable hardship for most insurers. "Hardship" means increased costs, extended time frames for product development, slower action on requests for rate increases. Required policy form variations have multiplied; the use of endorsements or amendments to standard forms may be found increasingly unacceptable.

State strategies are being adopted by the companies. For one thing, mandated benefits and minimum loss ratios may differ significantly, so that the differences cannot be absorbed on an equitable basis within a single plan code or rating structure. For another, claims experience may vary substantially by geographical area; use of area rating tends to conceal some of the differences. Some states want loss experience for their residents to be reported separately. This requirement relates to minimum loss ratio tests. Expected loss ratios may be compared with actual results, as a test of the assumptions in the initial rate filing.

Depending on the scope of the company's marketing operations, a point is reached, sooner or later, where compliance activities become as complex as those of a multinational corporation. It is a credit to the commitment of the health insurance industry to its policyholders' needs that so many insurers have chosen to meet the challenge of the evolving regulatory environment and stay in the business. Of course, not all insurers have done this. Some have withdrawn from certain states, while a few have dropped their individual health insurance lines completely.

The effects of regulation on the distribution system deserve more attention. Marketing cutbacks have hurt the agents. First, certain products have been dropped. Or they have been redesigned with more limited benefits and guarantees, and of course, with lower commissions. Agents specializing in health insurance must now sell higher cost policies with attained age premiums in return for lower compensation. Some of the pressure comes from inflation, some from competition, the rest from minimum loss ratio requirements.

Minimum loss ratio regulations indirectly control the expense factor; they also limit the margin for profit and contingencies. In fact, a product like non-can disability income cannot even be sold in a state with a minimum loss ratio of 65%, without a special dispensation.

Consumers are affected by regulations at least as much as the insurers and their agents: (1) some products will become unavailable due to regulatory strictures; (2) costs of compliance will have to be passed on to the policyholders; (3) agents can be expected to provide less service; (4) longer periods may be needed for claim settlement. The consumer will be paying for the enlarged regulatory staff as well as for the enlarged compliance staffs needed by insurers. But then, this condition pervades our society. What else is new? one well may ask.

II. A CANCER CARE POLICY

Cancer care policies have become very popular among the public. Consider these results for the leading writer of this coverage, the American Family Life Assurance Company of Columbus (Ga.). The table below shows premiums and claims, as reported in the Argus Chart[1] for health insurance, for the guaranteed renewable category, which is mostly cancer coverage.

American Family Life Assurance Company

Guaranteed Renewable Business
(Amounts in 000's)

Calendar Year	Premiums Earned	Claims Incurred	Loss Ratio
1978	$220,439	$107,487	48.8%
1977	169,816	73,466	43.3
1976	127,550	49,706	39.0
1975	81,092	33,484	41.3
1974	55,261	22,964	41.6
1973	43,107	17,033	39.5
1972	31,874	12,026	37.7
1971	22,915	8,511	37.1
1970	15,139	5,725	37.8
1969	10,530	4,300	40.8

Much commentary has been published on the cancer care policy. Despite the evidence of its popularity, it is difficult to find a kind word about it from insurance regulators, consumer advocates, or the popular press.[2] At least one state prohibits its sale, while others require that it be sold only with a comprehensive basic coverage for all causes.[3]

The thinking which underlies the opposition to cancer-only insurance seems to include the following objections: (1) the method

---

[1] Argus Chart of Health Insurance, The National Underwriter Company, Cincinnati, Ohio.

[2] See [20]. [3] See [24].

of sale is unacceptable, that is, "scare tactics" are said to be used or benefit levels misrepresented; (2) the return to the policyholder, measured by the loss ratio, appears inadequate; (3) a large profit factor obtains (the success of one company and the salary of its CEO are often cited[4]).

Possible responses to these objections include: (1) all insurance is purchased out of concern for or fear of financial loss; (2) benefits under most cancer plans are designed and advertised to be supplemental coverages; (3) low premium, low frequency risks will result in relatively higher expense components and lower loss ratios; (4) many companies have suffered losses on their cancer plans. For example, in an apparent effort to meet objections like these, at least one company has attached a return of premium rider to its cancer policy, an approach unlikely to result in excessive profits.

More and more companies now offer cancer policies. If your company does not, don't be surprised if your marketing committee brings up the question: Shall we develop and market a cancer plan? If preliminary considerations --company's image, company's distribution system, regulatory impacts-- can be accommodated satisfactorily, work may begin.

The first phase of development is to set the benefit structure. Cancer benefits currently being marketed should be studied. Two main approaches can be identified. Type A is most popular and provides defined or scheduled benefits by service category. Type B

---

[4]See [20], page 17.

is less used and pays lump-sum benefits by type of cancer. A sample of Type A benefits usually includes these items:

Hospital daily benefits: $50 per day, first 7 days; $30 a day after that. Includes common ancillary services. New stay begins if patient is out of hospital for at least 30 days.
Drugs and medicines: Often pays actual charges up to 10% of the hospital daily benefit payable.
Special nursing services: Up to $24 daily, $1,000 max.
Blood and blood plasma: Actual charges, $300 lifetime maximum, but no limit for leukemia.
Anesthesia: Up to $70 per operation, but $30 for skin cancers.
Ambulance: Up to $50 per confinement, $500 maximum.
Radioactive therapy: Up to $1,000; some plans pay the same for chemotherapy.
Physicians visits in hospital: Up to $10 per visit (one visit per day) and $600 maximum.
Surgical procedures: By schedule, up to $500 maximum.
Transportation: By air or rail to distant treatment centers, up to $500 maximum.

Other features sometimes found include: "bonus" payment of 10% of claim amount to cover non-medical loss; return of all premiums if death occurs before age 65; shift to 100% basis after 90 days in hospital, with monthly maximum of $5,000. A further option may allow benefits to be paid as if for loss of time, to avoid possible benefit reduction through application of COB provisions in other coverage the insured may have.

The above benefit array has been found acceptable in most states; specified minimum benefits (as in California) may be expected to increase periodically as inflation boosts costs.[5] Variations of Type A can also be found. One variation provides the same services without the internal limits, only an overall maximum of $10,000 or $20,000. Such a structure may become unworkable; it requires as much re-rating activity as a major medical product, but with a much lower premium base.

Another variation of Type A has benefit levels that vary by age. To insureds under age 45, benefits are paid at 150% of the scheduled amounts. To those age 65 and over, benefits are paid at 75%. In a couple of plans, for attained ages over 65, benefits may reduce to 20% or 25% of regular levels, to recognize the presence of Medicare; there is likely to be a corresponding drop in premiums. Note that the varied benefit attempts to achieve greater equity, since cancer plan premiums typically do not vary by age, while claim costs climb steeply as age increases.

The Type B cancer plan pays lump sum benefits at the time of diagnosis. Four categories are distinguished: (1) leukemia, for which the benefit amount is highest; (2) internal cancer, which is next highest; (3) skin cancers; and (4) lip cancers. Skin cancer benefits may be paid for up to 10 locations, while lip cancer benefits are limited to no more than two sites. The Type B approach typically includes an accident benefit so as to provide more benefits at the younger ages and thereby achieve greater equity by age.

Cancer benefits also may be used in all-cause plans; for instance, the daily hospital benefit may be doubled in a hospital in-

---

[5] See [23] for details.

demnity contract for confinements due to cancer. A comprehensive major medical policy provides broad coverage, but it still does not cover everything. The catch-all "bonus" payment of 10% could be used to meet some of the expense not covered by major medical.

Most companies have chosen the Type A approach. This means they sell a cancer-only specified package of benefits, with internal limits to minimize the need for rate increases. All sales materials should emphasize the need for other coverage, and, for Type A plans, avoid undue emphasis on the aggregate maximum benefit amount payable. Most companies sell one policy per family, but some allow the purchase of double benefits, or two "units" of coverage. To clarify the extent of coverage, a realistic sample claim should be shown, along with a breakdown on benefits under the plan. If such disclosure were required, it might have more impact on the marketing of cancer plans than loss ratio requirements can, since misunderstanding of the scope of benefits seems to be a major cause for complaints.

Once the benefit structure has been set, trial gross premiums may be calculated. An expected age distribution for new issues is needed because, although claim costs increase steeply by age, most cancer plans use very simple rate structures. There is one premium for individuals, one premium for families. Currently there is a move towards greater rate refinement, with some premiums coming out by sex and individual age at issue. This should help attract more of the younger lives.

The marketing method --group or individual, agent-sold or direct marketed-- will influence the premium structure. For instance, agent-sold individual policies presumably can make use of a more complex rate classification, since the agent is present to give assistance. Any of the group approaches, or the direct-mail individual method, would probably require easy-to-understand rate structures. Where does state regulation come in?

State regulation will have an impact on premium structure in the form of minimum loss ratio requirements and policy reserve requirements. The renewability provision of the coverage will affect the minimum required loss ratio. The kind of coverage and the average premium size may also control. Since the expense factor is effectively limited by the minimum loss ratio, the marketing method may be restricted in turn. These interrelationships can become quite complex.

Probably the best place to begin is to decide on the renewability provision. Few cancer plans are non-cancellable, most are guaranteed renewable with the right reserved by the insurer to increase premiums. A number of cancer plans use a limited right to non-renew provision. Very few will be strictly optionally renewable, that is, cancellable individually for any reason.

Assuming that the pressure of competition in the cancer insurance market limits the choice here to G.R. or to non-renewable for stated reasons only: the NAIC model guidelines will call for a 55% minimum anticipated loss ratio for either case. Furthermore, the guidelines allow a 5-point reduction for average premiums under $200, and another such reduction for cases where the average premium is under $100. The basis is the average premium for a given policy form, including any riders or endorsements.

[Note: Not all states will adopt the NAIC model, and many that do will introduce variations. Therefore, each state's rules must be confirmed.]

In addition, certain coverages call for special consideration, according to the model guidelines; cancer is one of these. Combining all these points, it may begin to appear that a target loss ratio of 40% to 45% may be used for a Type A limited package of cancer care benefits. Such a level is "reasonable," but only to someone who is familiar with the nature of the risk and the problems of marketing a relatively low-premium product. A more realistic assumption is that most states will require an expected loss ratio of at least 50%.

If an insurer has opted for a policy that is non-renewable for stated reasons only, he will enjoy these advantages: this provision allows for action on a state by state basis; additional policy reserves are not required in most states. Also, for the policyholder, although the plan is not G.R., no individual cancellation can occur; premiums will be lower than if the plan were G.R.

Exception: Under one state's rules[6] the classification of the policy will change, for reserving purposes, to be equivalent to guaranteed renewable, if premiums are level and a rate increase has been effected. At that time, additional reserves must be set up, to be funded out of future premiums, treating the date of the pre-

[6]Illinois, Rule 20.04.

mium increase as the date of issue and using the attained ages of the insureds at that time as the "issue age." Premiums must be level by issue age otherwise.

With cancer plans that have a single rate base, the question may come down to whether this "structure" represents a "level" premium or an "average" premium. An "average" premium implies a group or quasi-group rating approach. The closest example we encountered of this question involved a cancer plan with two premium classes, under 65, and age 65 and over. Premiums did not change at age 65. Benefits were the same for all ages. Two marketing methods were used: franchise group and individual issues. Premiums and policy forms were virtually identical. The franchise plan required continued membership in the group, but allowed non-renewal of the whole group. The individual plan allowed non-renewal of all policies in one state. After due consideration, an additional reserve was required by the state in question for the individual coverage but not for the franchise plan. This result emphasizes the need to better define the function of additional policy reserves for plans that are not guaranteed renewable.

Underwriting of cancer policies occurs in only three of the five usual ways: in the application, in the contract, and at time of claim. There is no medical examination; there is no APS. A question in the app may ask whether any person for whom coverage is requested has ever been diagnosed to have cancer (as defined). Sometimes, instead of a question, the applicant must acknowledge his understanding of this limitation, that is, that the plan will pay benefits only for cancer first-diagnosed at least say, 90 days after the effective date of the policy.

The policy repeats the provision on the 90-day wait. Also, the policy design, in the form of a package of benefits, is a method of underwriting, since it limits the risk on any one person. The plan promises to pay benefits without regard to any other coverage, but there may be a limit "in this insurer." In such case, a person with more than one policy in the same insurer will be paid under only one of them; any others will be void, and premiums will be refunded.

The key underwriting task occurs at claim time, since an examination of all applicants cannot be done at time of issue due to the expense. First, the presence of the malignancy must be established by review of a qualified pathologist's diagnosis. The rest is in the timing. Evidence must support the contention that the manifestation of the disease first occurred at least 90 days after the plan's effective date. Any other finding effectively voids coverage for that person; there may be a return of premium.

Regulatory requirements call for prompt and fair action by the insurer in settling claims. Privacy must be guarded. Delays may occur on an initial claim if information is lacking; but there is no defense for actions which may be prejudicial to the insured's rights. On the other hand, although the regulatory and judicial climate may currently favor the insured, there is nothing to prevent an insurer from bringing an action in response to a fraudulent claim, except, of course, the burden of proof.

In summary, the proposal for a cancer plan has resulted in a policy providing a relatively broad package of scheduled benefits, the same for all ages, designed to meet minimum benefit requirements and avoid the need for premium rate increases. Premiums have been set on an average basis to achieve a 50% anticipated loss ratio. Commissions, expense assumptions, and profit and contingency margins have been established to reflect this target loss ratio. The marketing program is expected to be mixed, in order to minimize any need for additional policy reserves. This means that the same package will be sold either individually or on a franchise group basis, as the situation may require. The franchise approach will be emphasized to realize expense savings and obtain a better spread of risks.

III. A MEDICARE SUPPLEMENT POLICY

Much attention has been paid by regulators to marketing practices used in selling health insurance to persons over age 65.[7] This age group is growing steadily, both in size and in political influence --that's part of the reason. Also, it has developed its own organizations and advocates. When the Medicare law in 1965 made basic health protection available to this segment of the population, it became apparent that this would be a good market for health insurance products that were supplemental in scope.

First of all, the Medicare program was not designed to cover

[7]See [25] through [35].

the entire health care needs of those over age 65, as these are broadly defined. Probably less than one-half of such costs were covered initially, and currently, it is estimated that only about 38% of these costs are covered by Medicare. Medicare supplement (M/s) policies are estimated to cover about 5% of costs; 19 million of these policies are now in force with annual premiums of about $4 billion (fall, 1979).[8]

A second general reason for interest in supplementing Medicare lies in the nature of the supplemental benefit package itself. Benefits supplemental to Medicare Part A (HI) will be fairly well insulated from inflation, since they are usually scheduled amounts. Although the amounts change from year to year, gross premiums also are allowed to change automatically in most jurisdictions. As to benefits which supplement Part B (SMI), they are subject to inflation, but in M/s plans their scope is much more limited than that of a typical major medical plan sold under age 65.

A third reason many insurers find this a viable market, although they may not recognize it as a factor, is the presence of utilization controls in the Medicare program itself. This is especially true for medical care benefits, where "allowable" charge levels as defined by Medicare rules are generally lower than "reasonable and customary" charge levels as recognized by most insured plans. Obviously, to take advantage of this control, the M/s medical care benefit level must be stated in terms of "Medicare allowable" charge levels.

---

[8] See [27].

There are some other good reasons for entering this market. If the need is there, the sale should not be particularly difficult. The M/s plan is therefore a good source of premium income for the insurer and of commissions for the agent. Another good reason, from the agent's and insurer's viewpoint, is that the contact with persons over age 65 can provide referred leads not only to others in the same age group, but also to children and other relatives with a variety of insurance needs in all lines. The most compelling reason of all, of course, is that A&H insurers may have no choice in that state but to offer an M/s program.[9]

State regulations applicable to this specific health insurance product have grown to staggering proportions in recent years. The statutes and regulations of the following states may serve as a starting point in any attempt to understand what is happening: California, Colorado, Delaware, Florida, Illinois, Iowa, Massachusetts, Michigan, Minnesota, New Jersey, New York, Oregon, South Dakota, Vermont, and Wisconsin. In some of these states, rule-making may still be in the initial stages.

The NAIC has developed model provisions as part of its Minimum Standards Act. The Health Insurance Association of America has formed a committee on the subject; the United States Congress has its own committees also.[10] The Federal Trade Commission is promoting legislation to require the HEW "Seal of Approval" for M/s plans issued in states which do not have regulations of their own. Additional Federal legislation is being proposed to allow insurance

[9]Michigan, MCLA §500.2265. [10]See [29].

commissioners to take jurisdiction over direct mail sales in their states, not now regulated by them.[11]

This growth in state regulation and Federal interest can be attributed to the following: (a) unusual marketing abuses, especially in the area of disclosure of benefits; (b) inability of the public to make meaningful comparisons of dissimilar products; (c) general vulnerability of the over age 65 population, combined with lack of information; (d) relatively high sales compensation coupled with a low return to policyholders, when measured by "loss ratio" results.[12]

Because of these conditions, Medicare supplement regulations have emphasized these elements: mandated benefits; minimum loss ratios; adequate disclosure. Buyers' guides are becoming more common,[13] and these must be provided to prospects at or before the time of sale. The "ten day free look" has been enforced and extended to a longer period in some cases. A final regulatory element should be repeated here: mandated availability of M/s coverage. So far, this coercive approach to handling the problem has not become widespread.

Cancer care policies and Medicare supplement plans show many similarities in design and marketing. For instance, high cancer incidence rates make the over 65 age group a prime market for can-

---

[11] H.R. 2602; H.R. 4000.

[12] See [26], pages 78-79. Also see [27].

[13] See [31] through [35].

cer coverage; for both plans sales techniques have involved arousing the fears of the prospect; both plan types are designed to provide supplemental benefits; early loss ratios may appear to be low for certain plans of each type (this relates to the scope of benefits, waiting periods, etc.).

There are important differences too: more benefit variations have been used in the design of the M/s plans; more complaints have probably been made about benefits which were thought to be covered under the M/s plans; active promotion of M/s plans by some regulators has occurred, a far different stance from that adopted towards cancer plans.

Perhaps the biggest problem --at least the most dramatic-- that has been found with Medicare supplement plans is that of multiple sales, many policies to the same insured. The solution to this problem should be one of the basic goals of the plan design. Benefit structure should be understandable; the possibility for overlapping coverage should be minimized. Currently-marketed M/s policies are designed in at least three fundamental ways: limited benefit plans, comprehensive plans, and "building block" plans.

Limited benefit plans have scheduled benefits, but usually no out-of-hospital coverage; maximums are low. Comprehensive plans may be scheduled or unscheduled; they cover expenses incurred in or out of the hospital. Plan maximums tend to be high.

"Building block" plans combine certain features of the first two types, using a limited in-hospital benefit as the starting point. Additional benefits are available by rider to complete the program;

this is convenient if at first the full premium for the more comprehensive program is not available. Still, the final package may provide less overall coverage than a comprehensive plan.

As a practical matter, the comprehensive plan approach must be adopted unless the insurer can avoid marketing in certain states; the only alternative is a more complex Medicare supplement series, where several policy forms are developed, geared to groupings of states. In any case, the significant decisions left open narrow down to about half a dozen features of the benefit structure, as follows:

1. The Part A Medicare deductible and copayments are usually covered, through the 60-day lifetime reserve. Some regulations require full coverage after this reserve has been exhausted. An alternative here would be to offer a daily benefit equal to the deductible, or perhaps up to twice the deductible for each day of hospital confinement after the reserve is used. The deductible amount is nominally supposed to represent the cost of one day in the hospital, but it is probably too low. Relating this extended daily benefit maximum to the deductible simplifies pricing and keeps pace with inflation.

2. Extended care in a skilled nursing facility is usually covered in the amount of the copayment for days 21-100. Beyond 100 days Medicare benefits cease, for that spell of illness. Some M/s plans provide benefits for stays longer than 100 days, out to two or three years. Such long stays are rare; the average stay is under 30 days. Long stays tend to involve other types of care, such as

intermediate care or custodial care. State regulation may mandate inclusion of intermediate care facilities as providers of skilled nursing services if that level of care is actually provided.

3. Medicare does not cover the first three pints of blood or blood plasma. This benefit is becoming more common in M/s plans, reversing the assumption that voluntary donors or credits are usually available and preferable.

4. Provision for reimbursement of the Part B $60 calendar year deductible has had the most variations. First, it may be completely excluded. Second, it may be covered in-hospital only. Third, it may be covered as a disappearing deductible. Fourth, it may be covered 100% if in hospital, and ignored for expenses incurred out of hospital, that is, treated as part of the eligible expenses that are reimbursed at 20% of R&C. Fifth, it may be reinbursed fully.

5. The Part B 20% coinsurance (after the first $60 per year) has several variations, too. The minimum is to pay it only if due to hospital confinement. The maximum, one may surmise, would be to pay the whole 20%, in or out of hospital. This is wrong. The maximum benefit here is to pay the excess of reasonable and customary medical expense charges over 80% of what Medicare allows, since "allowable" charges will be less than R&C. One state may require this maximum.[14]

6. Out-of-hospital prescription drugs and private duty nursing are not covered by Medicare at all. Insurers are often criticized for not providing benefits in these areas. They are also criticized

[14]Massachusetts.

for not covering custodial care, or outpatient psychiatric care beyond what Medicare provides. To what extent can any of these benefits be covered, if at all?

The plan features outlined above are relisted below with suggested coverages to be used in a comprehensive plan. This plan will not be sold where state regulation is "coercive," because there the choices have already been made. Note that certain alternatives fit the concept of catastrophic coverage, if this is the marketing image desired.

(1) Hospital deductible and copayments are always covered. After the 60 days' lifetime reserve has been used, coverage under the M/s plan should take over, running out to a full year of hospitalization, or unlimited if catastrophe needs are stressed. A stated maximum daily benefit should be used, relating to the Part A deductible if possible.

(2) Skilled nursing care can be handled best by paying the copayment for days 21-100 and stopping there. Catastrophe emphasis calls for an extension, consistent with that for hospital.

(3) Coverage of the first 3 pints of blood involves a significant cost. It may become necessary for competitive reasons.

(4) The Part B $60 deductible should be either completely excluded or else completely ignored. If it is covered, there will be many small claims. For many insureds, the annual gross premium for it will exceed $60 (in and out of hospital both).

(5) The 20% coinsurance should be paid on a reasonable and customary basis, after the first $60 per year. If the gap between R&C

and Medicare "allowable" charge levels becomes too great, this percentage can be increased. The fixed percentage approach speeds up claim settlement, since there is no need to hear from Medicare. Also, it leaves a small gap in the charges so that the insured remains interested in expense levels. An out-of-pocket maximum can be used to keep this gap from becoming a hardship.

(6) Medicare did not find it feasible to cover these items; insurers may come to the same conclusion. Prescription drugs are high frequency, low cost items. Special systems are required for management control. Since this item is seldom covered in M/s plans, its inclusion is likely to result in selection against the insurer, at least experience shows this to happen. This is a problem that even relatively high deductibles cannot solve. Leave it out. Private duty nursing is a high cost, low frequency item, just the opposite of drugs. Aged persons who need this level of care are likely to be hospitalized. If they are ambulatory, home health visits are available. The benefit is little used, but may be included if the stress is on catastrophe care. Custodial care is uninsurable. Outpatient psychiatric care can be covered 50%-50% as done by Medicare, out to $1,000 without much problem, if this is desired. Medicare pays half of the first $500 only.

One other benefit may be considered for the over age 65 market: a daily hospital benefit for the first 60 days in a spell of illness. Since Medicare covers this period fully (except for the deductible), this benefit should be sold as an income benefit. It may be offered where the agent find a comprehensive M/s plan already in place, if the agent's own M/s plan is not superior. The use of this benefit in combination with an M/s plan may be necessary to comply with minimum standards tested by equivalency rules. This happens if the insurer prefers not to cover the Part B $60 deductible. The hospital income benefit for days 1-60 will provide the extra points.[15]

Specific plan design features are detailed in the regulations of Massachusetts, Michigan, Minnesota, and Wisconsin (and others, no doubt). Massachusetts rules mandate very comprehensive minimum benefits, including drugs, although deductibles are allowed. Michigan rules require full coverage of all gaps, no exclusions except those for Medicare, and no limits on pre-existing conditions. Minnesota defines a qualified plan and applies an equivalency test to deviations from it. The Wisconsin rule outlines four plan types; at least one other state may follow this pattern. In other states, the general rule for a plan to be sold as a Medicare supplement is that it must provide the same scope of benefits as Medicare; it need not go beyond this, it may have like exclusions.

Premium structure, renewability, and reserve requirements have not been unusually affected by state regulation. But premium levels have been hit by loss ratio tests. The NAIC filing guidelines call for 60% as the target, as do a number of states. Several states require a 65% loss ratio; Congress talks of a 70% requirement.

Premium structure may be very simple: one premium, unisex basis, same for all ages. Or it may be complex: male and female rates in five-year age groupings. Premiums are almost always level, based

[15] Minnesota test of actuarial equivalence.

on original issue age, but with automatic changes as Medicare provisions change. Policies are usually guaranteed renewable for life (sometimes mandatory). Level-premium, G.R. policies require additional reserves. Statutory minimum standards have not yet been adopted, although the 1974 Medical Expense Tables have been proposed.[16] Unfortunately, these tables do not provide factors for comprehensive Part B benefits, making their applicability limited. Also, they appear not to have been tested against actual Medicare experience. The most practical approach under the circumstances is to base additional reserves on the expected morbidity assumed in the premium calculation (ultimate basis), as would be done for a major medical plan.

But benefits change each year. One way to cope with changing Part A amounts is to adjust reserves annually, using a dual calculation. Those benefits subject to change may be valued per $4 of Part A deductible; all other benefits would be grouped and valued per policy. Both valuations would use original issue ages. This approach is convenient and not overly conservative. At least one state applies its loss ratio test ignoring the increase in additional reserves.[17] No state has yet specified a required method of reserve strengthening, either for M/s plans (subject to benefit changes) or for major medical plans (subject to inflation), beyond

---

[16] Anthony J. Houghton and Ronald M. Wolf, "Development of the 1974 Medical Expense Tables." Transactions, Society of Actuaries 30: 9-69; discussion, 71-123.

[17] Colorado; see 10-8-101(1).

the general requirement of adequacy. In this connection, note that although a state may require automatic benefit increases as Medicare changes, the corresponding adjustment to premiums may require a demonstration that target loss ratios are being met. If they are not, the rule may call for a rate reduction.[18]

Loss ratio regulation limits methods of distribution. The master GA arrangement, with gross allowances of 80% first year and 25% in renewals years, may be a thing of the past in most states. To achieve a loss ratio of 65%, agent compensation must be reduced to about 40% first year, 10% thereafter. A direct marketing insurer, paying no commissions, may be able to operate within a 75% expected loss ratio level. However, the product may not be the answer to everyone's needs, or it may not be obtainable.

Initial underwriting of applicants for M/s policies has taken two common forms: (1) accept or reject, simplified app, single rate table; (2) substandard approach, standard app, up to four rating tables. Probably about the same number of rejections occur in both systems. It is likely that agents will not submit apps if they anticipate a rejection; but the general idea is to avoid the "sure claim." State regulations have allowed both approaches. An exception is Michigan which allows no restrictions unless the applicant was without group or individual medical expense insurance (reimbursement type) throughout the five-year period just prior to

---

[18] Colorado Rule 78-1 requires 60% loss ratio; 10-8-102.5(2) provides for rate reductions if this test is not met. Michigan has required such justifications for a number of years, not just for M/s plans, however.

the date of application. In such case, there may be a 6-month wait on pre-existing conditions.

Most contracts normally include such a waiting period; commonly it is six months. Then pre-existing conditions are defined as those for which treatment has been received in the 6-month period just prior to the policy's effective date. Waits of three, five, and 12 months have been used, with corresponding variations in the definition of pre-existing conditions.

In summary, the likely choice for the Medicare supplement plan will be one which provides relatively comprehensive benefits in a single package. The anticipated loss ratio will be 60% to 65%. Acquisition and maintenance expenses, along with profit and contingency margins will be limited by the required loss ratio. The plan will be guaranteed renewable for life and will require additional reserves. A hospital income policy and/or rider will be available as a companion product for this market. Underwriting will be on an "accept or reject" basis, the goal being to avoid the sure claim situation. A single premium scale (one class) will be used, with unisex rates and five-year age groups. Careful compliance with disclosure rules will be emphasized throughout the marketing program.

IV. THE RANGE OF REGULATORY ATTITUDES FUTURE TRENDS

Most insurers who market individual health insurance coverages are aware of the widely differing regulatory attitudes among the states. In this context "attitude" means something like what "competitive stance" means for an insurer. It reflects the perceived commitment of the regulatory agency to the carrying out of its mission. This parallels the insurer's commitment to meet the health insurance needs of its market. Just as holds true for the company, the commitment of an insurance department to a regulatory program can be measured by the resources allocated to the job, in terms of money, time, and personnel. A certain priority among regulatory programs can probably also be observed.

One actuary has divided the states into three categories, based on policy filing results for his company (which writes in all states but New Jersey and New York). Some are found to be "reasonable," others tend always to find objections, and the rest lack a uniform response pattern.[19] This response distribution has already been illustrated in the above discussions of two common supplemental health insurance products.

For the cancer care policy: some states prohibit its sale entirely; other states require that it be sold with all-cause basic coverages or not at all; still others control its use through loss ratio requirements and minimum benefit standards. Most states do not prohibit its sale.

For the Medicare supplement policy: one state requires its sale on a guaranteed-issue basis even if the insurer has never had this kind of policy; another state requires that it be available in a "qualified" plan that meets minimum requirements; still other

[19] See [12], page 736.

states require that any plan labeled "Medicare Supplement Policy" must provide minimum benefits and must be sold following specified procedures, primarily disclosure rules. Most states currently permit the sale of the M/s policy with relatively few restrictions.

Insurance department prohibitions of certain products exemplify the ultimate regulatory solution. For instance, is cancer-only coverage undesirable per se? The need for it continues to grow. If "only" 1 out of 4 persons is afflicted by cancer, does that make the plan a "bad buy"--just because the other three never collect?[20] Improved treatment techniques will undoubtedly lead to more survivals and the need for more hospital and medical care.

The prohibition of premium refund riders involves a similar judgment of undesirability. It has been demonstrated that profits for this benefit flow from withdrawals without value (tontine effect). Why not require a cash value consistent with reserve requirements? The real problem with the ROP rider is that premiums have proven to be inadequate and reserves have been based on favorable expectations that have not materialized, especially as to persistency and claims offsets. The regulatory alternative to prohibition of the ROP rider lies in insistence on its proper pricing and reserving. The market will do the rest.

Other examples may be found of regulations that need a "course correction." The diversity of regulatory response raises a number of questions, whose answers generate still more questions.

1. Is uniformity of state regulation necessary or desirable? If

---

[20] See [20].

it is, and it cannot be achieved under the <u>status quo</u>, what alternatives are there? Conditions vary from state to state, as do population characteristics. Some consumer groups may want or need greater protection than others. There have always been differences among the states. The same kinds of variations that we see among the states may be found among the countries who are joined in the European Economic Community.[21] Both here and there, it may be noted, some of the matters in contention appear to represent trivial differences in the way of doing things. So far, the alternatives to state regulation have not been considered practicable or desirable.

2. <u>More specifically, shall premium rates vary by state according</u> to differences in loss ratio requirements? Shall commissions vary? If not, such differences may lead to subsidizations that are difficult to rationalize. Many states now require that claim experience for their residents be reported by itself in addition to aggregate data. How may the insurer best cope with rules that lean towards one-way protectionism?

3. <u>If a state fails to create a "proper" regulatory environment</u> -- that is, one deemed sufficiently responsive to consumer interests-- shall its authority be pre-empted by Federal rules? Recent events seem to point towards such a result. However, it should be kept in mind that state regulation involves considerable extra-territoriality. The marketing program of an insurer operating in many states (rather than in just a few) will be influenced by the rules of

---

[21] See [7].

those with the greatest reasonable commitment, with a "spillover" effect into the less active states. For instance, competition is now tending to increase target loss ratios above the required minimums in some states, due to this spillover effect. An M/s plan with a 75% loss ratio will appear more attractive than one with a 60% loss ratio, if both are equally accessible and provide comparable benefits. As long as enough of the states are active, Federal intervention in the regulation of insurance will be hard to justify.

4. <u>At what point should insurers challenge state regulations</u>? Do we need a set of guidelines within which the regulators must confine their activities, or is the U. S. Constitution enough? There are signs that regulatory activity is reaching a plateau (see list below). Insurers have been exhorted to "act and not react." Where are they to begin? Insurers and regulators cannot operate at arm's length; both need to appreciate the goals of the other. Regulators must be concerned about insurance company risks and profits; insurers must be concerned about benefit returns and policyholder rights. Beyond the current plateau lies a mutual educational effort.

*   *   *   *

As answers are sought for these questions, the underlying one remains: What has happened to individual health insurance markets as a consequence of regulatory activity? The answer here will provide the basis for insurer planning and activity.

It appears that the market for individual health insurance has eroded over the past decade. There are at least two reasons for this: (1) expanded government programs of health care and income protection in direct competition to private insurance programs; (2) expanded regulatory activity at both state and Federal levels, touching different aspects of the marketing process. As one Congressman has concluded:

> Our government has closed off opportunity, discouraged entrepreneurs, limited productivity and stifled freedom. Yet the government's moral attitude is that it's doing just the opposite.[22]

The level of regulation and the level of insurer response to it may have reached plateaus. Insurer responses in this recent regulatory growth period have included:

State strategies: This makes the marketing scene something like entering the presidential primaries, win some, lose some, but hope to end up with the nomination; or like playing the new Monopoly game, tailor-made to each metropolitan area.[23]

Avoidance of regulation: A different vehicle, such as a trust or quasi-group arrangement, or self-insurance, removes the product from control of the regulators; re-design of the benefit structures may accomplish the same purpose.

Cessation of marketing: The insurer ceases marketing, at least directly, opting out of the coercive environment, and contracting its premium base rather than ignore insurance principles or endure forced or inflexible marketing constraints.

---

[22] Newt Gingrich, letter to Wall Street Journal, December 10, 1979.

[23] "Stock Block," ©1978 John F. Majors (J.F.M. Games Co. Seattle WA)

The tide may be turning. A number of events foreshadow changes in regulatory emphasis. Here is a sample:

--An apparently successful challenge has been made to the Minnesota Comprehensive Health Insurance Act of 1976;

--The New York Department has "exhibited concern that individual accident and health insurance availability is greatly diminished since the enactment of the maternity law" there;

--The Massachusetts Minimum Standards Regulation is to be challenged, especially as to the prohibition of cancer-only coverage; an injunction will be sought to bar enforcement;

--The trend to deregulation has taken hold in Canada[24]; perhaps the fallout will be felt in the United States;

--Mandated health insurance benefits may encounter greater resistance, and require a new rationale for justification[25];

--Congress appears more inclined to take action to reduce FTC rule-making activity.

The trend exemplified in the development of the Wisconsin rule (Ins. 3.39) on Medicare supplement marketing may be expected to influence regulatory activity in the 1980's. Other states are looking at this approach, no doubt because so far it seems to have successfully balanced the interests of the concerned parties.

The essence of this trend is education of the consumer and preservation of the market place.[26]

---

[24]See [12], page 739.

[25]See [10].

[26]See [12], page 740.

The vehicle of this educational thrust will be manifold. It will involve the schools and the media. If it succeeds, we may all begin to agree on the following:

1. That the price for a retail product is greater than that for a wholesale product.
2. That the loss ratio test is not a measure of product suitability in given circumstances.
3. That agents deserve adequate and proper compensation for services performed for both the insured and the insurer.
4. That consumers deserve an insurance product that does what they think it will do, while giving them this protection at a fair price.
5. That "insurance" is not defined as protection provided to "those who need it the most."
6. That government-sponsored or self-insured health programs operate under the same basic principles as do private health insurance programs.
7. That the appointment of experienced and knowledgeable insurance persons to state insurance departments will not compromise the regulatory mission.
8. That regulation should foster competition.
9. That product availability is inversely proportional to the <u>coercion index</u> of the regulation that governs it.
10. That insurance companies are private business enterprises serving public needs, but are not public utilities and are not consumer co-ops.

Other learnings may result, but general acceptance of these will serve to better balance the critical interests of all parties. Such acceptance will also allow market forces to resume their proper role in benefit design and pricing of health insurance products.

V. LIST OF READINGS

A. General

[1] Anderson, William R., "Is the Supreme Court Narrowing the Scope of State Regulation?" Life Association News 74 (September, 1979): 151-154.
Discusses case of medical malpractice insurance in Rhode Island; 3 of 4 insurers dropped it, the fourth shifted to claims made basis. Result: complaint of conspiracy, boycott, etc.

[2] Doherty, Neville, and Crakes, Gary, "The Impact of a Change in Regulations on Costs in an Experimental Program." Inquiry 16 (Summer, 1979): 154-157.
Discusses "nominal" and "opportunity" cost aspects of requirement of informed consent, on a program in progress. Interesting study, although limited.

[3] Gold, Melvin, "Improving State Insurance Supervision." Best's Review (L/H) 80 (July, 1979):24-25.
Author appears to support radical changes in procedures.

[4] Habeck, Charles, "Coping with Minimum Loss Ratio Regulation." Best's Review (L/H) 79 (May, 1978): 19+.
Discusses varying state requirements and definitions of "loss ratio," including marketing & actuarial implications.

[5] Kafka, Franz, The Castle. Modern Library, New York, 1969. Also available in Schocken Books series.
Allegorical treatment of problems encountered in dealing with a bureaucracy.

[6] Kristol, Irving, "The 'New Class' Revisited." Wall Street Journal, May 31, 1979 (editorial page).
New Class has "little use for our commercial civilization or its market economy," seeks power through public and non-profit sectors, including media. Creates a convenient stereotype, but more applicable to a few consumer advocates than to most regulators.

[7] Kronholz, June, "Consumerism European-Style." Wall Street Journal, November 20, 1979 (editorial page).
"Competition" in consumer programs results in uneven

effects on members of European Economic Community.

[8] Pharr, Joe B., "The Individual Accident and Health Loss-Ratio Dilemma." Transactions, Society of Actuaries 31 (1979).
Discusses problems related to loss ratios for individual health insurance. Suggests adjustments to allow more meaningful analysis of results.

[9] Warsh, David, "The Great Hamburger Paradox." Forbes 120 (September 15, 1977): 166+.
Investigates cost elements of a finished product, including requirements of public sector and other institutions.

[10] "Mandated Extra Coverage May Violate Constitution." Article in The National Underwriter, December 9, 1978.
Reports on speech of ACLI counsel; Preston says retroactive changes in existing contracts violate 14th Amendment.

[11] "The Evolving Regulatory Environment for Health Care." Record, Society of Actuaries 3 (October, 1977, Boston): 835-854.

[12] "Effects of Consumerism & Regulation on the Health Insurance Industry in Canada and the United States." Record, Society of Actuaries 5 (June, 1979, Banff): 725-747.
Various aspects of regulation, including Federal v. state. Mr. Wood's remarks on deregulation and breadth of responsibility are pertinent.

[13] "The Extent of Federal Insurance Activities." Best's Review (P/C) 80 (August, 1979): 18+.
Some items under FTC duplicate areas of state concern.

[14] "Competition in Health Planning Enacted in Amendments to Law." Health Services Information 6 (October 9, 1979).
Sees health planning amendments as shift to planning through competition, moving away from planning by regulation.

[15] "Michigan Revamping Plans." Health Lawyers News Report 7 (November, 1979): 7+.
To slow costs, legislature will revamp Blues (they cover 58% of people); may limit plans reserves. Goal is to change Blues to consumer organization or co-op.

B. Cancer

[16] Benson, E. F., "Caterpillars," in Great Tales of Terror and the Supernatural, ed. by Herbert A. Wise and Phyllis Fraser. Modern Library, New York, 1972, 760-768.
Illustrates element of fear connected with cancer.

[17] Epstein, Samuel S., The Politics of Cancer. Sierra Club Books, San Francisco, 1978.
Reviews chemical industry resistance to regulation.

[18] McMennamin, Breeze, "A Heck of a Sales Force." Forbes 119 (March 1, 1977): 53+.
Traces progress of American Family Life Assurance under John B. Amos; marketing techniques; foreign markets.

[19] Schwartz, Harry, "A Look at the Cancer Figures." Wall Street Journal, November 15, 1979.
Rational discussion of cancer statistics, with age adjustments. No cause for alarm, unless you smoke. Dr. Epstein responds in letter to WSJ of December 10, 1979, page 23. Says "burgeoning cancer toll" now affects one out of four.

[20] "Why Cancer Insurance Is a Bad Buy." Changing Times 33 (December, 1979): 15-17.
Details good; conclusions doubtful. For instance: since only 1 in 4 Americans gets cancer, other 75% do not and won't get any benefits, making insurance a "bad buy." Compares loss ratios for dissimilar marketing situations.

[21] "NALC Urges Ingram Not to Ban Cancer Insurance." The National Underwriter, June 16, 1979, page 21.
NALC says fear of cancer exists apart from insurer activity.

[22] Three articles from Scientific American:

Old, Lloyd J., "Cancer Immunology." Scientific American 236 (May, 1977): 62-79.
How do cancer cells evade the immune systems of the body?

Croce, Carlo M. and Koprowski, Hilary, "The Genetics of Human Cancer." Scientific American 238 (February, 1978): 117-125.
Shows how to identify chromosome involved in transformation of a normal cell into a tumor cell.

Nicolson, Garth L., "Cancer Metastasis." Scientific American 240 (March, 1979): 66-76.
Investigates types of tumor cells that can travel through the body and what they have in common.

[Cancer research goes on and on and on.]

[23] California regulation: CAC 10 Chapter 5 Subchapter 2 Article 1.5 Section 2220.24.
Outlines minimum benefits of Type A plan.

[24] New York: Regulation 52.16
Bans sale of cancer-only coverage without all-cause basic coverage. Allows 6-month waiting period.

C. Medicare Supplement

[25] Gornick, Marian, "Medicare Patients: Geographic Differences in Hospital Discharge Rates and Multiple Stays." Social Security Bulletin 40 (June, 1977): 22-41.
The data on re-entries are valuable; these, together with results by region, affect cost of Part A deductible.

[26] Hoecker, James J., "Section Ins. 3.39, Wisconsin Administrative Code: The Origins and Development of a Medicare Supplement Insurance Regulation." The Insurance Law Journal 673 (February, 1979): 73-101.
Valuable account of rule-making procedures, including industry participation. Thorough documentation.

[27] Montgomery, Jim, "Predators Find Elderly Are Often Easy Prey for Array of Rip-Offs." Wall Street Journal, November 9, 1979, front page.
Lists scams perpetrated on elderly. Notes that Medicare pays 38% of total health costs, supplementary plans pay 5%. Multiple sales cited.

[28] "What Medicare Will (and Won't) Do For You." Changing Times 33 (January, 1979): 39-42.
Concentrates on explaining how Medicare works, with stress on its complexities and limitations.

[29] "Medicare Supplement Probe Hears Regulators." The National Underwriter, December 9, 1978.
House Select Committee on Aging hears views of commissioners from four states; views differ on need for Federal activity and its degree.

[30] "Pledges Solution to Medigap Abuses." The National Underwriter, March 31, 1979.
HIAA President Robert Froehlke pledges effort at state and company levels to solve problems of abuse. Cites multiple sales, undesirable sales methods, inadequate coverage, and high rates.

The following items are available to the public on request:

[31] "What You Should Know About Health Insurance When You Retire," Health Insurance Institute, 1850 K Street, N.W., Washington, D.C. 20006. 18 pages.
Describes Medicare program, ways of closing "gaps." Suggests health emergency fund for anticipated out of pocket expenses.

[32] "Advice on Health Insurance for Senior Citizens in Illinois." Illinois Department of Insurance, Springfield, Illinois, 62767. Free; send self-addressed mailing label.

[33] "When Medicare Is Not Enough." Albany, New York, 1979. Send 67¢ in stamps to Medigap, New York State Consumer Protection Board, 99 Washington Avenue, Albany, N.Y. 12210.
This source describes Medicare program, ranks supplementary programs. A discussion of this guide, including industry responses, appears in the following article:

Herman, Tom, "More on Medicare Supplementary Insurance." Wall Street Journal, August 20, 1979, page 28.
Four other reports on the subject are listed at the end of this article.

[34] "Health Insurance Advice for Senior Citizens." Prepared by State of Wisconsin, Office of the Commissioner of Insurance, 123 West Washington Avenue, Madison, WI, 53702. Revised each year.
Outlines benefits in Wisconsin-approved plan types. Discusses "limited" policies; warns about nursing home plans.

[35] "Approved Medicare Supplement Policies." Available from same address as for Item [34].
This chart, updated often, shows all approved Medicare Supplement plans in Wisconsin. Company, policy form, plan type, age 65 premium rate, underwriting, pre-existing condition limitations, commission scale, and expected loss ratios are all shown. Available on request.

Two problems: a single plan type can encompass range of benefits; marketing methods are not distinguished.

IMPACTS OF STATE REGULATION ON THE MARKETING AND PRICING OF INDIVIDUAL HEALTH INSURANCE

by CHARLES HABECK

reviewed by ROBERT SCHULER

Mr. Habeck's timely article presents a clear view of the impact of regulation on individual health insurance practices and policies which has heightened in recent years as a result of perceived and/or imagined shortcomings in the industry by consumer groups, legislators and regulators. The author's discussion of the instruments -- mandated benefits, minimum loss ratios, policy readability, reserve requirements and risk classification -- used by government regulators provides the reader, if he or she has not already experienced it, a sense of the pervasiveness of government rules and regulations. In fact, in a recent book by Murray L. Weidenbaum, The Future of Business Regulation, Mr. Weidenbaum notes "At times it seems that each and every move that business makes is studied with almost obsessive attention by one or more regulatory agencies, far out of proportion to the inherent need for government attention."[1] Yet, in spite of the regulator's growing omnipresence there is still considerable room for private initiative and action.

(1) Weidenbaum, Murray L. The Future of Business Regulation. Amacom, New York, 1979.

In the opening paragraph of the paper, Mr. Habeck raises the issue of relative effect or importance of "market forces" versus "regulation" in benefit design and pricing of individual health contracts. It is doubtful as contended by the author that regulation has become the "dominant objective" in benefit design and pricing of individual health contracts. Rather, the rules for playing real life monopoly -- old and new -- have become more cumbersome and pervasive thus requiring the players to spend more time studying the rules before playing the game. The "market forces" or the "game" remains -- to provide the challenge of obtaining a fair market share while meeting the company's objective in underwriting results. Certainly, at one time or another, we all have probably agreed with the author on the "regulation dominance." However, in our more rational moments we usually accept some government regulation as necessary and work to limit its scope and influence to only those activities that provide government "oversight" or "review" and restrict or eliminate government design or structure of policy benefits. For example, minimum loss ratio requirements appear to speak to the results of marketing products and hence provide government oversight or review opportunities. In contrast, minimum standard legislation encroaches on an insurer's benefit design practices. In this respect, Timothy B. Clark writing in a recent article in the

National Journal[2], notes that new approaches to regulation are needed. Certainly the insurance industry would welcome some innovation in this area.

Following are some observations on the various causes of the increases in regulations and mechanisms used by regulators.

Mandated Benefits

Much of the mandated benefit pressure comes, indirectly, from special interest groups which in some cases are sponsored by providers of care. EEOC and women's rights groups have also been influential. It seems that the concept of "insurable hazard" is gradually being replaced by "planned, budgetable expense" concept. This phenomena seems to be spreading from the group insurance market to the individual health insurance market.

Statement of Contract Terms

The intent of the regulators on this aspect is laudable. However, the "Flesch Test" criterion used in some states is not the answer. The October, 1979 issue of the Actuary, the monthly newsletter of the Society of Actuaries, gives a superb example which would pass though few readers would understand the mathematical proof, and a rather prurient passage which is readily comprehensive, but

(2) Clark, Timothy B., "New Approaches to Regulatory Reform -- Letting the Market Do The Job". National Journal, August 11, 1979.

would not pass. Expanded use by the Industry of communication and legal experts in the drafting of policies would seem to offer a more desirable route.

Minimum Loss Ratio

While regulators have found these useful comparative tools, the author's point that more thought should go into the drafting of the regulations is well taken, especially in such matters as statutory reserve, agency compensation, and return on stockholder equity.

Classification of Risks

Even though some classifications of risk have been under attack by regulators, insurers have been generally successful in maintaining proper classification systems. A real danger exists that regulatory actions could result in serious antiselection and could result in citizens of some states being denied access to needed insurance protection.

Sales Materials, Product Names, etc.

In many instances, the restrictions here are generally desirable and for the public good. Too often sales material - by the mere policy name - implies broader coverage than the policy actually provides.

In the author's concluding comment on these regulatory forces, he sums up one of the most serious impacts of state regulation when he notes, "The consumer will be paying for the enlarged regulatory staff as well as for the enlarged compliance staffs needed by insurers." Certainly, with these costs increasing daily, it is in the public interest for regulators to look for "new approaches" and "innovation" in their actions as called for by Timothy B. Clark.

Mr. Habeck's discussion of the market, underwriting, and regulatory considerations needed in designing and underwriting a cancer care policy and a Medicare supplement policy clearly illustrates the importance that regulatory requirements now play in benefit design. However, it should again be pointed out that regulatory requirements remain secondary to the carrier's dual requirements of meeting the market demands of policyholders and its underwriting practices and objectives.

While I have had limited experience in the cancer care policy, I can understand its popularity and equally the controversy about such policies. Cancer spells fear and emotional reaction. Consequently, there is no doubt a certain segment of the industry is using this to their financial benefit. Unfortunately, they have given the more responsible insurers in this market a bad image and hurt the marketing of this useful product. I personally believe that responsibly set minimum loss ratio will bring a respectability to this type coverage and help its future availability.

The author provides a good overview of the Medicare program and the reasons for insurers becoming interested in the Medicare supplement. However, the author seems to imply that rate adjustments for the supplemental policies are somewhat automatic. It should be noted that several company's success in obtaining necessary rate relief has been relatively poor, even when the Insurance Department actuaries stipulated the actuarial soundness of such requests. While these companies are still in the market, they have chosen to limit efforts for market expansion until a resolution of philosophical and political issues is reached - social goals versus sound underwriting practices.

The author's observation that Medical supplement plans supplementing Part B (SMI) are not impacted by inflation to the extent "of a typical Major Medical plan sold under age 65" does not coincide with this writer's experience. We have found the claim trend factors for Medical supplement plans and the typical Major Medical plans to be comparable.

Despite the difficulties experienced by some insurers in obtaining rate approvals, Medicare supplement is a viable market for an insurer to pursue. Level premiums for either Medicare Part A or Part B supplemental policies have built-in pricing difficulties and marketing advantages of implied - but not guaranteed - future premium outlay. An alternative utilized extensively by A & H insurers, including Blue Cross and Blue Shield Plans and the government for Medicare Part B, is one-year term pricing. This pricing strategy can be coupled with age and/or sex differentials, or a single "average rate" may be employed.

The advantages include not having to establish statutory reserves and the ability to react quickly to changes in the pattern of utilization and to cost trends. The major disadvantage - frequent rate increases - is shared by the so-called level premium plans as they respond to the constantly changing benefits supplementing Medicare.

The author's suggested design of a Medicare supplement benefit package is generally good. One conclusion I do not share, though. With proper administrative controls, prescription drugs can be a worthwhile inclusion in a Medicare supplement policy. My company's most widely held Medicare supplement offering includes a post-discharge drug benefit which pays 80 per cent of a pharmacy's usual charge after the insured has met a $20 deductible. Three factors are at play here to keep the cost of claims and administration reasonable. First, the benefit is available only for six months following a covered inpatient stay. Secondly, the deductible eliminates the "nickel and dime" claims for patients who require short-term medication immediately following discharge. Finally, since record keeping and claim submission is the insured's responsibility, substantial underreporting can be expected. With these three factors in place, prescription drugs account for less than ten per cent of total claims cost for this Medicare supplement policy.

The author concludes the paper with a discussion of the range of regulatory attitudes and future trends which sets forth a perceptive analysis of the various regulatory viewpoints and the crucial interdependency of the insurer and the regulator objectives. He also raises the question of the appropriate amount of government regulation.

It is doubtful that very few individual health insurance practitioners would quarrel with the author's observation that the individual health insurance market has eroded over the past decade due to expansion of government health care and income protection programs and expanded regulatory activity. Perhaps of equal importance if not more significant is the ever expanding role of group health insurance. In addition to increase in the number of people covered by group policies, the scope and level of benefits have exploded over the past decade. The following table summarizes recent experience for insurance companies.

Insurance Company Statistics
Group and Individual Health Insurance
Number of Persons Covered (000)

Type of Coverage	Group			Individual and Family Policies		
	1977	1967	% Inc.	1977	1967	% Inc.
Hospital	89,219	71,454	25	28,687	24,619	17
Surgical	91,904	72,038	28	14,409	17,603	(18)
Physician	88,818	61,028	46	10,964	8,541	28
Major Medical	101,925	67,394	51	6,101	4,552	34
Disability						
Short Term	28,176	24,805	14	14,302	13,188	8
Long Term	12,481	3,722	235	6,883	3,056	125
Dental	32,216	2,330	1383	----	----	---

Source: Source Book of Health Insurance Data 1978-79
Health Insurance Institute

In conclusion, Mr. Habeck's excellent paper with a well documented List of Readings brings in focus the expanding influence of governmental regulation on designing, underwriting, and marketing individual health insurance. His concluding observation of the importance of greater policyholder awareness and education is perhaps one of the best weapons against further government encroachment into the individual health insurance market.

TITLE: CREDIBILITY AND SOLVENCY

AUTHOR: Mr. Philip Heckman

Mr. Heckman is Senior Actuarial Analyst with CNA in Chicago. He earned his PhD from the University of Chicago in 1969 and his ACAS in 1978.

REVIEWER: Ms. Janet Fagan

Ms. Fagan is Assistant Secretary, Actuary for The Home Insurance Company. She received her FCAS in 1978. Janet earned a BS degree at State University of New York-Stony Brook and her MS degree from the University of Wisconsin.

I. Introduction

An important fact brought into sharp focus by the papers submitted to last year's program is that a healthy insurance enterprise not only must produce adequate earnings but must do so steadily and predictably. That is to say: the risks that confront the enterprise must be tightly controlled. Management must steer a course mindful not only of reasonable expectations but also of unforeseen deviations therefrom. This is true of the insurance enterprise in particular because of the urgent character of its obligations to policyholders who have suffered insured losses.

Last year's contributors also made clear that the risks inherent in the operation of the insurance enterprise cannot be isolated or sequestered: every source of financial uncertainty either in the environment or internal to the enterprise inevitably impinges on its ability to make that essential transaction - the loss payment. One of the most conspicuous areas where control of financial risk is essential is in the pricing of the business. It is misleading to think - let alone say - that the actuary can exert detailed control of the pricing process by prescribing and enforcing a static, cost-plus formula. The keyword here is "process". Pricing is a dialectical process engaging the potential insurer and insured within the market environment. Neither participant in the process can count on having complete control. Much that passes under the name of "pricing" is what engineers call "costing" and provides the insurer with the indispensable knowledge of the region in which to break off negotiation: between the economist's "break-even" and "shutdown" points. Successful pricing involves using the results of the costing formulas with art and judgment to anticipate the condition of the market. This is the first source of risk which the enterprise must face: that of misjudging the market and rushing in at a price at which too little (or too much) can be sold.

A second source of risk is that involved in estimating the aggregate cost of benefits which have been purchased in the market. Minimizing this risk is the chief endeavor of the pricing actuary. His chief ally is the law of large numbers; his chief foe is the passage of time: the time it takes for data to reach him and the time that elapses between his calculation and the losses it purports to forecast.

A third source of risk arises from a unique characteristic of insurance: the true ultimate meaning of the insurer's contractual obligation is as varied as the insureds themselves: no two alike. This is the joint domain of actuary and underwriter and this risk of inhomogeneity within the aggregate is attacked by such devices as classification ratemaking and experience rating, and, after all other bolts are shot, underwriting judgment. Since the underwriting decision is essentially yes/no, it is less risky to the extent that the actuary can propose a price that is graduated to be appropriate to the individual applicant. That is to say, as we all know, the underwriter is on the razor's edge and needs all the help he can get in deciding whether or not to expose the insurer's assets to risk. It is on this problem of resolving the inhomogeneity of the actuarial aggregate, in a way that minimizes the insurer's financial risk, that we shall concentrate in this paper: the problem of credibility. We have been careful in setting the stage because we shall follow the same course mathematically as we have conceptually: we shall define the problem in a way that leaves it distinct, as much as possible, from the other great problems of market strategy and aggregate valuation and forecasting. On the other hand we shall argue strenuously that credibility in classification ratemaking and credibility in individual risk rating are essentially the same problem on different scales, differing only in the relative importance of small sample corrections, and are not distinct fields of study as traditional actuarial practice would lead us to suppose.

This program will not lead us to present new work on credibility or even the most advanced. Our emphasis will be on simplicities rather than complexities, on the robust rather than the ethereal. For example it will be emphasized that it is more important for a credibility scheme to be well tuned - adjusted to minimize practical risk - than that it should embody a sophisticated model with the most general assumptions. And we shall spend significant space discussing the estimation problem for the parameters of the credibility model. Further we shall restrict ourselves to quadratic risk: that is, we shall consistently characterize risk in terms of the second moments of the financial variables. This is time-honored practice in risk theory, and we adopt it here on the premise that it is better to have a primitive definition that one can work with than to waste time and effort refining it to something unworkable.

We shall attempt to focus attention on the conceptual streams which feed into current practical work on credibility, particularly that going forward at I.S.O. These two converging streams are that initiated by Bülhmann and Straub in their landmark 1972 paper (1), which established the requirements for practical application of Bayesian Credibility, and that began by Charles Stein in his work on estimates with minimal quadratic loss (2).

In a concluding section we shall examine the benefits which flow from a credibility program tuned and managed for optimal risk control. In particular we shall see the implications for market strategy, social responsibility, and financial management.

Finally, in a series of technical appendices, we review a number of simple stochastic models for the loss ratios of a class of individual risks (or a group of classes). We show how these models all lead to a simple model for the covariance structure of the loss ratios over time and how this covariance model leads naturally to the familiar credibility formula as a predictor of future loss ratios. A final appendix shows an application of a simplified model to numerical data.

II. Decomposing the Chain of Risks

In the introduction we identified the sources of risk in the pricing process and declared that these can be dealt with separately. This is a declaration of hope because these risks are dealt with separately in practice. Indeed, the structure of the pricing process makes such separate treatment almost a practical necessity. Market risk is dealt with at the executive levels of underwriting, marketing, and financial management using as input knowledge of market conditions, aggregate loss costs, and administrative expenses. These deliberations can be knocked badly askew if aggregate loss cost projections are inaccurate.

These projections, in turn, are of diminished utility if the means of dealing with inhomogeneity in the aggregates are inadequate. This is particularly true in a competitive environment where the insurer's competitors are using accurate class rates and experience rating plans to reach pricing and underwriting decisions. In such a situation a company with a crude pricing apparatus will find that its aggregates are in fact unstable, with better business leaving and worse business coming in.

Thus we see that each level of the pricing process depends on the adequacy of risk control at the next, more detailed level. The question confronting us on the technical plane is whether or not these levels can be modeled independently:

Can the computation and trending of the aggregate be carried out without regard to its composition in terms of classes and in such a way that classification relativities can be defined free of secular trend?

Can the rate for a particular classification be further split by an individual experience model utilizing only information pertaining to that classification?

The first question can be answered in the affirmative if we give due consideration to two issues: first, the obvious problem encountered in tracking any aggregate through time, that of shifts in the makeup of the insured population by class; second, the rather more arcane problem that arises because the aggregate from which we split the classification relativities is not exactly known but is subject to sampling uncertainties.

The problem of shifting risk populations can be dealt with by a method familiar to demographics: define or project a standard risk population and restate historical aggregates in terms of it. Once this is done, there is an extensive repertoire of trending models that can be applied to the adjusted numbers - pure premiums are the most likely candidates for such treatment. Not to get too far from our main thrust, we should remark that the ubiquitous least squares trend line is neither the simplest nor - it is likely - the most robust of these models. The process has twin but complementary goals: to achieve an accurate projection of aggregate rate level and to allow classification relativities to be treated as a bundle of stationary time series, that is, to allow the problem of assigning relativities to be detached from that of projecting secular movements.

The more arcane problem; that the relativities will be based on aggregate numbers which are themselves uncertain, will be addressed in succeeding sections.

The second question, whether individual risk relativities can be split reasonably from the classification relativities, is somewhat more subtle and elusive. Classification rates are dependent on objective characteristics of the individual risks. The information used is limited to what can be recorded economically and more or less reliably. The most efficient use of this information is a matter beyond the scope of this paper. But there can be no dispute that, no matter how efficiently classification variables are utilized, a residuum of useful information will remain in the form of individual risk experience records. It has long been recognized that this information can be utilized to refine the pricing of individual contracts. How to do this most efficiently and, indeed, how to define efficiency in this situation is one of our principal topics. Some difficult problems make this task less than straightforward. One of the hardest is how to define and establish the identity of an individual risk whose characteristics may be changing over time. We shall have more to say about this later.

We have advanced about as far as we can using words only. In the next section we shall invoke a particular stochastic model to illustrate our previous remarks.

III. A Stochastic Model for Credibility

Without attempting to settle all the questions raised in the last section, we shall propose here a simple and plausible covariance model for the behavior of individual risk relativities within a given class. Assuming stationarity and a stable covariance structure, we shall arrive at a best linear unbiased estimator for the individual relativity, in terms of past experience, presuming a known aggregate value. The model has a classification structure also, and the best method for computing class rates and assigning credibility emerge from the model as well.

A. Variables and Notations

For the sake of concreteness we shall couch the model in terms of loss ratios. This allows closer comparison with the pioneering work of Buhlmann and Straub. The loss ratio is also a recognized index of equity, and the goal of most experience rating plans is to equalize loss ratios - properly defined - across all insureds, insofar as possible.

The variable of interest, then, we shall denote as

$$\tilde{x}_{Aa}(t),$$

the loss ratio for the a-th risk in class, A, during period, t. The superposed tilde denotes treatment as a random variable. For convenience, the time label, t, will be reckoned backwards, and we shall attempt to forecast experience in period, 0.

The only assumptions we shall need to characterize the variable for our purposes apply to the first two moments:

$$E[\ \tilde{x}_{Aa}(t)] = \xi, \tag{1}$$

$$E\left[(\tilde{x}_{Aa}(t)-\xi)(\tilde{x}_{Bb}(u)-\xi)\right] = \delta_{AB}\,\sigma^2\left[\frac{1}{K} + \delta_{ab}\left(\frac{1}{K_A} + \frac{\delta_{tu}}{W_{Aa}(t)}\right)\right]. \tag{2}$$

The first condition is the global expectation, without reference to class membership or individual experience records. The notation, δ_{ij}, is the Kronecker delta which is unity if the indices match, zero otherwise. The parameter, K, measures the homogeneity of the entire aggregate; K_A measures the homogeneity of the particular class, A. The quantity, $W_{Aa}(t)$, measures the relative statistical weight of individual risk experience, thus characterizing the size of the risk. For practical purposes, we may take it to be the individual risk premium computed at the aggregate rate before classification and experience rating. Statistical research with copious and accurate data might yield a better choice, but experience teaches not to hold one's breath but to proceed as best one can.

B. <u>Credibility Estimator</u>

The next step in our conceptual process is to propose an Ansatz for the form of the best estimator for predicting individual risk experience in time period, zero. This estimator should be:

1) linear in the aggregate and experience variables,
2) confined to experience within the class (since we have assumed a fixed, known aggregate),

3) unbiased, as our method of construction will ensure.

These requirements lead to the Ansatz,

$$\hat{x}_{Aa}(\theta)-\xi = \sum_{b,t} A_b^{Aa}(t)\left(\hat{x}_{Ab}(t)-\xi\right) + \tilde{\varepsilon}_{Aa}(o), \quad (3)$$

where $\tilde{\varepsilon}_{Aa}(o)$ is a random variable expressing the residual variation after our attempt at prediction. It is this variation that we wish to minimize in some sense by seeking an optimal form of the coefficients, A.

How best to go about this requires some thought. We might follow Mayerson's approach in his famous paper on Bayesian credibility (3) and minimize $E[\tilde{\varepsilon}_{Aa}(o)^2]$. This optimizes individual equity. More important than individual equity, however, is the insurer's aggregate risk in this class of business since the security of all policyholders depends in part on the control of this risk. This would lead us to seek a solution minimizing $E\left[\left(\sum_a W_{Aa}(o)\tilde{\varepsilon}_{Aa}(o)\right)^2\right]$. If there were indeed a conflict, we should have to choose the latter. The marvelous fact is that no conflict exists: the individual equity solution also minimizes the aggregate risk, as we shall demonstrate in Appendix B.

Let us now compute the expectation of the squared individual risk residual. In our abbreviated notation an asterisk denotes summation on the indicated index.

$$R_{Aa} \equiv \frac{1}{\sigma^2} E\left[\tilde{\varepsilon}_{Aa}(0)^2\right] = \frac{1}{K} + \frac{1}{K_A} + \frac{1}{W_{Aa}(0)} \tag{4}$$
$$- \frac{2}{K} A_*^{Aa}(*) - \frac{2}{K_A} A_a^{Aa}(*) + \frac{1}{K}\left(A_*^{Aa}(*)\right)^2 + \frac{1}{K_A}\sum_b \left(A_b^{Aa}(*)\right)^2$$
$$+ \sum_{bt} \frac{1}{W_{Ab}(t)}\left(A_b^{Aa}(t)\right)^2 + 2\lambda\left(A_*^{Aa}(*) - \sum_b A_b^{Aa}(*)\right) + 2\sum_b \mu_b \left(A_b^{Aa}(*) - \sum_t A_b^{Aa}(t)\right).$$

The extra terms with Lagrangian coefficients, λ and μ_b are added to enforce the sum constraints while treating the starred objects as independent variables. This model is solved in Appendix B. Here we merely state the solution:

$$A_b^{Aa}(t) = \delta_{ab} Z_{Aa}(t) + \left(1 - Z_{Aa}(*)\right)\frac{K_A Z_{Ab}(t)}{K + K_A Z_{A*}(*)},$$

$$\text{where} \quad Z_{Aa}(t) = \frac{W_{Aa}(t)}{K_A + W_{Aa}(t)}. \tag{5}$$

This leads to the prediction,

$$\tilde{x}_{Aa}(0) = \left(1 - Z_{Aa}(*)\right)\left\{\left(1 - \frac{K_A Z_{A*}(*)}{K + K_A Z_{A*}(*)}\right)\xi + \frac{K_A \sum_{bt} Z_{Ab}(t)\tilde{x}_{Ab}(t)}{K + K_A Z_{A*}(*)}\right\}$$
$$+ \sum_t Z_{Aa}(t)\tilde{x}_{Aa}(t) + \tilde{\varepsilon}_{Aa}(0). \tag{6}$$

The above is a beautiful result which should gladden the hearts of *all* actuaries. What it tells us is that, in the context of our covariance model, credibilities can be nested at different levels of aggregation, providing only that the aggregation is carried out properly.

Within this bristling expression, we may identify:

the partial credibilities, $Z_{Aa}(t)$, applying to the experience of risk, Aa, in period, t,

their complement, $(1-Z_{Aa}(*))$, applying to the class aggregate,

the experience estimate of the class aggregate,

$$\frac{1}{Z_{A*}(*)} \sum_{b,t} Z_b^{Aa}(t)\ \tilde{X}_{Ab}(t)$$

and its credibility, $\dfrac{K_A Z_{A*}(*)}{K+K_A Z_{A*}(*)}$, relative to the global aggregate, ξ.

All this falls out of the model and a simple Ansatz with no need for approximation.

This nested structure, with its easy identification and separation of individual and aggregate elements, is the structure that has always been assumed in actuarial practice. It is gratifying to know that it also has an appealing axiomatic justification in the context of risk and estimation theory. One loose end remains, however: the parameters K and K_A, one for the aggregate, one for each classification in the scheme. These must be estimated from the data available and this considerable problem will occupy us in the next section. Before embarking on this project, however, let us consider some straightforward extensions of our model.

C. Extensions of the basic Model

It is fairly easy to compute the consequences of relaxing the assumption that the aggregate loss ratio, ξ , is known and fixed. Doing so, we are led to the predictive relationship,

$$\tilde{x}_{Aa}(o) = \sum_{Bbt} A^{Aa}_{Bb}(t)\,\tilde{x}_{Bb}(t) + \tilde{\varepsilon}_{Aa}(o),$$

with the bias constraint,

$$A^{Aa}_{**}(*) = 1$$

Note that the sum now extends outside the class, A. Also, the global mean, ξ, does not appear: the model tells us how to compute it from the experience data. The result is identical to equation (6), except that where ξ appears in (6) we have now

$$\hat{\xi} = \frac{1}{\zeta_{**}(*)} \sum_{Bbt} \zeta_{Bb}(t)\,\tilde{x}_{Bb}(t),$$

where

$$\zeta_{Bb}(t) = \frac{K_B Z_{Bb}(t)}{K+K_B Z_{B*}(*)} .$$

Another natural extension of the model is to take into account the problem raised in the last section: that the identity of an individual risk changes as time elapses and that the relevance of experience decays with age. Let us suppose that this decay is exponential in time. The indicated covariance model becomes

$$E\left[(\tilde{x}_{Aa}(t)-\xi)(\tilde{x}_{Bb}(u)-\xi)\right] = \delta_{AB}\left[\frac{1}{K} + \delta_{ab}\left(\frac{1}{K_A} + \frac{\delta_{tu}}{W_{Aa}(t)}\right)\right] e^{-\lambda|t-u|},$$

$$\lambda > 0.$$

This model is much more difficult to solve (and to apply in practice) than the model of equation (1). At this writing, the general solution is not available. We can, however, exhibit the partial credibilities in a simple, one-class model using two periods of data:

$$Z(1) = \frac{b\,W(1)\left[1 + \frac{W(2)}{K}(1-b^2)\right]}{K + W(1)\left[1 + \frac{W(2)}{K}(1-b^2)\right] + W(2)},$$

$$Z(2) = \frac{b^2 W(2)}{K + W(1)\left[1 + \frac{W(2)}{K}(1-b^2)\right] + W(2)},$$

where $b = e^{-\lambda}$.

This form is less than appealing, but not out of the question in a computerized experience rating system. The form becomes progressively more complex as more periods of data are introduced. It is seen that, as $b \to 1$ these expressions approach the more familiar partial credibilities defined with equation (5).

IV. Estimation

As we have remarked, our basic model, with the linear Ansatz, prescribes the form of the credibilities at all levels of aggregation and even tells us how to compute the most efficient linear estimates of the aggregates at each level. (Our results apply to a simple, cellular scheme of classification. More complex schemes, involving several class differentials, require separate study.) We have, however, to estimate the homogeneity parameters, and our preferred forms make that task singularly awkward. Regardless of difficulties, this task of estimation cannot be bypassed. Reliance on judgment produces an untuned credibility scheme that can do more harm than good. The administration of an experience rating plan is two expensive and time consuming to leave its risk-control aspect out of reckoning since it is just this that justifies the trouble and expense.

In this section we shall discuss several aspects of the estimation problem, starting with the algebraic estimation scheme of Bühlmann and Straub. We shall then discuss a more elementary approach: that of non-linear estimation by searching on the parameter space. We shall then conclude with some brief remarks on alternative estimation schemes associated with the name of the statistician, Charles Stein.

A. Algebraic Estimation

We use the term algebraic estimation to denote the use of statistics which can be computed directly and combined to yield a solution for the parameter in question. Following the method of Buhlmann and Straub in their 1972 paper (1), we may define the statistics, for our model defined in equation (1) of Section II,

$$\tilde{S}_A^{(1)} = \sum_{at} W_{Aa}(t)\left[\tilde{x}_{Aa}(t) - \bar{x}_{Aa}(*)\right]^2,$$

and

$$\tilde{S}_A^{(2)} = \sum_{at} W_{Aa}(t)\left[\tilde{x}_{Aa}(t) - \bar{x}_{A*}(*)\right]^2,$$

with

$$\bar{x}_{Aa}(*) = \frac{1}{W_{Aa}(*)} \sum_t W_{Aa}(t)\,\tilde{x}_{Aa}(t) \quad ,$$

and

$$\bar{x}_{A*}(*) = \frac{1}{W_{A*}(*)} \sum_{at} W_{Aa}(t)\,\tilde{x}_{Aa}(t),$$

for each class, A, and the global statistics,

$$\tilde{T}^{(1)} = \sum_A \tilde{S}_A^{(2)} ,$$

$$\tilde{T}^{(2)} = \sum_{Aat} W_{Aa}(t)\left[\tilde{x}_{Aa}(t) - \bar{x}_{**}(*)\right]^2 .$$

Using the defining equation of the model, we find,

$$E[\tilde{S}_A^{(1)}] = \sigma^2\left\{N_{A*}(*) - 2 + \sum_{at}\left(\frac{W_{Aa}(t)}{W_{Aa}(*)}\right)^2\right\},$$

$$E[\tilde{S}_A^{(2)}] = \sigma^{-2}\left\{\frac{W_{A*}(*)}{K_A}\left[1 - \sum_a\left(\frac{W_{Aa}(*)}{W_{A*}(*)}\right)^2\right] + N_{A*}(*) - 2 + \sum_a\left(\frac{W_{Aa}(*)}{W_{A*}(*)}\right)^2\right\},$$

$$E[\tilde{T}^{(1)}] = \sum_A E[\tilde{S}_A^{(2)}],$$

$$E[\tilde{T}^{(2)}] = \sigma^2\left\{\frac{W_{**}(*)}{K}\left[1 - \sum_A\left(\frac{W_{A*}(*)}{W_{**}(*)}\right)^2\right] + \sum_A \frac{1}{K_A}\left[W_{A*}(*) - \sum_a\left(\frac{W_{Aa}(*)}{W_{A*}(*)}\right)^2\right] + N_{**}(*) - 1\right\}.$$

The important thing to notice about this system of expressions is that it can be solved for the K_A's and K in terms of the expectations and the exposures and risk counts (the N's). Replacing the expectations by the actual sums of squares then gives asymptotically unbiased estimators for the parameters, K and K_A, (all A).

For ease of computation, these estimators are unmatched. They do, however, possess some drawbacks. First, estimation of the K's involves taking ratios of sums of squares, and small sample bias may be substantial. This problem can be circumvented only by introducing a hypothesis for the detailed form of the distributions - something we would rather avoid. Second, the estimation scheme involves computing the aggregates in a way we have already seen to be suboptimal. Third, and most important from a practical viewpoint, the solution for the K's involves taking the difference of sums of squares with appropriate (positive) coefficients. This can, and sometimes does, lead to a negative estimate for $\frac{\sigma^2}{K}$ or $\frac{\sigma^2}{K_A}$, as should never be since these are components of variances and must be positive to make sense. One recourse, in such a situation is just to set $\frac{\sigma^2}{K}$ to zero ($K \to \infty$), a solution corresponding to perfect homogeneity. Considering the extreme ease of computation in this scheme, along with the countervailing drawbacks, it seems sensible, in the age of the electronic computing machine, to use these estimates as starting values in a more accurate iterative scheme. This brings us to our next topic.

B. Non-linear Estimation

On this subject, there is always much less to be said than to be done. This is the brute force method for estimating non-linear models when short cuts prove unacceptable. Given a model for the individual risk residuals, the method is applied in two stages: first define a loss function, usually a linear combination of the squared residuals, which we wish to minimize: second, vary the parameters of the model, following an efficient search algorithm, until no further improvement in the loss function can be achieved. Both stages are fraught with difficulty, and a discussion of search algorithms is entirely beyond the scope of this paper. Although refinements are possible, and perhaps desirable, a choice for the loss function of the form,

$$\tilde{\mathcal{L}} = \sum_{Aa} W_{Aa}(o)\, \tilde{\varepsilon}_{Aa}(o \mid \hat{K})^2$$

will usually serve. (Here $\hat{K}$ refers to estimated values of K and all the K_A, one for each class.) For the purpose of evaluating the efficiency of the model, $\tilde{\mathcal{L}}$ may be compared to

$$\tilde{\mathcal{L}}_o = \sum_{Aa} W_{Aa}(o)\left(\tilde{x}_{Aa}(o) - \hat{\xi}\right)^2$$

where $\hat{\xi}$ is the assumed aggregate value or the estimated value defined in Section III-C. The efficiency can be defined as

$$E = 1 - \frac{N}{N-n-1}\left(\frac{\tilde{\mathcal{L}}}{\tilde{\mathcal{L}}_o}\right),$$

where N is the number of risks and n is the number of classes. Small sample corrections have been applied to account for the number of parameters (K's) that must be estimated. If the estimated value, $\hat{\xi}$, is used, it is preferable to minimize $\tilde{\mathcal{L}}/\mathcal{L}$.

If there are many classes in the model, it may be desirable to minimize in stages, guessing at the aggregates, optimizing K_A within each class, combining these results to find a value of K, recomputing the aggregates, and iterating until the loss function ceases to improve. Such an approach was suggested by Morris and Van Slyke in their recent paper. It is impossible to specify more detail outside a practical situation, but a simplified treatment using data from Bühlmann and Straub's 1972 paper is described in Appendix C.

C. Stein Estimators and Small Sample Corrections

In recent years, some very interesting developments have taken place in the statistical community which have an important bearing on actuarial credibility concepts. We refer to work in estimation theory initiated by Charles Stein (2). There is not space for thorough discussion, but the basis of these estimation techniques lies in shrinking the individually estimated means for an inhomogeneous ensemble toward some prior estimate or guess at the true means. The shrinkage factor varies according to the expected variance of the individual member and is chosen to minimize a quadratic loss function summarizing the deviations of the entire ensemble.

If all this sounds familiar, it is no wonder. If the center of the shrinkage is taken to be an estimate of the common aggregate mean, the procedure satisfies the definition of an actuarial credibility technique. Further the explicit goal of the estimation, minimization of quadratic loss, is identical to the goals of Bayesian credibility analysis. Indeed, a recent paper under review by the adventurous and innovative Subcommittee on Credibility at ISO, treats a scheme nearly identical to that proposed in Section III of the present work. The paper is that of Morris and Van Slyke (4), and their formal results are identical to ours except that the individual risk credibility turns out to have the form,

$$Z_a = 1 - \frac{n-3}{n} \frac{K}{K+W_a}$$

where n is the number of risks in the sample. As n becomes large, this approaches the usual Bayesian result. The disturbing point is that the small-sample correction is not derivable from their formalism, but must be introduced freehand, mainly to compensate bias introduced due to the use of maximum likelihood estimation assuming normal distributions.

It is interesting to compare with the Bayes-Mayerson formulation that we used in Section III. In this paper we let the minimal risk condition dictate the computation of aggregates, and the small-sample corrections did not appear. If, on the other hand, we had used straight premium weighting, as in the

Buhlmann Straub procedure, the model would have bristled with small-sample corrections of the clumsiest sort. These corrections are eliminated because they are built into the aggregates which are not fixed but slide around as the homogeneity parameters are varied. Since Morris and Van Slyke compute their aggregates in the same way, their small-sample correction must arise entirely from the method of estimation. Their method of maximum likelihood assuming normal distributions imposes a level of hypothesis which one would like to avoid in insurance statistics, if at all possible. It seems that a quest is in order for a method of estimation which does not have these drawbacks. We propose the method advanced earlier in this section as a possible candidate, deserving more study.

V. Conclusion

In conclusion, we shall review our main points, remark on some connections with financial theory, and close with a brief discussion on the systematics of practical use of these ideas.

Some of the lessons we have learned deserve underlining:

- A credibility scheme for pricing that is well-tuned for risk control can have significant impact on the marketing, underwriting, and financial aspects of an insurance business, on marketing through improved individual equity, on underwriting by allowing greater flexibility in the choice of insured and more rational review of individual risk experience, on financial by decreasing uncertainty in underwriting results consequently allowing more lucrative employment of the investment portfolio.

- An untuned credibility scheme cannot be expected to do any of these things optimally, or even reliably. In particular, the "classical" claim count credibility used in classification ratemaking is very rigid and difficult to tune and poorly adapted to risk control requirements.

- A credibility scheme cannot be expected to follow trends accurately, nor to prescribe the most efficient system of classification. These must be treated separately, though the credibility scheme may produce technical corrections to the global and class aggregates.

- Our proposed model, extended to estimate the overall aggregate corrections, dictatesthe splitting out of relativities from the top to the bottom of the process, including class credibilities, and has the convenient nested structure always assumed in class and experience rating.

- Although the question of estimation is far from settled. The recent work of Morris and Van Slyke shows that it is possible, by an iterative scheme, to go beyond the limitations of the Buhlmann-Straub estimators. Extensive lore on non-linear estimation techniques is available as a guide in such work.

- A great deal of credit is due to the ISO Subcommittee for bringing these ideas close to practical application.

It is perhaps enlightening to examine our results in the light of some aspects of financial theory which came to the fore in last year's call paper program. In his prize-winning paper, (5), Robert Butsic brought forward the distinction between systematic and non-systematic risk: non-systematic risk is subject to the law of large numbers and can be reduced, relatively, by increasing the book of business; systematic risk cannot. The significance of this distinction to pricing is that an insurer can load its rates to cushion against systematic risk since it affects companies of all sizes in the same way. In a market with efficient price competition, however, it is not possible to load rates against non-systematic risk since smaller companies will price themselves out of the market. Inspection of the covariance assumptions of our model shows that it combines both elements: inhomogeneity is a key element of systematic risk, and credibility theory is aimed directly at controlling it. The lesson is clear: well-tuned pricing credibilities can give a company a real market advantage by allowing it to run safely on slimmer margins. The result is a more financially efficient operation from which all parties benefit.

Some final clarifications are due on how such ideas can be put into practice. We shall draw a broad outline, omitting - unfortunately but necessarily - some important questions of detail regarding data consistency and other matters.

First, tuning a credibility scheme requires data; where are they to come from? The answer is that companies and bureaus are awash in such data. The main problem is that they throw it away too soon - before the feedback loop is complete. A proper credibility scheme for ratemaking and individual risk pricing must be an information <u>system</u> and not just a set of assumptions, however apt.

A complete feedback loop implies correlation of premium and loss data on the individual risk level. This is done in the <u>operation</u> of all experience rating plans. However, these data must also be kept around for tuning the model for the next rating cycle and testing the stability of the optimal parameter values. Such tuning will yield aggregate corrections, class relativities, and a simple experience rating formula, using the K_A tabulated for each class, A, in which all complications are kept behind the scenes and away from the rating clerk's desk.

I wish to thank my colleague Glenn Meyers for keeping me up to date on happenings at ISO and developments in the literature. Special thanks go to Barbara Dudman for typing the manuscript.

Appendix A: Detailed Stochastic Models

Bayesian models are usually presented in terms of underlying variables, which can be observed only indirectly through Bayesian inference. We have not done so here because a variety of models leads to the covariance structure which determines the credibilities, and there is no need to commit oneself to a particular model. To illustrate this we show two simple models, one additive, one multiplicative which lead to the same structure.

1. Additive Model

Let $\tilde{x}_{Aa}(t) = \xi + \tilde{\xi}_A + \tilde{\eta}_{Aa} + \tilde{\zeta}_{Aa}(t)$, where

ξ is the aggregate mean;

$\tilde{\xi}_A$ is the systematic departure of class A's experience from the aggregate.

$\tilde{\eta}_{Aa}$ is the departure of risk a's experience from the mean of class, A.

$\tilde{\zeta}_{Aa}(t)$ is the random fluctuation producing risk a's actual experience in period, t.

The random variables , $\tilde{\xi}$, $\tilde{\eta}$, $\tilde{\zeta}$, have unconditional mean, zero, and are mutually independent. Their covariance structure is

$$E[\tilde{\xi}_A \ \tilde{\xi}_B] = \frac{\sigma^2}{K}\delta_{AB}, \quad E[\tilde{\eta}_{Aa}\tilde{\eta}_{Bb}] = \frac{\sigma^2}{K_A}\delta_{AB}\delta_{ab},$$

$$E[\tilde{\zeta}_{Aa}(t)\tilde{\zeta}_{Bb}(u)] = \frac{\sigma^2}{W_{Aa}(t)}\delta_{AB}\delta_{ab}\delta_{tu}.$$

These assumptions lead to

$$E\left[\left(\tilde{x}_{Aa}(t)-\xi\right)\left(\tilde{x}_{Bb}(u)-\xi\right)\right] = \sigma^2 \delta_{AB}\left[\frac{1}{K} + \delta_{ab}\left(\frac{1}{K_A} + \frac{\delta_{tu}}{W_{Aa}(t)}\right)\right],$$

our model from Section III.

2. Multiplicative Model

Let us apply exactly the same symbols and assumptions to a different model,

$$\tilde{x}_{Ac}(t) = \xi\left(1+\tilde{\xi}_A\right)\left(1+\tilde{\eta}_{Aa}\right)\left(1+\tilde{J}_{Aa}(t)\right)$$

The same covariance for this model is

$$\sigma^2 \xi^2 \delta_{AB}\left\{\frac{1}{K} + \delta_{ab}\left(1+\frac{\sigma^2}{K}\right)\left[\frac{1}{K_A} + \left(1+\frac{\sigma^2}{K_A}\right)\frac{\delta_{tu}}{W_{Aa}(t)}\right]\right\}.$$

This has the same algebraic structure as the additive model and, after redefinition of the parameters, will lead to the same credibility formula.

Appendix B: Solution of Least-Squares Conditions

Our purpose here is to record calculations too involved to include in the main text: the derivation of our chief result, equations (5, 6) of Section III and the proof of our assertion of harmony between individual equity and aggregate risk control.

1. Derivation of Credibility Formula

Requiring the risk function, R_{Aa}, of Section III, equation (4) to be stationary to variations of the coefficients, A, leads to the equations,

$$\frac{1}{K} A_{*}^{Aa}(*) - \frac{1}{K} + \lambda = 0 , \qquad (1)$$

$$\frac{1}{K_A} A_{b}^{Aa}(*) - \frac{\delta_{ab}}{K_A} - \lambda + \mu_b = 0 , \qquad (2)$$

$$\frac{1}{W_{Ab}(t)} A_{b}^{Aa}(t) - \mu_b = 0. \qquad (3)$$

Solution:

$$(3) \longrightarrow \quad A_{b}^{Aa}(t) = \mu_b W_{Ab}(t) ;$$

$$(2) \;\&\; (3) \rightarrow A_{b}^{Aa}(*) = \delta_{ab} + K_A(\lambda - \mu_b) = \mu_b W_{Ab}(*) ;$$

whence

$$\mu_b = \frac{\delta_{ab}}{K_A + W_{Ab}(*)} + \frac{K_A \lambda}{K_A + W_{Ab}(t)} ,$$

and

$$A_{b}^{Aa}(t) = \delta_{ab} Z_{Aa}(t) + K_A \lambda Z_{Ab}(t), \quad (4)$$

where

$$Z_{Aa}(t) = \frac{W_{Aa}(t)}{K_A + W_{Aa}(*)} ;$$

$$(1) \;\&\; (4) \longrightarrow A_{*}^{Aa}(*) = 1 - K\lambda = Z_{Aa}(*) + K_A \lambda Z_{A*}(*),$$

whence $\lambda = \left(1 - Z_{Aa}(*)\right) \dfrac{1}{K + K_A Z_{A*}(*)}$.

which on substitution into (4) gives

$$A_b^{Aa}(t) = \delta_{ab} Z_{Aa}(t) + \left(1 - Z_{Aa}(*)\right) \frac{K_A Z_{Ab}(t)}{K + K_A Z_{A*}(*)} .$$

This is equation (5) of Section III, which, with equation (3) of that section gives equation (6), the desired result.

2. Harmony of Equity and Aggregate Risk Control

Our assertion in the main text was that

$$E\left[\left(\sum_a W_{Aa}(0)\, \tilde{\varepsilon}_{Aa}(0)\right)^2\right]$$

is minimized by the same solution which minimizes

$$E\left[\ \tilde{\varepsilon}_{Aa}(0)^2\ \right],$$

so that there is no conflict between the two. This would be obvious except that these model residuals involve the same individual risk records through the classification aggregates and are not manifestly independent.

To proceed, we first define

$$\bar{A}_b^A(t) = \sum_a W_{Aa}(0) A_b^{Aa}(t),$$

so that the aggregate residual is

$$\sum_a W_{Aa}(0)\, \tilde{\varepsilon}_{Aa}(0) = \sum_a W_{Aa}(0)\left(\tilde{x}_{Aa}(0) - \xi\right) - \sum_{bt} \bar{A}_b^A(t)\left(\tilde{x}_{Ab}(t) - \xi\right)$$

and the expectation of its square is

$$\sigma^2\Big\{\frac{1}{K}W_{A*}(0)^2+\frac{1}{K_A}W_{A*}(0)^2+W_{A*}(0)-\frac{2}{K}W_{A*}(0)\bar{A}^A_*(*)-\frac{2}{K_A}\sum_b W_{Ab}(0)\bar{A}^A_b(*)$$
$$+\frac{1}{K}\bar{A}^A_*(*)^2+\frac{1}{K_A}\sum_b\bar{A}^A_b(*)^2+\sum_{bt}\frac{1}{W_{Ab}(t)}\bar{A}^A_b(t)^2$$
$$+\;2\Lambda\Big(\bar{A}^A_*(*)-\sum_b\bar{A}^A_b(*)\Big)$$
$$+\;2\sum_b M_b\Big(\bar{A}^A_b(*)-\sum_t\bar{A}^A_b(t)\Big).$$

giving the equations,

$$\bar{A}^A_b(t)=M_b\,W_{Ab}(t)$$
$$\bar{A}^A_b(*)=W_{Ab}(0)+K_A(\Lambda-M_b)$$
$$\bar{A}^A_*(*)=W_{A*}(0)-K\Lambda$$

These can be solved by the same steps used earlier in this appendix, using the sum constraints; the result is

$$\bar{A}^A_b(t)=W_{Ab}(0)Z_{Ab}(t)+\sum_a W_{Aa}(0)\big(1-Z_{Aa}(*)\big)\frac{K_A Z_{Ab}(t)}{K+K_A Z_{A*}(*)},$$

which is easily recognized as

$$\sum_a W_{Aa}(0)A^{Aa}_b(t)$$

where the A is the solution derived earlier for the individual risk problem.

Appendix C: Estimation of Single Class Model with Data

It is desirable to test our model against data; however, data are hard to come by. We shall content ourselves by working with the tabulation presented in (1). This is a set of reinsurance data, gross premiums and excess loss ratios on a uniform, 'as-if' basis. We present these in the following table:

Treaty/Year:		4	3	2	1	0
1	w=	5.	6.	8.	10.	12.
	x=	0.0	0.0	4.2	0.0	7.7
2	w=	14.	14.	13.	11.	10.
	x=	11.3	25.0	18.5	14.3	30.0
3	w=	18.	20.	23.	25.	27.
	x=	8.0	1.9	7.0	3.1	5.2
4	w=	20.	22.	25.	29.	35.
	x=	5.4	5.9	7.1	3.1	5.2
5	w=	21.	24.	28.	34.	42.
	x=	9.7	8.9	6.7	10.3	11.1
6	w=	43.	47.	53.	61.	70.
	x=	9.7	14.5	10.8	12.0	13.1
7	w=	70.	77.	85.	92.	100.
	x=	9.0	9.6	8.7	11.7	7.0

In the above table, w stands for the premium and is used as a statistical weight, while x stands for excess loss ratio (in percents).

We have used the data from years 4 through 1 to predict experience in year zero. The collective consists of a single class and our model is

$$\tilde{x}_a(0) = \sum_{t=1}^{4} Z_a(t)\,\tilde{x}_a(t) + \left(1-Z_a(*)\right)\hat{\xi} + \tilde{\varepsilon}_a(0),$$

$$Z_a(t) = \frac{W_a(t)}{K+W_a(*)}\ ; \quad 1-Z_a(*) = \frac{K}{K+W_a(*)}\ ;$$

$$\hat{\xi} = \frac{1}{Z_*(*)} \sum_{\substack{a=1,7 \\ t=1,4}} Z_a(t)\,\tilde{x}_a(t).$$

The asterisk denotes summation on treaty, a=1, 7, or on year, t = 1, 4.

The model has a single parameter, K. Its structure, however, involving the aggregate estimator, $\hat{\xi}$, is too complex for the non-linear fitting modules of most statistical software packages. In this instance, though, it was a simple matter to construct a Fortran program to scan across prescribed parameter values and write out the value of the loss function,

$$\tilde{\mathcal{L}} = \sum_{a=1,7} W_a(0)\, \tilde{\mathcal{E}}_a(0)^2$$

and other statistics.

The results are shown in Exhibit 1 and are somewhat surprising. Buhlmann and Straub, using their algebraic estimators, arrived at a value, $K_{BS} \doteq 17.3$. Optimizing the loss function in our model gives a value, $K \doteq 0.1$, or, essentially, zero. This is the same as saying that the members of the collective are so diverse that the aggregate information is of no use, and each insured is rate should be determined from his experience alone. Statistically, the Bühlmann-Straub result is not much different from this since 17.3 is a good deal smaller than the cumulative premium on any one treaty. Unfortunately, the collective is very small, and any estimate will be highly uncertain.

Exhibit 1

Fitting of Single Class Model

Parameter, K	Loss Function	Sliding Mean
- 10.0	3631.13	--
- 1.0	3334.86	--
- 0.1	3332.94	--
0.0	3332.89	8.644
0.1	3332.87	8.646
1.0	3334.00	8.663
10.0	3425.40	8.787
∞	6431.49	9.496

Fitting Residuals

Treaty	K = 0.0	K = ∞	Weight
1	6.54	-1.80	12.
2	12.58	20.50	10.
3	0.31	-4.30	27.
4	1.80	-1.20	35.
5	2.17	1.60	42.
6	1.32	3.60	70
7	-2.83	-2.50	100.
Average	0.59	0.32	296.
Root Mean Square	3.36	4.66	

References

(1) H. Bühlmann and E. Straub, 'Glaubwürdigkeit fur Schadensätze', Mitteilungen Vereinigung Schweizerischer Versicherungsmathematiker, Band 70, Heft 1 (1970). Translated as 'Credibility for Loss Ratios' by C.E. Brooks, ARCH (1972).

(2) W. James and C. Stein, 'Estimation with Quadratic Loss', Proceedings of the Fourth Berkely Conference on Mathematical Statistics and Probability, Vol. 1, Berkely, University of California Press (1961), pp. 361-79.

(3) A. Mayerson, 'A Bayesian View of Credibility', PCAS LI (1964). p. 85.

(4) C. Morris and O. Van Slyke, 'Empirical Bayes Methods for Pricing Insurance Classes', Proceedings of the Business and Economic Statistics Section, American Statistical Association (1978).

(5) R. Butsic, 'Risk and Return for Property and Casualty Insurers', Total Return Due A Property Casualty Insurance Company, C.A.S. (1979), p. 52.

CREDIBILITY AND SOLVENCY

by PHILIP HECKMAN

Reviewed by JANET FAGAN

Much work has been done in the past few years on the applications of Bayesian credibility to insurance pricing. This work has been born of necessity due to the failure of "classical" credibility theories. Recent work by Bühlman and Straub as well as Morris & Van Slyke incorporate the Bayesian concept of utilizing as much information available from historical data as possible in predicting behavior for a segment of a population.

The Bühlman-Straub work develops a method for individual risk rating using a credibility scheme for each risk based on historical data of the risk and the total population. The work by Mr. Heckman is an attempt at extending this approach to the problem of individual risk rating within the classification ratemaking exercise. In so doing he emphasizes the point that individual risk credibility can be approached in the same way as class credibility and he derives a very neat nested credibility structure. A second point which is driven home by this paper is the extent of work which remains to be done in investigation of suitable and workable estimation schemes.

The Model

Mr. Heckman begins his work by defining a model which assumes a certain structure for the first and second moments of the loss ratio distribution function as follows:

$$E\left(\tilde{x}_{Aa}(t)\right) = \zeta \;,$$

$$E\left[(\tilde{x}_{Aa}-\zeta)\,(\tilde{x}_{Bb}-\zeta)\right] = \delta_{AB}\,\sigma^2\left[\frac{1}{K} + \delta_{ab}\left(\frac{1}{K_A} + \frac{\delta_{tu}}{W_{Aa}(t)}\right)\right].$$

This model structure assumes a global mean, though this is later shown to be an unnecessary assumption. The covariance can be given the following verbal interpretation. Classes are assumed to be independently distributed. Thus, there is no contribution to the covariance unless the two risks belong to the same class, therefore K is a measure of the homogeneity of the class. Similarly, K_A is a measure of the correlation of an individual risks' experience over time under the assumption that one risks' experience is independent of anothers (ie: no contribution to covariance between two risks). The $W_{Aa}(t)$ represents a random component of the model and σ^2 is a constant.

By defining $\tilde{X}_{Aa}(0) - \zeta = \sum_{b,t} A_b^{Aa}(t)\left(\tilde{X}_{Ab}(t) - \zeta\right) + \tilde{\varepsilon}_{Aa}(0)$

where $b \in A$, a risk function can be defined as $\frac{1}{\sigma^2} E\left(\tilde{\varepsilon}_{Aa}(0)^2\right)$

It is this function which is to be minimized to arrive at the optimal set of coefficients $A_b^{Aa}(t)$. Mr. Heckman asserts that the solution which minimizes the individual risk also minimizes the global risk. Unfortunately, this result flows directly from this covariance model structure and is not true in general. The resulting nested credibility structure is very neat.

The credibility for a risk at time t is calculated to be $\frac{W_{Aa}(t)}{K_A + W_{Aa}(*)}$

and the class aggregate is a credibility weighted average of class experience with credibility

$$\frac{K_A Z_{A*}(*)}{K + K_A Z_{A*}(*)}$$

This workup is quite similar to the Bühlman-Straub methodology and the major obstacle now becomes the estimation of the parameters K and K_A.

Estimation of Parameters

A reasonable estimation procedure is presented by Mr. Heckman but nothing is said regarding the speed at which such an iteration converges, nor is it clear that it must converge in all cases. Somehow building a theoretical model with all its niceties and then having to resort to a trial and error approximation routine seems less than perfect. It is in the area I would like to see some work done.

Mr. Heckman refers to the small sample corrections introduced in the Morris & Van Slyke work as being "gratuitously introduced". At this point I must take exception, the correction term is required due to the estimation procedures used. It is true that the current work does not require such correction factors but this is due to alternative, not superior, mathematics.

Implementation

At the end of the paper, Mr. Heckman asserts that the K and K_A values can be updated on a regular basis by bureaus who are "awash in" in the required data. While it is true that the operation of an experience rating plan is based on the correlation of premium and loss data for an individual risk, only Workers' Compensation experience rating modifications are computed from bureau data. The Statistical Plan for this line is specifically designed so that the unit report data is on a policy basis. For other lines (ie: ISO lines) this level of detail is not possible nor is it necessarily desirable. Data for these lines is not compiled on a policy basis and cannot be retrieved on such a basis from bureau files.

Returning to Workers' Compensation, the only line where such a scheme may apply, the experience rating plan generates a single modification. The proposed scheme would generate separate credibilities by class for each risk thus adding a great deal more complexity to the plan. I suspect it would also add alot more variability to an individual risks' modification. Whether this is desirable is open to question.

All credibility work I have seen done by the Bayesian method in the last few years has concentrated on a one way class scheme. We are nearing the point of workable credibility methodology using this approach and it is time for thinking about the much more difficult problem of 2-way schemes. Hopefully more work will be done this year and next on both problems.

TITLE: PRICING FOR CORPORATE OBJECTIVES

AUTHOR: Mr. Frank J. Karlinski

Mr. Karlinski is Actuarial Director with Prudential Property and Casualty Insurance Company. He received his FCAS in 1977 and is a member of the CAS Examination Committee.

REVIEWER: Mr. Robert A. Anker

Mr. Anker is Vice President-Corporate Planning at the American States Insurance Company. Bob received his FCAS in 1972 and is a member of the American Academy of Actuaries. He has served on a number of committees including the Education Committee (Vice Chairman) of CAS, Public Relations Committee of CAS and Committee on Property and Liability Insurance-Financial Reporting Principles of the American Academy of Actuaries.

The Determination of a Profitable Rate Level

The exercise of determining a rate level for some future period usually takes place within the context of an already existing rate level. Some amount of experience is usually available under the current rates and under other older rate levels which can be adjusted to current levels through a mathematical process which is familiar to us all. In any event, data is usually available for at least a somewhat credible determination of the adequacy of the current rates. So ordinarily the ratemaking process is actually a rate review process as the results of the analysis are usually expressed in terms of changes to the present rates. When the process is considered as a review of current rates, major steps in that process naturally follow. They are:

1. Determination of the adequacy (or excessiveness) of the present rates for the present time.[1]
2. Identification of the perceived differences between the present time and the future time and quantification of those differences.
3. Translation of the results of steps 1 and 2 into changes to the current rates to create adequate but not excessive rates for some future period.

Step 1 requires that experience period premiums and losses be adjusted to reflect the current levels of premium collections and loss incurrments.[2] This step requires judgments as to, among others,

[1]Possibly it would be more precise first to speak of the adequacy of current rates for the time represented by the experience, but that would just slow us down without adding significantly to the discussion.

[2]The reader is advised not to attempt to find "incurrments" in any standard English dictionary. Instead let us define by analogy: pay is to payments as incur is to incurrments.

data sufficiency and accuracy. This type of judgment deals with somewhat known quantities and events of the past and present rather than with predictions of future conditions or events. Such judgments are usually suseptable to a rather objective deductive type of reasoning. Let us refer to such judgments as deductive judgments.

In step 2 we compare the present with the future and therefore we must make judgments about events that have not yet occurred and conditions that may not yet exist. The usual procedure for doing this involves the examination of phenomena of the recent past such as inflation rates and frequency trends and extraction from those phenomena of inferences about the future. Let us therefore call those types of judgments inferential judgments. Inferential judgments usually weigh quite heavily in the amount of trend that will be used to bridge the gap between the present and future periods.

Step 3 involves the application of trends and other perceived differences between the present and future periods to the current rate level, with adjustment for the current inadequacy or redundency, to derive the required rate level for the future period. We now have the basis for a new set of rates which, if our reasoning, assumptions and inferences are good, will bring us to within epsilon (defined by chance variation, the law of large numbers and luck) of a profitable position in the future period.

But let us look at step 2 once more. There we identified the differences between the current and future periods. However, there are differences that could not possibly have been evaluated at that time; they are the differences caused by the change in rate level itself determined in step 3. These differences could include the change in the company's competitive position with other companies writing in the same market place, changes to the types and amounts of coverages purchased by present and prospective insureds prompted by the movement in rates, and the amount and kind of new business that will be generated by the new rate level. We make both deductive and inferential judgments many times along the way to arriving at a rate level. Now we see that some additional judgments are necessary. We must decide whether the rate level that is a product of our assumptions and judgments is compatible with those assumptions and judgments. This final judgment is an inferential one in that it involves prediction of the future rather than deductions about the past or present. For example, we may have assumed in step 2 that our own book of business would experience a frequency trend which does not differ from that of the rest of the industry, i.e. that the relative quality of our book of business will not change between the present and future periods. But given the rate level change derived in step 3, can we expect to be able to attract risks that are as good as the ones we have been attracting with the current rate level? If the rate change indication is for a substantial increase, that assumption could be questioned, especially if the competition is not expected to take similar rate action. Or suppose the data indicate a substantial decrease in rate level. The resulting improvement in competitive position could be enough to cause a significant increase in the number of new risks. This influx of new business with its usually higher loss ratios may contradict our assumptions about the new/renewal split in the future and could create an inadequacy in the indicated rate level. A small judgmental reduction in the indicated decrease could bring that rate level into equilibrium with its underlying assumptions.[3]

So we see that both deductive and inferential judgments are an appropriate and necessary part of the rate review process and that inferential judgments must be made both during the process and in the evaluation of the compatibility of the results with the underlying assumptions. So rate level is both cause and effect. It is the effect of a certain level of losses but it also is, in a sense, a cause of those losses since the rate level determines the type and quality of risk that will be written.

For example, suppose we are to determine the rate level for Company A in a very competitive and price conscious environment such as private passenger automobile. The current rate level situation is:

	Rate Level as % of Company A Level	% of Total Market
Company B	95%	20%
Company A	100%	4%
Companies C, D, E	105%	25%
Others	120%	45%
Involuntary	150%	6%

3 The question of whether an assumed change in the new/renewal distribution should be reflected in the rate level at all is an interesting and important one and the answer is not at all clear, but neither is it pertinent to this discussion.

Now suppose our rate level indication is for a 25% increase in rates. Such an increase would totally change the competitive picture for Company A. Instead of being in a position to attract some of the best risks, Company A would have the highest rates in the voluntary market. The indication for +25% assumes the ability to attract the better risks, for that was implied by the competitive position during the experience period which produced that indication. But clearly that assumption is incompatable with our taking that increase. Yet we need the higher rate level to support the level of losses that we expect. What should we do?

Look at the table of rate levels again. If we, as Company A, need a 25% increase in rate level, then either we are doing something very wrong or Companies B, C, D and E will also need large increases in the near future. Assuming the latter, we should increase rates by 10% or so and plan to increase again later after the other companies have changed. In this way we have preserved as nearly as we can the equilibrium between our determined rate level and the assumptions underlying it.

This dynamic relationship between rate level decision and actual rate level needs is a strong argument for making the actuary a part of the decision making process instead of merely a provider of information. Let us now explore further the role the actuary might play in company decision making. Specifically let us examine company goals as they relate to the actuary or decision maker.

GOALS AND INFERENTIAL JUDGMENTS

The first question to be addressed by the actuary with regard to company goals might be: How does achievement of the stated goal in some future period change the rate level need for that period? An equally important question but one which may not be as obviously within the strict domain of the actuary is: How can the product be priced to give the stated goal the best possible chance to be achieved with the greatest positive implications for profitability? A slightly more immediate question which should also be of interest to the actuary is: What information that the actuary is most qualified to obtain and interpret would be most useful in making final pricing decisions in light of the stated goals? Let us examine these questions through a model. Suppose a nationwide insurer of private passenger automobile is considering a growth goal for itself that translates into a 100% countrywide increase in new business writings for the coming year. That growth goal need not translate into a doubling of the new writings in each state; only the countrywide goal is important to the company. The company enjoys a renewal ratio of 90% and is currently in a no growth (in exposures) situation and is earning a 2% underwriting profit on premiums. If an effective annual investment rate (that is, considering the amount of time for which policyholders' funds are held for investment) of 3% on premiums is earned and 8% is earned on invested surplus, then the current overall rate of return on surplus given a 2 to 1 premium to surplus ratio is:

$$2(2.0\% + 3.0\%) + 8.0\% = 18.0\%$$

The overall return of 18%[4] is considered adequate but is generally not sufficient to attract significant amounts of new capital. If new business generates loss ratios which are 20% above those for renewal business, the premium dollar may have components like the following:

---

4 Let us simplify by ignoring taxes and other peripheral nuisances.

	New (10%)	Renewal (90%)	Combined
Losses	72.0%	60.0%	61.2%
Expenses[5]	36.8	36.8	36.8
Profit	- 8.8	3.2	2.0
Total	100.0%	100.0%	100.0%

A doubling of new business spread equally among the states would result in a combined loss ratio of about 62.2% as follows:

	New (20)	Renewal (90)	Total
Losses	72.0%	60.0%	62.2%
Expenses	36.8	36.8	36.8
Profit	- 8.8%	3.2	1.0
	100.0%	100.0%	100.0%

The return on surplus would then be:

$$2.2\ (1.0\% + 3.0\%) + 8.0\% = 16.8\%$$

We have given up some current earnings for growth which we hope will translate into future higher earnings. But suppose the insurers domain consists of only two states, each with one half the total premium volume and with experience as follows:

	State A		State B		
	New (5)	Renewal (45)	New (5)	Renewal (45)	Total
Losses	66.0%	55.0%	78.0%	65.0%	61.2%
Expenses	36.8	36.8	36.8	36.8	36.8
Profit	- 2.8	8.2	- 14.8	- 1.8	2.0
Total	100.0%	100.0%	100.0%	100.0%	100.0%

5 Of course new business expenses are higher but that isn't needed here; the point is that the profit is negative.

Now if our growth goal countrywide can be achieved by writing all our extra new business in State A, the profit picture would be quite different. The countrywide loss ratio would be 61.6%, the underwriting profit would only be reduced to 1.6% and the rate of return would be:

$$2.2\ (1.6\% + 3.0\%) + 8.0\% = 18.1\%$$

a higher current rate of return on surplus with achievement of our growth goal, the best of both worlds. But notice that the level of new business writings would have to triple in State A to achieve this result. Such a level of new production may go beyond the efficient limits imposed by either internal manpower constraints (such as the number of underwriters familiar with that state) or the availability of such a large number of new and acceptable risks in the state. It is probably more reasonable to assume that the loss ratio, the expense ratio or both would increase for new business in State A as we do the things, such as advertising or loosening of underwriting standards, which are necessary to attract large numbers of new risks. The result would probably reduce the rate of return to a figure below the 18% no growth rate. The model could be taken further to include such factors; but let us instead return to the three questions of interest to the actuary and attempt to answer them within the context of the model.

The first question was: How does achievement of the goal in the future period change the rate level need for that period? The question could be addressed each time a state's rates are reviewed, but we have already seen that it is better from a profitability standpoint to concentrate growth in the profitable states within certain limitations. To address the question state by state is to give an incomplete and unsatisfactory answer. How then should we proceed? We must answer the question on a countrywide basis by developing a state by state plan of growth that adds up to the countrywide objective. Each state's part in the overall goal can then become a factor in the inferential judgments used to develop the "goal oriented" rate level indications for that state. But this cannot reasonably be done without reference to the second question of interest to the actuary: How can the product be priced to give the stated goal the best possible chance for achievement with the greatest positive implications for profitability? For example, suppose the stated goal is moderate growth with no reduction to the overall rate of return. Most probably the influx of unprofitable new business would have to be offset by an increase in the general rate level so that the same overall rate of return is achieved. But if we want to concentrate our growth in the profitable states, those states would require a substantially higher growth rate prompting an equally substantial offsetting rise in the overall rate level. This rise in rate level may not be possible without compromising the competitive position in the state and destroying the possibility for the desired growth, a "Catch 22" situation. The solution may be to concentrate the growth in the profitable states but spread the needed rate level increase among all states so the effects on competitive position will be negligable.[6]

THE NEED FOR A PRICING POLICY

The details of the particular problem of growth and profitability being discussed here are not important. What is important is the generalization that springs from the exercise. The best way to achieve the overall corporate goal in this case is to plan a strategy for each state so that the sum of the states' objectives equals the countrywide objective. But all states' goals must be set at the same time to insure that the whole will equal the sum of its parts. But the nature of the task will not allow all states' rates to be reviewed at the same time. Therefore, it is imperative that the corporate objective be translated into a state by state pricing strategy which can be referenced as each state's rates are reviewed. It can be argued that if the role of the actuary, or more precisely the function of the pricing area of the company's actuarial department, is to derive the best possible estimate of the future rate level needs in each state, then the corporate goals need not be translated into a pricing policy at all; they need only be recognized by management as it makes state by state rate level decisions (as contrasted with the calculation of rate level needs). Then the decision maker would receive two separate inputs:. the overall corporate goals on the one hand and the state by state "no goals"

6 Such action should not ordinarily cause regulatory concern since the rates state by state are usually at such a point within the range of reasonableness that small increases would not produce excessive rate levels.

rate level indications on the other. The decision maker's task then would be to synthesize the two to determine the rate level that will actually be used in the future. No doubt corporate goals could be addressed in this way but the method needlessly obscures the ramifications for the overall goal of the individual state decisions. The decision maker is left in the position of having to make an inferential judgment with practically no guidance. For example, he may have a rate change need of +10% in a particular state in which he wishes to generate substantial growth. He may estimate that, with growth, the rate level need would be +12% so that growth has virtually eliminated the 2% profit in the state. This is a rather subjective foundationless inference.

Now remember the third question of interest to the actuary: What information that the actuary is most qualified to obtain and interpret would be most useful in making final pricing decisions in light of the stated goals? The answer here is that if corporate goals are translated into pricing strategy which is then communicated down to the level at which the rates are actually reviewed, then "no goal" rate level needs and "corporate objective" rate level needs can both be calculated and compared. The decision maker can then see precisely what the goal is costing in profitability and competitive position. He can also see the ramifications for the countrywide goal of choosing the "no goal" rate indication for a particular state. In this way the decision maker can constantly reevaluate the goal itself in terms of its profitability cost and also track the state by state progress toward the goal as each state is reviewed. The decision maker can do all this because the actuary, aware of corporate goals and armed with a pricing policy based on them, has provided a precise calculation of the relationship between those goals and the otherwise applicable rate needs for each state. The need for inferential judgments has not been eliminated, but a vehicle has been provided by which those judgments can be made within the rate review process itself. That vehicle is a pricing policy based on overall corporate objectives. We have replaced the 12% vs. 10% crude judgment of the decision maker above with the mathematical evaluation by the actuary based on statistical knowledge of the relationship between new and renewal loss ratios.

Let us summarize in just a few sentences. Rate level indications are not static inputs into the decision making process, rather they form a dynamic interrelated system with decisions, either rate level or growth decisions, so that indicated rate level needs determine and are determined by those decisions. This realization does two things. First it argues persuasively for including the actuary in the decision making process. Second it demonstrates the need for a direct link between corporate goals and the rate level calculation (a pricing policy) so the actuary can calculate rate levels in a manner consistent with those goals and provide other information of value in the decision making process.

PRICING FOR CORPORATE OBJECTIVES

by Frank Karlinski

Review by Robert A. Anker

My initial impression of Mr. Karlinski's paper was that it is a sales piece aimed toward fuller utilization of the actuary's training and skills in the pricing decision process. Subsequent readings confirm that impression.

The premises of the argument Mr. Karlinski espouses go like this:

1. Often the ratemaking/rate review process performed by the actuary is a pro forma data analysis function. (Mr. Karlinski later refers to this as "no goals" ratemaking.)

2. The nature of the insurance pricing/underwriting/marketing process is such that each of these components are interdependent with each other and with profitability.

3. The quality of decisions made is directly related to the quantity and quality of information available at the decision point.

Mr. Karlinski first argues that, where the first premise applies, it represents underutilization of the actuary's skills and that, at a minimum, the relevant interdependence inherent in "no goals" ratemaking should be valued and fed into the decision process. This would represent, as a first step, an enhanced utilization of the actuary's skills, some recognition of interdependence, and some improvement in the information available at the decision point.

Without an explicit identification, Mr. Karlinski then points out that the corporate goals tend to be marketing and profitability oriented and set independently of pricing/underwriting considerations and, often, independently of each other. To provide the link among all variables, he identifies the need for pricing policy. In discussing the pricing policy, he often refers to it as a pricing strategy. The term strategy is highly preferred, both because it is more descriptive and because it is the more commonly used term in the literature on corporate planning which is, to a large extent, what this paper is about.

Once the strategy link is included, complete with all its necessary feedback loops, the pricing/underwriting/marketing/profitability interdependencies have been recognized and can be valued. The valuation provides the vehicle for maximizing the quality and quantity of information to the decision maker. The actuary becomes key to the process because he is most qualified to accomplish and interpret the valuations to the decision maker(s).

Thus, to the extent that each of the premises indicated a deficiency in existing processes or an opportunity for improvement, corrective mechanisms have been identified.

There are companies who have achieved the processes for which Mr.

Karlinski argues, and some that have gone beyond them. Where the full loop exists, the actuary is intimately involved in the goal setting process itself. It is the actuary who provides the analytic guidance on all past and prospective goal interdependencies within the context of the corporate mission and long range plans and strategies.

I agree with Mr. Karlinski's premises and his conclusions. However, there are a few deficiencies or unrecognized opportunities that might be noted:

1. Coverage and underwriting standards are as equally a part of the interdependent structure as rate and market position. They also need to be recognized in the process and, by that recognition, can afford additional flexibility to the process.

2. There are also opportunities to affect the balance of the various items through changes in the expense portion of premium dollar.

3. The same arguments for strategizing from corporate goals to a state-by-state level are valid for strategizing the ratemaking within a state. Indeed, this is probably the key place for an actuary to start demonstrating the skills that can apply at a higher level.

4. In attacking the problem of sacrificing some profitability in a given state in order to gain market position, it is more

proper to use surplus funds for that expansion and require that the necessary return be gained over time from the specific state. This is both a better business basis for decision making and an approach less open to possible regulatory criticism.

Finally, speaking as both an actuary and a corporate planner, I feel that the questions Mr. Karlinski articulates as needing to be addressed by the actuary are "right on." But, even more importantly, they are questions that must be asked by management in order to realize optimum use of corporate resources.

TITLE: RATEMAKING FOR THE PERSONAL AUTOMOBILE PHYSICAL DAMAGE COVERAGES

AUTHOR: Mr. John J. Kollar

Mr. Kollar is Associate Actuary and Manager for Insurance Services Offices. John received his FCAS designation in 1975 and is a member of the American Academy of Actuaries. He is a past member of the CAS Examination Committee.

REVIEWER: Mr. Galen Barnes

Mr. Barnes is Actuary, Property/Casualty Pricing with Nationwide Insurance Company. He received his FCAS in 1977 and is a member of the American Academy of Actuaries. Galen is also a member of the Examination and Editorial Committees of the CAS.

INTRODUCTION

As with any other line of insurance, the ratemaker's goal is to develop rates that will cover losses and expenses (including underwriting profit) arising from policies in force during a specified future period. In order to accomplish this goal, a proper match between premiums (or exposures) and losses plus expenses must be established. This is particularly important when starting from an experience period which may reflect conditions which have changed or which are expected to change prior to or during the period for which rates are being made.

The ratemaker must know what coverage was and will be provided. Has the insurance policy itself changed? Has the legislature, insurance regulator or court changed the interpretation of the policy resulting in a de facto change in losses? Has the term of the policy been modified? Has there been a shift in insureds by deductible?

The ratemaker must know who was and will be insured. Has there been a change in the company's marketing or underwriting policy? Has there been a change in the involuntary market mechanism? Has there been a change in cancellation or nonrenewal laws? Has there been a shift in insureds by class or territory?

The ratemaker must estimate what effects changes in economic and other conditions will have on insurance costs. What is the change in the cost of goods and services for which insurance pays? What is the change in claim frequencies?

The ratemaker must know what the rating system itself was and will be. Has a rating variable, e.g. damageability-repairability, been modified, introduced or eliminated? Has the overall rate level been revised?

All of the above factors and their interaction should be considered in making rates. The following sections of this paper will concentrate on them, particularly those most important or unique to ratemaking for the personal automobile physical damage coverages.

DEDUCTIBLES

A significant shift in the distribution of insureds by deductible during the experience period may lead to an improper matching of premiums and losses. For example, if there were a significant shift of insureds from the $100 deductible to the $200 deductible, the distribution of premiums and losses might be comparable to Table 1 for a given calendar year. If precise payments and reserves were known immediately with no effects from prior calendar years, then the actual incurred losses and the actual incurred loss ratio would be those in columns (2) and (3), respectively.

TABLE 1 (premiums and losses in thousands)

Deductible	(1) Earned Premiums	(2) Actual Incurred Losses	(3) Actual Loss Ratio(2)/(1)	(4) Calendar Year (C.Y.) Paid Losses
$100	$1,200	$ 840	70%	$1,250
200	800	560	70	200
Total	$2,000	$1,400	70	$1,450

Deductible	(5) C.Y. Loss Ratio(4)/(1)	(6) Accident Year (A.Y.) Paid Losses as of (12 mos.)	(7) A.Y. Loss Ratio(6)/(1)
$100	104%	800	67%
200	25	400	50
Total	73	$1,200	60

It is clear that the calendar year paid loss ratios in column (5), which differ from the actual incurred loss ratios in column (3), would produce improper matchings of losses and premiums due to the lag in reporting, processing and paying of claims.

The accident year paid losses as of 12 months in column (6) have to be developed to an ultimate level. Unless the overall loss development factor (or factors by deductible) reflects the shift in deductibles (more development on the growing $200 deductible than the declining $100 deductible), a difficult task, the losses will not precisely match the premiums.

As almost all physical damage claims are paid within 60 to 90 days of occurrence, few (if any) companies establish individual case basis reserves.

Consequently, as of the end of the experience year the actual incurred losses will not be precisely known. As of 15 months (three months after the end of the calendar year) the accident year paid losses should have developed to the actual incurred losses, and the distortion would have been eliminated.

Another way of correcting for the distortions in the data as of 12 months, particularly if calendar year paid losses are used, is to adjust the data to a common deductible basis as set forth in Table 2.

TABLE 2 (premiums and losses in thousands)

Deductible	(1) Earned Premiums at Current Levels	(2) Relativity to \$200 Deductible	(3) Premiums on \$200 Level (1) X 1.00/(2)	(4) Calendar Year Paid Losses
\$100	\$1,200	1.25	\$ 960	\$1,250
200	800	1.00	800	200
Total	\$2,000		\$1,760	\$1,450

Deductible	(5) L.E.R. to \$200 Level	(6) 1.0 - L.E.R. to \$200 Level 1.0 - (5)	(7) Losses on \$200 Level (4) X (6)	(8) Adj. Loss Ratio (7)/(3)
\$100	20%	.80	\$1,000	104%
200	0	1.00	200	25
Total			\$1,200	68

Deductible	(9) Number of Paid Claims	(10) C.E.R. to \$200 Level	(11) 1.0 - C.E.R. to \$200 Level 1.0 - (10)	(12) Claims on \$200 Level (9) X (11)
\$100	2,500	10%	.90	2,250
200	500	0	1.00	500
Total	3,000			2,750

Generally this adjustment is to the higher deductible because the long term shift is to higher deductibles. Premiums, losses and claims should be known by deductible in order to do this properly. (Depending on the trend procedure, claims may not be needed. See the section on trend.) The $100 deductible premiums are adjusted to the $200 deductible level by multiplying by the ratio of the $200 to $100 rate relativity as shown in columns (1) through (3).

The $100 deductible losses and claims are adjusted to the $200 level by loss and claim elimination ratios (L.E.R.'s and C.E.R.'s), respectively, as illustrated in columns (4) through (12). L.E.R.'s and C.E.R.'s are developed from a distribution of losses and claims by size and reflect the dollars of losses and numbers of claims eliminated by switching from a lower to a higher deductible. These ratios should reflect loss levels and distributions comparable to the experience period. The total loss ratio of 68% in column (8) of Table 2 does not equal the actual incurred ratio of 70% in column (3) of Table 1 because the former is on a paid basis and the latter is on an incurred basis. Because of the inherent lag in paid data the paid losses in column (7) of Table 2 have an average date of accident two or three months earlier than the actual incurred losses and consequently reflect earlier loss levels. The paid losses can be adjusted to an incurred basis by multiplying by a ratio of incurred losses to paid losses (generally about 102% for a sample of companies).

In Tables 1 and 2 it was assumed that earned premiums and actual incurred losses were in equal proportions by deductible, i.e., deductible rate relativities were appropriate during the experience period. If this were not the case, then distortions might be shrouded as illustrated in Table 3. Though the losses and total premiums are the same as in Table 2, the premiums by deductible in column (1) have been changed to reflect the inappropriate relativities assumed in column (2).

TABLE 3 (premiums and losses in thousands)

Deductible	(1) Earned Premiums at Current Levels	(2) Relativity to $200 Deductible	(3) Premiums on $200 Level (1) X 1.00/(2)	(4) Actual Incurred Losses
$100	$1,400	1.40	$1,000	$ 840
200	600	1.00	600	560
Total	$2,000		$1,600	$1,400

Deductible	(5) Actual Loss Ratio (4)/(1)	(6) Calendar Year Paid Losses	(7) Paid Loss Ratio (6)/(1)
$100	60%	$1,250	89%
200	93	200	33
Total	70	$1,450	73

Deductible	(8) L.E.R. to $200 Level	(9) 1.0 - L.E.R. to $200 Level	(10) Losses on $200 Level (6) X (9)	(11) Adj. Loss Ratio (10)/(3)
$100	20%	.80	$1,000	100%
200	0	1.00	200	33
Total			$1,200	75

While the overall calendar year paid loss ratio and the overall actual incurred loss ratio equal those in Table 1, the overall adjusted loss ratio of 75% in column (11) is much greater than the overall adjusted loss ratio of 68% in column (8) of Table 2. The actual rate level need is much greater than a superficial review of the overall data would indicate. Consequently it is essential to rate each deductible appropriately. This can be accomplished by developing rates independently for each deductible having a credible volume of data. However, problems may result from this type of an approach. A smaller deductible might indicate a lower rate than a larger deductible. A smaller deductible might cost unreasonably more than a larger deductible. These problems can be rectified by requiring reasonable relationships between rates of different deductibles.

Another way of pricing the deductibles appropriately is to adjust the data to a common deductible basis as discussed previously. This will result in a proper pricing of this key deductible. Rates for the other deductibles will then be established by relativity to the key deductible.

Regardless of whether or not the deductibles were rated appropriately in the past, it is necessary that they be reevaluated for the future. This can be accomplished by recalculating L.E.R.'s after the distribution of losses and claims by size has been adjusted to future loss levels. Each L.E.R. would equal the rate discount to shift from the lower to the higher deductible (assuming that expenses are treated separately).

It should be noted that the deductible rates developed by the above procedures are appropriate for the average insured. If certain deductibles are generally purchased by atypical insureds, then the results for those deductibles will be atypical. Rate relativities should be adjusted to reflect such atypical distributions. An example of such a situation would be the very high deductibles, e.g., $1,000. Such a deductible is generally purchased by an insured who has a car valued much greater than that of an average insured. The expected percentage loss savings for such an insured with a $1,000 deductible would undoubtedly be significantly less than the size of loss distribution for the average insured would indicate due to the greater value of the car. The rate relativity should be adjusted to reflect this.

In the foregoing paragraphs the proper matching of premiums and losses was discussed as it pertained to the experience period and the future rate relativities between deductibles. The shift in distribution by deductible may also impact trend data. If comprehensive or collision data were examined for all deductibles combined, e.g., Fast Track data, the trend for the underlying experience period would be understated due to a shift to higher deductibles. In Table 4 an extremely simple example illustrates how a shift to higher deductibles would result in an apparent downward "trend" in losses when in fact there were no trends in the loss components.

TABLE 4

Year	(1) Exposures	(2) Claims	(3) Losses	(4) Claim Cost	(5) Claim Frequency	(6) Loss Cost
x	10,000	1,000	$1,000,000	$1,000	10%	$100
x + 1	10,000	900	800,000	889	9%	80
Change from x to x + 1	0%	-10%	-20%	-11%	-10%	-20%

Assumptions:

1. All coverage in year x and x + 1 was $100 and $200 deductible, respectively.
2. The L.E.R. to go from $100 to $200 deductible was 80%.
3. The C.E.R. to go from $100 to $200 deductible was 90%.
4. There were no other differences in conditions from year x to year x + 1, i.e., no change in claim cost or claim frequency by deductible.
5. (4)=(3)/(2); (5)=(2)/(1); (6)=(3)/(1)=(4)X(5).

The problem of shifting distributions by deductible can be reduced by examining trend data separately by deductible. Nevertheless the innate lag in the payment of claims might create an improper match between paid loss and exposure data used in determining claim frequency and loss cost (pure premium) trends.

Substantial shifts in the distribution of insureds by collision deductible might even impact property damage liability (P.D.L.) loss data. For example, if insured A had $100 deductible collision coverage and incurred $180 worth of collision damage, then insured A could collect $80 from his own company. If insured B caused the accident, then insured A's company could collect $180 from insured

B's company and should return the $100 deductible to insured A. The total P.D.L. loss would be $180 for that claim. To simplify the example it is assumed that exactly the same situation occurred one year later except that insured A had a $200 deductible. With a $200 deductible insured A could not collect from his own insurance company. Insured A would have to seek recovery of his loss on his own. Collecting on a liability claim requires substantially more effort than collecting under first party coverage. If insured A did not seek recovery, there would be a reduction of one claim and $180 in losses for P.D.L. This would result in a reduction in P.D.L. claim frequency and loss cost. As the average claim cost for P.D.L. has been greater than $180, the average claim cost would actually increase due to the elimination of a small claim. It should be noted that such a situation would only be expected to affect a small number of P.D.L.claims, and therefore, the effect on P.D.L. trend data would be much less than on collision trend data.

THE INSURED

Physical damage insurance is not compulsory, but it may be required for the life of the loan if the car is financed. As the car ages many insureds drop collision and possibly comprehensive coverage. Thus there is a gradually shifting mix of insureds in the data over time. While paid loss data is not as current as premium data in calendar year ratemaking, the match of premiums and losses is only minimally affected. To make rates for a somewhat different group of insureds in the future, it is necessary to develop proper class and territory rate relativities between risks so that a change in distribution will not result in changes in overall loss ratios.

The shifting mix over time of insureds purchasing coverage might impact physical damage trend data. On the other hand the mix of insureds would be relatively consistent in P.D.L. trend data because the need for coverage is not a function of the insured's car. However, there are overall demographic shifts which might have some impact on trends for both P.D.L. and physical damage coverages. There has been a gradual population movement from urban to suburban and rural areas. This shift has largely been to areas with lower loss levels and lower rate levels. With the shift to more rural areas has come an increase in multi-car families. Multi-car insureds receive a multi-car discount on each car because of lower losses per car than single car insureds. Both of these shifts may not continue in the future. There has also been a gradual increase in the average age in the general American population. This has resulted in a decreasing percentage of youthfull operators and an increasing percentage of adult and "over 65" operators. This gradual shift has been to insureds with lower loss and rate levels. All of these shifts have resulted in a small declining effect on average rates and on loss trends for all coverages.

ECONOMIC & OTHER CONDITIONS

Recent inflationary trends of over 10% a year in automobile damage repair costs have exceeded the cost increases in the overall economy. These large repair cost increases were primarily due to the rise in crash parts prices resulting from the increasing cost to produce such parts and probably from the monopolistic nature of crash parts production. Consideration of inflation as the most important economic factor impacting insurance costs must recognize that inflationary rates have been volatile, and therefore, past trends may not be the best indicator of future conditions.

Inflation also causes increases in the cost of both new and used cars. As insurance rates for the physical damage coverages reflect the price of the car when new (symbol), they increase with the price of the car. In recent years auto manufacturers have increased the prices of their new cars during the model year as well as at the beginning of the model year. As symbols have been assigned to a car at the inception of the model year, these subsequent auto price increases have not resulted in any additional premium revenue which was needed to offset the increased loss potential. While used cars generally decrease in value as they age, inflation generally helps to reduce the magnitude of this decrease and thereby lessen the decrease in losses on these cars. Thus losses remain high through the life of the car and losses on new cars are larger then losses on old cars.

A recession or a severe slowdown in economic growth generally includes a decrease in new car sales. This results in a short term reduction in premium levels from what would otherwise have been expected. In the past new car sales have rebounded so that in the year(s) following the recession a large number of new cars would be sold. In the long run the distribution of cars by age has remained relatively constant. To make rates for a different mix of insured cars by age (and symbol too) in the future, it is necessary to develop age (and symbol) rate relativities between cars as precisely as possible.

In the past reduced economic growth had meant a reduction in the inflation rate. Recently stagflation, reduced economic growth with a continuing high rate of inflation, has resulted in continued large cost increases with no premium increases, or even decreases, due to the decline in the sale of new cars.

In recent years the exchange rate for American dollars has vacillated and has generally decreased for most countries exporting cars to the United States. While resulting increases in the cost of foreign cars have led to higher rates and losses for them, increases in the cost of foreign car parts have only led to higher losses. These higher losses have been in addition to the increased losses due to inflation. Repairs of foreign cars have also been more costly because of the limited availability of replacement parts and repair services. Also foreign cars have been increasing their share of the market steadily through the 1970's.

One of the reasons for the increased popularity of foreign cars is their greater fuel efficiency. Only in the late 1970's did domestic auto manufacturers seriously begin to develop fuel efficient cars in response to increased consumer demand and federal regulation. Fuel efficiency has become an important consideration because of the uncertain availability and cost of gasoline.

The gasoline shortage in 1973-4 resulted in fewer miles driven and reduced claim frequencies. As the reduction in mileage driven exceeded the decrease in claim frequencies, it is likely that the mileage eliminated was of a lower frequency nature. The recent and continued shift to more fuel efficient cars should reduce the need for gasoline from what it otherwise would have been. Thus a decline in gasoline usage would not necessarily translate into an equal decline in claim frequencies. This would be particularly true for comprehensive claim frequency which includes many perils largely unaffected by gasoline usage.

For a ratemaking experience period reflecting reduced claim frequencies due to a temporary gasoline shortage, e.g. 1973-4, it is necessary to adjust claim frequencies to expected levels as if no gasoline shortage occurred. Such an adjustment should probably vary by region as the claim frequency reductions probably varied by region. In addition, this atypical experience might cause distortions in both collision and P.D.L. trend data unless adjustments are made to remove the effects of the temporary gasoline shortage.

Offsetting to at least some degree any claim frequency reductions due to reduced gasoline usage has been the increase in claim costs due to higher energy costs. In addition, more fuel efficient (smaller) cars have also been shown to have worse loss experience than larger cars.

If claim frequencies might be reduced in the future because of the uncertainty of reduced gasoline usage, there would undoubtedly be public pressure to reduce rates or expected rate level needs. If claim frequencies did not decline as anticipated or they rebounded to past levels, then rate increases would be needed, and needed immediately, sooner than they could be implemented.

While the prior paragraphs dealt with uncertain reductions in gasoline supplies, it is possible that the federal government could ration or otherwise substantially limit gasoline supplies for an extended period of time. Under such conditions consideration should be given to reducing expected claim frequencies in the ratemaking formula. At the same time premium income would probably be reduced because of a reduction in the number of cars driven and a shift to lower rated classes. The reduction in gasoline supplies is one example of an economic factor which could impact both premium and loss data.

Comprehensive data can be distorted by a catastrophe or a series of catastrophes. To make adequate and stable rates, losses from such an occurrence(s) should be excluded from the experience period. A provision based on a long-term average of such losses should be included in the rates even if no catastrophe occurred.

TREND

While the prior section discussed changes in economic and other conditions in general terms, this section will concentrate on the more specific reflection of these changes through trend data.

Experience period losses can be trended to future levels by use of physical damage data, property damage liability data, econometric indices or some combination of these items. Of course, physical damage data most closely reflects all past changes in physical damage losses. It is also distorted by the change in the distribution of insureds by deductible. As discussed in the section on deductibles this distortion can be reduced, but not eliminated, by examining trend data by deductible. The alternate approach of adjusting trend data to a common deductible basis requires modification of at least some points of trend data which might distort the resulting trend factors. If individual size of loss distributions are not available for each point to be adjusted, then the adjustment factors themselves have to be estimated further increasing the likelihood of inaccurate results.

Because comprehensive provides coverage for losses due to catastrophes, storms and other irregular occurrences, the use of comprehensive data for trend may require additional judgmental adjustments. Changes in the distribution of losses by peril can be due to unique conditions or continuing long term trends. While the latter should be reflected in the trend data, the former should not be reflected. Thus trend factors based on comprehensive data may be even less accurate than trend factors based on collision data.

Like collision data property damage liability (P.D.L.) data reflects damage to primarily automotive parts. Damage to non-automotive parts and the third party nature of P.D.L. losses may cause some small distortion in using P.D.L. trend data to measure trends in collision losses. As P.D.L. claims require longer time to settle, P.D.L. trend data is not as current as physical damage trend data. As noted in the section on deductibles, P.D.L. data is only minimally impacted by the change in distribution of insureds by deductible.

P.D.L. coverage pays claims from the first dollar of loss whereas physical damage coverages are generally subject to a deductible. The trend in physical damage deductible losses, therefore, exceeds the trend in P.D.L. losses. The trend in physical damage deductible losses is analogous to the trend in excess losses explained by Mr. J. T. Lange in his paper "The Interpretation of Liability Increased Limits Statistics", PCAS LVI, 1969. For example, if the inflation rate applicable to automobile damage is 10%, then a $400 P.D.L. claim will be $440 in the following year. If the exact same occurrence were paid under a physical damage coverage with $200 deductible, the loss would be $200 one year and $240 in the following year, or 20% higher. The use of a P.D.L. claim cost factor of only 10% to trend physical damage losses would obviously understate the increase in physical damage losses. To adjust for this understatement in expected physical damage losses, the deductible can be added back on each claim, then the P.D.L. trend factor can be applied to the total damage amount, and finally the deductible can be removed from each claim as shown in Table 5.

TABLE 5 (Losses in Thousands)

(1)	(2)	(3)	(4)
$200 Deductible		P.D.L.	Trended $200 Deductible Losses
Losses	Claims	Trend	[(1)+(2)X$200]X[1.0+(3)]-[(2)X$200]
$1,600	2,000	10%	$1,800

The calculated trend in $200 physical damage losses is +12.5% ($1,800/$1,600-1.0) in this example.

As the mix of parts damaged in comprehensive losses differs from that in P.D.L. and collision losses, there may be a difference in trends for these losses. As noted earlier in this section, apparent trends in comprehensive losses may even be caused by shifts in the distribution of losses by peril. Any long term trends in comprehensive losses due to such shifts would not be reflected in P.D.L. trend data, although they should be.

Inflation leads to an increase in physical damage claim frequencies by causing previously uninsured (small) accidents to exceed the deductible level and thereby become collectible. Thus, the application of a P.D.L. claim frequency trend factor to physical damage losses would understate future levels of physical damage claim frequencies. As P.D.L. claim frequency has been decreasing about 2% or 3% per year, no change in collision claim frequency has generally been assumed. Because no consistent pattern in comprehensive claim frequency has been identified and comprehensive covers perils different than P.D.L. and collision, no change in comprehensive claim frequency has generally been assumed.

By definition insurance data reflects historical facts, and therefore, may not be responsive to current and future conditions. Economic indices may provide leading indicators of changes in physical damage losses. Furthermore, econometric models may eventually succeed in predicting future physical damage losses, or at least future claim costs. Of course, such models must provide an acceptable fit to actual physical damage loss trends which are affected by many different economic and other conditions.

As discussed in the section on economic conditions, the increase in the cost of new cars has resulted in a gradual increase in symbol (based on cost new), and therefore, rates for new cars. This increase has been partially offset by the shift from larger to smaller less expensive cars. The resulting increase in symbol can be measured by the change in the average symbol insured separately for comprehensive and collision coverages. As a car ages it generally has decreased in value, and therefore, its rates have gradually decreased in relationship to new car (age 1) rates. To reflect changes in the volume of new cars purchased as well as their increase in price, the average age and symbol factor can be examined separately for comprehensive and collision coverages. Average symbol and average age and symbol relativities must be examined separately by coverage and deductible because of different rate relativities by coverage and distributional differences by coverage and deductible.

As discussed in the section on the insured, there has been a gradual population movement from urban to suburban and rural areas with an increase in multi-car insureds. There has also been a gradual decrease in the percentage of youthful operators. Both of these factors, which result in a small gradual decrease in average rates, can be measured by examining the change in average rates on the current rate level after excluding any age and symbol changes.

Reviewing average class plan or average age and symbol relativities over some experience period may be difficult. When such relativities have been revised during the experience period, the change in the average relativity may be distorted. To adjust for this distortion, one set of relativities (preferably the current one) can be substituted for any other set of relativities in effect during the experience period. Because of the extensive degree of detailed data needed, it may not be possible to make this adjustment by class or age and symbol group. In such a situation approximate adjustments could be made.

To this point discussion of premium and loss trends has concentrated on historical data. This may not be the best indicator of future, or even current conditions. All factors impacting premium or loss trends must be considered. To ensure the responsiveness of trend factors to current and future conditions judgmental modifications should be made as necessary. Trend factors could be selected to be higher or lower than past trends. If trend factors vary by region or state, minimum and maximum trend factors could be used to ensure the reasonableness of individual trend factors.

RATING SYSTEM

In addition to the typical rating variables of class and territory, physical damage rates have traditionally varied by age and symbol of car. As illustrated in Table 6, each symbol has represented a dollar range to which each make and model of car in the new model year has been assigned based on its cost when new. As more expensive cars have generally cost more to repair or replace than less expensive cars, rates for more expensive cars (higher symbols) have been greater than rates for less expensive cars (lower symbols).

TABLE 6

Price New	Symbol	Comprehensive Rate	Collision Rate
$0 - 1,000	1	$ 90	$ 180
1,001 - 2,000	2	90	180
2,001 - 3,000	3	90	180
3,001 - 4,000	4	100	200
4,001 - 5,000	5	110	220
5,001 - 6,500	6	155	250
.	.	.	.
.	.	.	.
.	.	.	.

In the 1970's many insurance companies began to collect data by make and model of car. This was accomplished by recording data by auto manufacturer vehicle identification number (V.I.N.). Analysis of this data indicated that cars of similar value could have substantially different experience. This has led to the modification of symbols by make and model of car, or vehicle series rating (V.S.R.) as it is frequently called.

After a symbol has been assigned to each make and model of car (vehicle series) for a new model year based on price new at the beginning of the model year, an experience modification is made to it. The experience modification is based on the latest available combined comprehensive and collision data by V.I.N. for the predecessor of that vehicle series. Loss ratios are examined so that differences in distribution of insureds by class, territory and deductible are reflected. For example, a vehicle series that is driven by a disproportionate number of higher rated operators would be expected to have higher than average losses as its rate would also be higher than average. The loss ratio for each vehicle series is then compared to the loss ratio for all vehicle series combined for that model year. The resulting relativity indicates how much better or worse that vehicle series has been than the average. As both comprehensive and collision rates will increase or decrease significantly due to a change in symbol, a threshold must be exceeded before a symbol is increased or decreased. The threshold can be established in several ways. One amount, e.g. 20% better or worse than the average, can be used. A decision rule which requires an indicated change greater than the resulting change in combined comprehensive and collision rates is more precise but also more complex (see examples in Table 7.)

TABLE 7

Symbol Based On Price New	Indicated Change Required to Upsymbol	Indicated Change Required to Downsymbol
4	+10%	-10%
5	+25%	-10%
6	+25%	-20%

In both of these approaches a maximum change of one symbol at a time has been permitted, but greater changes might be indicated and could be implemented.

Reviewing data for every vehicle series poses obvious credibility problems. The Highway Loss Data Institute (H.L.D.I.) has been collecting the loss and exposure data by V.I.N. of many large auto insurers and publishing results by make and model. H.L.D.I. collects data separately by deductible and separately for youthful and non-youthful operators so that results by make and model can be normalized, i.e., adjusted to a common distribution by deductible and age category of operator. This removes most distortion due to a disproportionate distribution of youthful insureds by deductible or make and model.

Even with H.L.D.I. data many vehicle series have only a small volume of data. To produce meaningful results for every vehicle series a credibility procedure is required. This can be accomplished by credibility weighting the indication of the vehicle series with the indication of a similar group of vehicle series.

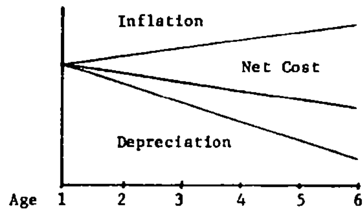
When an entirely new vehicle series is introduced, there is no data on which to base an experience modification. A similar situation occurs when a vehicle series is changed so that experience of past model years is substantially different from the expected experience of the new model year. In both cases, it can be assumed that the new model year of the vehicle series will have experience comparable to a similar group of vehicle series.

In addition to evaluating symbols for the new model year, symbols (including previous modifications) can be annually reevaluated on earlier model years (resymboled). By the second resymboling the review of symbols for a model year can be based on the data for that model year. Modifying a symbol on a new model year vehicle series has no impact on a company's policy issuing system because the new car is being covered on the policy for the first time. On the other hand, modifying a symbol on older model years requires the ability (by hand or automatedly) to rerate policies by V.I.N. The size and complexity of accomplishing this task automatedly cause a substantial impact on a company's policy issuing system as V.I.N. data must be accurate and accessible.

Rating cars by make and model as described above has little impact on overall premium levels. However, it does provide for more accurate rates by vehicle series. In the long term it may help to control increases in losses by encouraging auto manufacturers to build less damageable and more repairable cars. While the above approach modifies cost new symbols based on insurance experience, it is possible that future developments will comprise of more sophisticated rating by make and model including the modification of experience indications based on engineering analysis to reflect substantial changes in vehicle design which are expected to impact insurance losses.

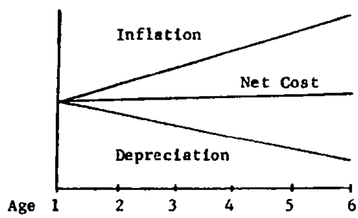
Physical damage rates have also traditionally varied by age of car. Older cars cost less to insure than newer cars; cars generally decrease in value as they age. Through the early 1970's the long term inflation rate was less than 5% per year. The increased cost of partial losses due to inflation combined with the decreasing value of total losses due to depreciation resulted in decreasing losses as a car aged as illustrated in Graph I. Thus it was appropriate to charge less for a car as it aged. This was accomplished by applying increasing discounts to the new car rates (age 1).

GRAPH I



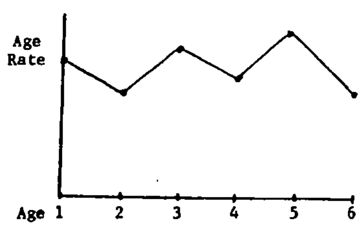
Since 1973 the high inflation rate has rapidly increased the cost of partial losses. It has also caused cars to depreciate more slowly. Consequently the net effect of inflation and depreciation has been no change or an increase in losses as a car aged as illustrated in Graph II.

GRAPH II



In reality insureds did not receive lower rates because rate revisions for substantial increases had to be implemented to overcome the inappropriate age discounts. The combination of age discounts and rate increases frequently caused rates to fluctuate as a car aged as shown on Graph III.

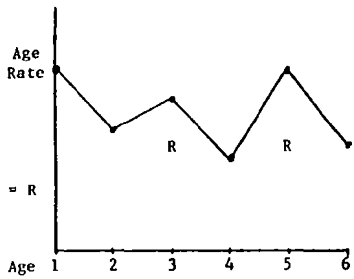
GRAPH III



While Graph III assumed modest annual rate revisions and age discounts in alternate years (ages 2,4,6), Graph IV illustrates a less stable situation with infrequent rate revisions for large increases (ages 3, 5).

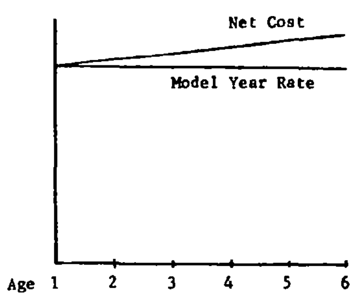
GRAPH IV

Rate Increase = R



Model year rating (M.Y.R.) is reducing these problems by assigning a rate to a car based on its model year. The model year rate generally stays the same until an overall rate revision is implemented. Thus M.Y.R. more closely matches rates and costs than age rating as illustrated on Graph V.

GRAPH V



As new cars continue to cost more to insure than old cars, rates for newer model years continued to be higher than rates for older model years. As each new model year is introduced, it is rated a percentage, e.g. 5%, higher than the preceding model year to reflect the higher insurance cost of new cars.

The absence of age discounts and the shift in distribution to newer, higher rated model years results in more stable rates and a significant growth in premium income under M.Y.R. as illustrated in Table 8. This growth translates into a comparable reduction in overall rate level need from what would otherwise have been indicated. The combination of premium growth due to M.Y.R. and reduced rate level need results in approximately the same indicated rates as under age rating.

TABLE 8

	Comprehensive			Collision		
	(1)	(2)	(3)	(4)	(5)	(6)
Model Year	Rate Rel.	Dist. in Year X	Dist. in Year X+1	Rate Rel.	Dist. in Year X	Dist. in Year X+1
x+1	1.05	-	9%	1.05	-	10%
x	1.00	9%	11	1.00	10%	12
x-1	.94	11	11	.92	12	12
x-2	.88	11	11	.85	12	12
x-3	.83	11	11	.78	12	12
x-4	.78	11	11	.72	12	12
x-5	.73	11	11	.66	12	12
x-6	.69	11	11	.60	12	10
x-7	.65	25	14	.55	18	8

	Comprehensive	Collision
(7) Average rate relativity in year x:	.786	.743
(8) Average rate relativity in year x+1:	.829	.801
(9) Change in rate relativity:	+5.5	+ 7.8%
(10) Approximate effect on rate level indication: 1.0/[1.0 + (9)] - 1.0	-5.2%	- 7.2%

While the same distribution of cars by age has been a reasonable assumption in the long term, the distribution has varied in the short term. However, a substantial change in distribution would be needed to significantly impact the effect on rate level line (10) of Table 8. The future distributions used to weight the relativities must be estimated. Consequently, exposures are generally used although premiums would be more precise.

Model year rate relativities must be reviewed frequently to ensure their continued appropriateness. Rates by model year allow responsiveness to better or worse experience between model years, which was not feasible under the age rating system.

STATISTICAL IMPLICATIONS

After rating and ratemaking requirements have been identified, all statistical implications must be determined. This is not to say that statistical implications are secondary. On the contrary, before any rating or ratemaking change is implemented, its statistical implications should be considered.

All data must have sufficiently high quality to have positive value in rating or ratemaking. Inaccurate data can lead to inadequate or excessive rates thereby damaging a company's fiscal or market position, not to mention its credibility when errors are discovered.

To develop quality data the coding of information should be logical and simple. Instructions should be provided with examples as necessary. Training should be a prerequisite for all new coders and for all coders when there is a major statistical change.

As information needed to rate risks is usually recorded more carefully than other data, code as rated is generally a good rule to follow. However, if the rating scheme is too complicated, a significant number of both rating and coding errors will result. In certain situations it may be important to collect data in greater detail than is used for rating. Review of such data should consider its possible lower quality because it is not used in rating.

When the statistical implications of a change in rating or ratemaking have been determined, then the cost of all facets of the change can be weighed against all benefits of the change. This may lead to a simplification of the rating system. While it may be necessary to make a change in the rating system, the decision may be made to forego the coding of less significant items of information. Detail desired for ratemaking may have to be modified. For example, collecting ratemaking data in complete detail by deductible, age and symbol, V.I.N., etc. may not be feasible for a company.

Whatever data is collected can be edited to enhance its quality. Field edits check the validity of certain columns of data: is a particular class code valid? Relationship edits check the validity of certain columns of data relative to other information about the insured: is a particular class code valid in a given state? Distributional edits compare summarized data of many insureds to determine if the distribution by a particular category is atypical: is there a disproportionate number of insureds in a particular class?

A similar problem to collecting data is that of summarizing the data in the detail that is required to produce needed reports. The cost of designing and implementing a particular report may exceed the benefits of the report.

Once the decision has been made on what data to collect and what reports to produce, methods of estimating required, unavailable data may have to be developed. In addition, other affected areas within the company must be informed of the limited availability of data in certain detail as this may impact their operations.

The final review of the data rests with the ratemaker. Knowledge of economic conditions and other factors affecting data during the experience period allow the ratemaker to determine the overall reasonableness of the data. For example, the presence of a large catastrophe may make comprehensive losses look overstated relative to years without a catastrophe. The ratemaker would recognize the effect of the catastrophe on the data and would be able to make the appropriate adjustments to the data as discussed in the section on economic and related conditions.

CONCLUSION

After rates are developed by the ratemaker, they are subject to management review. As the rates were developed on the basis of certain key assumptions about coverage, insureds, economic conditions and the rating system, these assumptions should be considered in the management review. If the rates are changed by management (or later by a regulatory body), the original assumptions should be modified to reflect the changes in rates. Whatever the final rates the modified assumptions underlying them should be transmitted to all involved in the selling of insurance.

As actual economic conditions can vary substantially from the expectations underlying the revised rates, the ratemaker should be continuing to monitor available sources, e.g. Fast Track data, so that management can be informed of changes in the appropriateness of the rates. This would allow the rate users in the company to adjust their assumptions about the appropriateness of the existing rates even before revised rates can be developed and implemented. The importance of good communications within a company is another aspect of properly matching premiums and losses to ensure appropriate rates.

As mentioned at the beginning of this section, the final formula rates should be evaluated for overall reasonableness and responsiveness to current and expected conditions. The formula is a tool for establishing such rates and not an end in itself. Therefore, it should be modified as necessary to reflect changing conditions...for today's formulas are tomorrow's antiquities.

RATEMAKING FOR THE PERSONAL AUTOMOBILE
PHYSICAL DAMAGE COVERAGES
by JOHN J. KOLLAR
Reviewed by GALEN BARNES

Introduction

When I was asked to participate in the call paper program by reviewing a paper, I approached the idea with some apprehension. I have never been involved in this side of the Society's endeavors. Rather, I have been a benefactor of the quality submissions of my colleagues and their reviewers in the past. That is not to say that I have not had the opportunity to study in fine detail some of the worthier papers as a student and examiner, but I have until now kept counsel only with myself.

My compliments go to John Kollar for the careful deliberation given and the time spent in the active role of providing us with a paper for discussion. In all honesty, however, I expected a paper much different in scope. When I was asked to review this paper I expected to receive a recipe guide for a beginning student in my office to use and read before he or she began asking the imponderable questions that I will never be able to answer.

Given that I am not reviewing the paper I expected to review, does that detract from my opinion of its value? Absolutely not. The paper will, I believe, foster the type of discussion for which this call paper program is intended. In addition, I believe it is particularly significant that John represents a rating organization and the reviewer an independent direct writer for automobile insurance. With different perspectives, surely the author and reviewer will identify some diverging opinions that will increase the level of discussion.

Deductibles

The paper is divided into various sections which I will abide by for ease of presentation. The first section addresses deductibles.

The theme of this section is that a significant shift in the exposure distribution by deductible during the experience period under review may lead to an improper matching of premiums and losses. An improper marriage will result in an inappropriate base upon which projections to the future are applied, (e.g. trend) for statewide rate indications, etc.

The distortion is to be eliminated by use of accident year results as of 15 months or the use of 12 months of presumably calendar year results adjusted to a common deductible. But we are to be wary if the premium conversion relativities and the loss elimination relativities are out of synchronization.

In this case it is stated that the actual rate level need may be much greater than a superficial review of the overall data would indicate. I believe that this statement needs to be made more precise as I feel that what has been shown is that the rate level for the higher deductible is relatively more inadequate than the lower deductible. When both deductibles are considered together, the overall rate level for the two combined may indeed be equivalent to the example when loss elimination ratios and premium relativities are more appropriately aligned.

For several years the reviewer has used a calendar year incurred loss approach for all deductibles combined and until recently this approach has worked reasonably well. With the advent of deductible roll-ups other approaches such as those suggested by the author have become more important. They have not, however, replaced the use of calendar year ratemaking based on incurred loss estimates.

Calendar year incurred losses are equivalent to accident year losses if required and carried reserves are the same. The usually small size of physical damage loss reserves relative to paid losses will generally ensure only small distortion of results if the required and carried reserves are not exactly the same. But periodic checking of loss reserve developments to assure reasonable accuracy is advised.

The use of total collision results will avoid some of the problems of exactly determining the price for each deductible if financial health is one of the main objectives and a balance of experience indications and reasonability is needed to price some of the infrequently used deductibles. The reviewer agrees with the author that it is essential to rate each deductible appropriately particularly if there is a shift occurring.

In addition to the methods for accomplishing appropriate deductible rating suggested by the author, the reviewer has used the technique of reviewing the trend in loss elimination ratios for judgmental selection purposes. For example, an LER for calendar year n of .12 and an LER for calendar year n +1 of .09 will produce an intuitive feeling if a judgmental selection is appropriate.

The deductible section of the author's paper also speaks briefly to the impact of deductible shifts on trend. The reviewer will speak to this issue in the section on trend. One facetious parenthetical comment is in order, however, in regard to the comments about the impact of deductible shifts on property damage liability. An actuary working for only one independent company always assumes that third party claimants will pursue claims, if only for $5.

The Insured

The theme of this section is that demographic shifts of insureds and changing vehicle population characteristics may impact trend and possibly the matching of premiums and losses.

The author has pinpointed some of the societal effects that may have an impact on both trends and the experience base. There are other societal, demographic, and vehicle changes that also come to mind which are likely impacting these actuarial measures.

While the list is not exhaustive, it includes the following:

a) Greater metropolitan density on average. While there are more suburbanites, they are clogging the urban streets and to a greater extent than in the past, the suburban streets.

b) Greater claims consciousness fueled by the perceived decline of affordability.

c) More damageable and expensive to repair vehicles than existed in the past.

d) More small cars being hit by big cars.

The reviewer agrees that the societal changes considered by the author will likely result in a small declining effect on average rates and loss trend for all coverages. There are other factors though that conceivably have a counterbalancing effect.

Trend

The theme of this section is to describe the strengths and weaknesses of various bases for selection of trends for physical damage coverages. Distortions caused by deductible shifts; comprehensive catastrophes; the use of first dollar PDL severity trends; and the use of PDL frequency trends to estimate physical damage frequency trends are explored among other considerations. Also considerations of premium trends are addressed.

This section of the author's paper is filled with items for discussion and thought. But to add emphasis to the author's remarks or rather to put them in perspective I attach Exhibits 1 and 2 which are rough calculations of the effect of deductible shifts on comprehensive and collision trends. The deductible shift effect on comprehensive is seen to be negligible on frequency, severity and pure premiums. For collision the effect is negligible on severity but noteworthy on frequency and pure premiums.

The point of these exhibits is to caution that the valuable information to be gained from reviewing physical damage trend data should not be ignored to the exclusive use of PDL data. Comprehensive and collision trends are subject to distortions as is the PDL data but as John states "all factors impacting premium or loss trends must be considered...judgmental modifications should be made as necessary."

Informed judgment is the actuary's best tool and I believe this is particularly true in the choice of trends.

Rating System

The theme of this section is to describe a new rating system to reflect vehicle series and model year rating. The new system is a clear example of the continued improvement in actuarial technology and the ability to handle complex information systems. The author's description is concise, well done, and highly recommended to those who have not been living with VIN for the past several years.

The only comments that I feel are in order are as follows:

1) It is my hope that we soon can support a differentiation in symbols for comprehensive and collision to reflect the likely expected loss difference between the two coverages by vehicle series.

2) I also hope that first party injury coverage differences by vehicle series can be reviewed.

3) I can testify to the large task of obtaining accurate VINs in computer records for resymboling needs.

Statistical Implications

The theme of this section is that statistical implications of rating plans or changes thereto should be considered. The author has identified several points to bear in mind as the ratemaker designs statistical plans and formats reports.

The reviewer would only add that in order for information to be valuable in managing a company that it must be understandable and significant to non-actuaries. Data that produces actions is the key and not actuarial full employment.

Summary

John's paper is a well written exposition on some of the perplexing problems to be faced by a ratemaker for personal automobile physical damage insurance.

While my perspective as an actuary for an independent company differs somewhat from John's, I believe that we both agree that the perplexing problems must be dealt with. Furthermore, I believe we also agree that actuarial judgment continues to be very important and that various ways of approaching the problems must be considered.

John deserves commendation for his excellent treatment of the subject. My review hopefully does not detract from the significance of the paper for it is meant only to facilitate open discussion. As a result the differences of opinion are highlighted rather than the agreements. The agreements far outweigh the differences but they have not been emphasized.

Company: ABC Insurance Company
State: X
Collision
Effect of Deductible Shift

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Exposure Distribution		Calendar Year 1975 Estimated Claims Statistics			1975	1978	1975	1978
Deductible	Cal. Yr. 75	Cal. Yr. 78	Severity	Frequency	Pure Premium	Col. 2 x Col. 5	Col. 3 x Col. 5	Col. 2 x Col. 6	Col. 3 x Col. 6
$80/20	60	30	$305	.155	$47.28	9.30	4.65	$ 2,837	$ 1,418
50	2,300	1,370	428	.093	39.80	213.90	127.41	91,540	54,526
100	7,560	8,040	419	.087	36.45	657.72	699.48	275,562	293,058
150	-	20	411	.083	34.11	-	1.66	-	682
200	-	80	418	.076	31.77	-	6.08	-	2,542
250	80	440	428	.069	29.53	5.52	30.36	2,362	12,993
500	-	20	481	.045	21.65	-	.90	-	433
Total	10,000	10,000				$86.44	$70.54	$372,301	$365,652

Notes: (1) Col. 4 $80/20 Actual 1975 average incurred cost (AIC)
50 Actual 1975 average incurred cost (AIC)
100 50 Ded. AIC x 100 Ded. LER = 428 x [(1032665 - 89448) ÷ (1835 - 121) ÷ (1032665 ÷ 1835)]
150 100 Ded. AIC x 150 Ded. LER = 419 x [(5831339 - 397451) ÷ (8108 - 403) ÷ (5831339 ÷ 8108)]
200 100 Ded. AIC x 200 Ded. LER = 419 x [(5831339 - 767562) ÷ (8108 - 1052) ÷ (5831339 ÷ 8108)]
250 100 Ded. AIC x 250 Ded. LER = 419 x [(5831339 - 1104772) ÷ (8108 - 1673) ÷ (5831339 ÷ 8108)]
500 100 Ded. AIC x 500 Ded. LER = 419 x [(5831339 - 2400635) ÷ (8108 - 3952) ÷ (5831339 + 8108)]

(2) Col. 5 $80/20 Actual 1975 frequency
50 Actual 1975 frequency
100 50 Ded. Frequency x 100 Ded. LER = [.093 x (1835 - 121) ÷ 1835]
150 100 Ded. Frequency x 150 Ded. LER = [.087 x (8108 - 403) ÷ 8108]
200 100 Ded. Frequency x 200 Ded. LER = [.087 x (8108 - 1052) ÷ 8108]
250 100 Ded. Frequency x 250 Ded. LER = [.087 x (8108 - 1673) ÷ 8108]
500 100 Ded. Frequency x 500 Ded. LER = [.087 x (8108 - 3952) ÷ 8108]

(3) Col. 6 = Col. 4 x Col. 5

Conclusions: a.) Frequency Based on 1975 Distribution = .0886
Based on 1978 Distribution = .0871 .0871 ÷ .0886 = .983 $(.983)^{1/3}$ ≐ .994

b.) Severity Based on 1975 Distribution = $420.00 = 372301 ÷ 86.44
Based on 1978 Distribution = 420.03 = 365652 ÷ 70.54 $420.03 ÷ $420.00 = 1.000

c.) Pure Premium Based on 1975 Distribution = $37.23
Based on 1978 Distribution = 36.57 $36.57 ÷ $37.23 = .982 $(.982)^{1/3}$ ≐ .994

Exhibit I
Sheet 1

Exhibit 1
Sheet 2

Company: ABC Insurance Company

State: X

Loss Elimination Study

Collision

$ 50 Deductible Size of Loss Study

Ded.	Amount of Settlement	No. of Claims	Total Settlement	Total Losses Eliminated
$100	$1 - 50	121	3,748	3,748
	Over 50	1,714	1,028,917	85,700
	Total	1,835	1,032,665	89,448

$100 Deductible Size of Loss Study

Ded.	Amount of Settlement	No. of Claims	Total Settlement	Total Losses Eliminated
$150	$1 - 50	403	12,201	12,201
	Over 50	7,705	5,819,138	385,250
	Total	8,108	5,831,339	397,451
$200	$1 - 100	1,052	61,962	61,962
	Over 100	7,056	5,769,377	705,600
	Total	8,108	5,831,339	767,562
$250	$1 - 150	1,673	139,522	139,522
	Over 150	6,435	5,691,817	965,250
	Total	8,108	5,831,339	1,104,772
$500	$1 - 400	3,952	738,235	738,235
	Over 400	4,156	5,093,104	1,662,400
	Total	8,108	5,831,339	2,400,635

Company: ABC Insurance Company
State: X
Comprehensive
Effect of Deductible Shift

(1) Deductible	(2) Exposure Distribution Cal. Yr. 75	(3) Exposure Distribution Cal. Yr. 78	(4) Calendar Year 1975 Estimated Claims Statistics: Severity	(5) Calendar Year 1975 Estimated Claims Statistics: Frequency	(6) Calendar Year 1975 Estimated Claims Statistics: Pure Premium	(7) 1975 Col. 2 x Col. 5	(8) 1978 Col. 3 x Col. 5	(9) 1975 Col. 2 x Col. 6	(10) 1978 Col. 3 x Col. 6
Full	9,920	9,860	$144	.072	$10.37	714.24	709.92	$102,870	$102,248
$ 50	80	130	157	.056	8.79	4.48	7.28	703	1,143
100	-	10	187	.041	7.67	-	.41	-	77
Total	10,000	10,000				718.72	717.61	103,573	103,468

Notes:

(1) Severity
- Full Coverage: Calendar Year 1975 actual average incurred cost (AIC)
- $ 50 Ded.: Full Coverage AIC x 50 Ded. LER = 144 x [(2310689 - 342747) ÷ (7525 - 1631) ÷ (2310689 ÷ 7525)]
- 100 Ded.: Full Coverage AIC x 100 Ded. LER = 144 x [(2310689 - 596055) ÷ (7525 - 3236) ÷ (2310689 ÷ 7525)]

(2) Frequency
- Full Coverage: Calendar Year 1975 actual frequency
- $ 50 Ded.: Full Coverage Frequency x 50 Ded. LER = .072 x [(7525 - 1631) ÷ (7525)]
- 100 Ded.: Full Coverage Frequency x 100 Ded. LER = .072 x [(7525 - 3236) ÷ (7525)]

Conclusions:

a.) Frequency
- Based on 1975 Distribution = .0719
- Based on 1978 Distribution = .0718   .0718 ÷ .0719 = .999 $(.999)^{1/3} \doteq 1.000$

b.) Severity
- Based on 1975 Distribution = $144.11
- Based on 1978 Distribution = 144.18 $144.18 ÷ $144.11 = 1.000

c.) Pure Premium
- Based on 1975 Distribution = $ 10.36
- Based on 1978 Distribution = 10.35 $10.35 ÷ $10.36 = .999 $(.999)^{1/3} \doteq 1.000$

Exhibit 2
Sheet 1

Exhibit 2
Sheet 2

Company: ABC Insurance Company
State: X

Loss Elimination Study

Comprehensive
Full Coverage Size of Loss Study

Ded.	(1) Amt. of Settlement	(2) No. of Claims	(3) Total Settlement	(4) Total Losses Eliminated
$ 50	$1 - 50	1,631	48,047	48,047
	Over $50	5,894	2,262,642	294,700
	Total	7,525	2,310,689	342,747
$100	$1 - 100	3,236	167,155	167,155
	Over $100	4,289	2,143,534	428,900
	Total	7,525	2,310,689	596,055
$250	$1 - 250	5,864	566,295	566,295
	Over $250	1,661	1,744,394	415,250
	Total	7,525	2,310,689	981,545
$500	$1 - 500	6,576	818,350	818,350
	Over $500	949	1,492,339	474,500
	Total	7,525	2,310,689	1,292,850

TITLE: USES OF CLOSED CLAIM DATA FOR PRICING

AUTHOR: Mr. R. Michael Lamb

Mr. Lamb is Casualty Actuary for the Insurance Division of the State of Oregon. He received his FCAS designation in 1975 and holds masters degrees from Purdue and the University of Washington in mathematics and business administration. Michael is chairman of NAIC technical task forces on workers' compensation and on medical malpractice insurance and serves as a member of several others. He is on the CAS Editorial Committee.

REVIEWER: Mr. Richard S. Biondi

Mr. Biondi is Manager and Associate Actuary for Insurance Services Office in New York. At ISO, he is the staff representative on several committees, including the Commercial Automobile, General Liability and Professional Liability Actuarial Subcommittees. Dick received his FCAS in 1979 and holds an M.S. degree from the Polytechnic Institute of New York.

A fundamental problem of pricing insurance is: When all is known about claims from an accident-or policy-year, that year is too old to be relevant for next year's coverage. Thus, our ancestors began using aggregate historical patterns to estimate how incurred costs of recent periods would mature to full ultimate value.

The common accident-year model will be referred to as a representative of these development methods. The cost of claims from an accident year can be estimated at each of several points of time. The estimate at one time divided by the estimate at the previous time is an observed development ratio. Development stages are defined by a series of evenly-spaced time intervals measured from the beginning of each accident year. The latest observed ratios for each stage are usually averaged to estimate how a recent accident year yet to reach that stage will develop when it does pass through. The compound product of development ratios over all stages after a certain stage until the end of time - or to some prudent horizon - is a development factor for an accident year which has reached that stage.

The costs used in the process are usually the estimated incurred costs of claims reported to date.

The assumption that this statistic will follow historic patterns rests on a belief that claim personnel who establish reserves are both consistent and uneducable. Using paid costs instead of incurred costs is more objective, but disregards all information about open reported claims. There is a tradeoff of advantages to be considered.

Pricing insurance is like predicting adult traits of the next unborn generation of a species. Offspring are born and then grow teeth, hair, claws or fins and learn to walk, swim, hunt or fly and grow to adult size and strength. We can observe how youngsters of past generations have passed through stages to become adults and so can predict how todays children and adolescents will ripen. This is even called "development".

But in predicting the next and future generations, we must allow for evolution. In times of rapid evolution, many previous patterns of development into adults may not be accurate because the adults will be different. It becomes necessary to examine the very latest information about members of the species at every stage.

Evolution is called "trend" by an actuary. Trend factors are calculated across accident years much as development factors are calculated across stages of maturity. So, the costs of a future accident year is estimated by essentially this formula:

$$\begin{bmatrix} \text{Cost of} \\ \text{a recent} \\ \text{accident} \\ \text{year} \end{bmatrix} \text{ times } \begin{bmatrix} \text{development} \\ \text{factor} \end{bmatrix} \text{ times } \begin{bmatrix} \text{trend} \\ \text{factor} \end{bmatrix}$$

This paper compares the common accident-year model with the uncommon closure-year model. Whereas an accident year includes all accidents or incidents occurring in a year, a closure year includes all claims reaching final disposition during a year regardless of when the incidents occurred.

Closed claim data offer the most recent objective information about final costs of insured risks. In times of uncertainty, this can be tremendously important-particularly when new methods of claim management or other aspects of claim disposition are significantly affecting costs independently from circumstances of true original incidents.

Closure-year models are uncommon because they do not represent an insurance product. Accident and policy years are more natural. Closed claim data is not temporally aligned with claims arising from, reported in, or covered by policies issued in a recent period.

Closure-year models are difficult to assemble. Ideally, you should have closure data for claims arising from all prior occurrence periods within a conservative horizon. Relying upon open claim reserves to represent early periods can cloud important distinctions between risks. Furthermore, unless the insured population is stable and your data source is universal, you must have exposure indices for each occurrence period. For application to a future coverage year, each occurrence period component of a closure year must be separately trended in the traditional algebraic model. Pure premium trends are the most natural, or they may be split into frequency and claim size portions. The use of external cost indices and trend residuals is not recommended without considerable study into how claim costs are determined by occurrence-period, closure-period, and intermediate-period influences. The simplest conception of a closure-year model for representing an accident year is

$$\begin{array}{l}\text{Accident Year (T+1)}\\ \text{cost per unit}\end{array} = \sum_{J=1}^{M} \frac{\begin{array}{c}\text{Accident Year (T-j+1)}\\ \text{Claims Closed in Year T}\end{array}}{\begin{array}{c}\text{Accident Year (T-j+1)}\\ \text{exposures}\end{array}} \times \begin{pmatrix}\text{Trend}\\ \text{Factor}\\ \text{to Year}\\ \text{T+1 from}\\ \text{(T-j+1)}\end{pmatrix}$$

where M is the number of years required for all claims to be closed. More interesting and useful models will be presented in later sections.

The most serious conceptual problems with closure-year models relate to the passage of time over long horizons. We toil and spin in a multivariate world of infinite dimension in which relations between finite sets of factors do not remain constant. Significant changes over only a few years, however, mean important variables have not been included. The inability to recognize and usefully measure important influences is the true conceptual difficulty in any model. Other time-related problems will be discussed later.

Another criticism of closure-year models is that they ignore information offered by open claims. They resemble an extreme of payment development models. This criticism, however, leads us to see the value of closed claim data as fully-developed factual information about claims now reaching final disposition. During any times of changing claim management approaches or disposition methods which may affect costs, closed claim data should at least be used to supplement an accident-year model. The algebraic construction of closure-year models suggests closed claim trends correspond to accident-year development factors and certainly can explain and guide their selection.

Illustrations in this paper will mostly be drawn from the medical malpractice claims study of the National Association of Insurance Commissioners, which is the most extensive closed claim research effort in the public domain.

AN ALGEBRAIC MODEL

The model described hereafter relies on the work of Archer McWhorter (2), with some important variation. Let us define M to be the number of years required for all claims to be closed, or at least a reasonable horizon where remaining claims may be aggregated with little loss of precision, and

$N(t)$ = the ultimate number of claims for occurrence year t,

$n(t,u)$ = *the number of claims from occurrence year t* closed during closure-year u,

$g(j)$ = the fraction of occurrence-year claims closed in the j-th year, j=1 through M,

$r(t)$ = the claim frequency trend in year t, or $N(t)/N(t-1)$.

We can first use the number of claims closed in year T to estimate N (T+1) by a set of M equations:

$$n(T-j+1,T) = \frac{N(T+1) \cdot g(j)}{\prod_{k=1}^{j} r(T-k+2)}$$

or, if the claim frequency trend is reasonably constant,

$$n(T-j+1,T) = N(T+1) \cdot g(j) \cdot r^{-j}$$

The following table illustrates the estimation of N(1979) and the set of gs for various assumptions of a constant frequency trend. Claims closed in 1978 in the jth year from date of occurrence were equally divided between the jth and (j+1)th years preceding 1979. The sensitivity of the projected claim volume to the assumed frequency trend is readily apparent.

			Estimated closing pattern for 1979 occurrence year:			
j	Paid Claims Closed in 1978	r=	1.00	1.05	1.10	1.15
1	441		.059	.049	.040	.032
2	916		.123	.106	.091	.077
3	998		.134	.121	.109	.096
4	1,194		.160	.152	.143	.133
5	1,308		.175	.175	.172	.167
6	1,047		.140	.147	.152	.154
7	676		.090	.100	.108	.114
8	388		.052	.060	.068	.075
9	200		.027	.033	.039	.045
10	112		.015	.019	.024	.029
11	57		.008	.010	.013	.017
12	35		.005	.007	.009	.012
13	24		.003	.005	.007	.009
14	18		.002	.004	.006	.008
15+	61		.008	.013	.021	.032
Projected number of claims arising from 1979 occurrences			7475	9537	12,221	15,743

Source: NAIC Malpractice Claims, Vol. 2, No. 2 (1980).

Once the rs have been specified, the above M expressions give us M+1 unknowns. Since the sum of the gs equals 1.000, solutions can be found for N(T+1) and each g(j), j=1,2,...,M.

Now we divide the range of claim sizes into L intervals using a sequence d(0), d(1), d(2), ..., d(L) where d(0) = 0 and d(L) is a coverage limit or else d(L-1) is some practical bound and d(L) is infinity. then we define:

C(k) = average claim cost between d(k-1) and d(k),

P(k|j)=probability of a claim closed in the jth year having a cost between d(k-1) and d(k),

P(j)=probability of a claim closing in the jth year,

Y(t)=total claim costs for occurrence year t,

S(t)=claim size trend in year t.

A straight forward algebraic construction of Y(T+1) is

$$Y(T+1) = N(T+1) \sum_{k=1}^{L} C(k) \sum_{J=1}^{M} P(k|j) \cdot P(j).$$

Ordinarily, P(j) = g(j). Evaluating each P(k|j) and C(k) from closed claim data would begin by examining the distribution of claims closed in year T for each of the latest M occurrence years. This subset of claims for each occurrence year may be more homogeneous then the whole and more likely to follow a theoretical pattern such as a log-normal distribution.

If a density function can be found to describe the size of claims closed in year T for occurrence year T-j+1, then $P(k|j)$ can be evaluated by the definite integral from a(k-1) to a(k), where

$$a(k) = \frac{d(k)}{\prod_{k=1}^{j} s(T-k+2)}$$

of, if the size trend is reasonably constant,

$$a(k) = d(k) \cdot S^{-j}$$

A first approximation for C(k) may be the average of closed claim amounts between d(k-1) and d(k). If L is large, then that may be sufficiently precise. Otherwise the effect on the average in each interval from a translation of the density functions could be determined using modern programmable calculators.

Those who think continuously may readily observe that if the density functions can be generalized to a joint function of both claim size and year j, then

$$Y(T+1)=N(T+1)\int_{0}^{M}\int_{0}^{d(L)} x\ f(XS^{-j}, j)\ dxdj$$

If the size of a claim is independent of the interval from incident to disposition, then

$$f(x,j) = f(x|j)g(j) = f(x)j(j) \text{ and so}$$

$$Y(T+1) = N(T+1)\int_{0}^{M}\int_{0}^{d(L)} x\ f(XS^{-j})g(j)\ dxdj.$$

A great many other algebraic models may be devised which rely on a knowledge of trends and claim size patterns.

Trending Methods

Some closed claim studies (1 and 6) have used trending methods to attempt to reduce the temporal alignment differences by "adjusting" the size of each claim to what might be expected for a common occurrence period. There have been two problems with these methods: (1) They are too simple - relying on elementary curves with only primative measures of significance or none at all, and (2) such techniques assume time passage affects claim costs totally independently of all other factors. The latter problem will be discussed in later sections. Some approaches to resolving the first problem appear here.

Closed claims have been commonly used to indicate claim size trends. Changes in the distribution of these claims among accident years, ranked by maturity, distort the patterns. Better understanding of both size and frequency trends can be gained by displaying closed claim data by closure period and maturity simultaneously. The NAIC publications (3, 4 and 5) give us good illustrations of how this may be done. Due to the obstinance of some insurers, however, we are unable to observe reliable frequency patterns from this source.

The NAIC maturity and closure period table provides ratios of claim sizes from consecutive closure periods for each maturity range. The measures of variance also shown allows us to use one-way and two-way analysis of variance on the array of ratios to determine if significant patterns are present. Regression methods may be used across closure periods within maturity ranges on either ratios or dollar amounts but only on ratios across maturity ranges within closure periods.

Interpretation of the horizontal and vertical patterns requires some premise of whether the closure period influences costs independently of occurrence periods and maturity. In the simple trending methods, this distinction is overlooked because the time spans for which trend factors make adjustment have the same width whether measured between occurrence dates or between closure dates. If closure periods have an influence, then significant differences observed between maturity ranges within closure periods could mean trend factors should differ between maturities. Otherwise, such differences describe changes in trend ratios across occurrence periods. A simple exponential trend is appropriate only of no significant differences are observed in the array of ratios. Otherwise, you must interpret the differences before you can select the series of trend factors to apply in an algebraic model such as previously described.

A reasonable way to determine whether costs are influenced by the closure period independently of the other factors is by reviewing correlations between claim costs and external indices for occurrence and closure periods. Again, the NAIC (5) has thoughtfully illustrated how this can be accomplished. Average paid claim costs are arrayed by closure and occurrence periods for various severity of injury ranges. Correlations can be tested for medical indexes, price indexes, and other economic indicators. One tenable theory is that temporary injuries or losses are compensated at actual costs in the period of occurrence while permanent disabilities are compensated with regard to prices and price changes at the time of disposition.

For a line of insurance like medical malpractice, precision may be gained by using trending methods to describe the residual claim cost changes remaining after adjustment using economic indexes and also by separately reviewing trends by type or severity of loss.

Claim Size Distributions

From the wealth of published material about size distributions come two principal mathematical probability functions: the logrithmic normal distribution and the gamma distribution. The rationale behind using the log-normal is the central limit theorem from the depths of probability theory. This theorem says the sum of items taken from similarly distributed populations tends to be normal. The size of a claim is the product of a great many factors in this multivariate world, so the logarithm of claim size is the sum of many terms. Even if we do not know all the factors, we can still consider whether observed claim size patterns are log-normal. The gamma distribution is a more general one.

Density Function:	$\dfrac{e^{(\ln x-\mu)^2}}{\sqrt{2\pi}\,\sigma\, x}$	$\dfrac{c^{\gamma}}{\Gamma(\gamma)}\, x^{\gamma-1}\, e^{-cx}$
Mean:	$e^{\mu+\sigma^2/2}$	γ/c
Variance:	$e^{2\mu+\sigma^2}\left[e^{\sigma^2}-1\right]$	γ/c^2
Mode:	$e^{\mu-\sigma^2}$	$(\gamma-1)/c$
Skewness:	$\left(e^{\sigma^2}+2\right)\sqrt{e^{\sigma^2}-1}$	$2/\sqrt{\gamma}$

The symbols μ and σ in the log-normal formulas represent the mean and standard deviation of the logarithms. The gamma formulas are determined by two constants, γ and c. For a gamma function to be of interest to us, usually γ must be greater than 1.000, which means that the mean must exceed the standard deviation. Unfortunately, this condition has ruled out gamma distributions for most types of insurance which have been critical problems recently because of the uncertainty of claim amounts as well as frequency.

A simple way to determine whether observed data might reasonably be described as log-normal is to see how well the mean and variance of the logarithms of observed claims fit the formulas above. The hypothetical mean and variance of the logarithms can be found by an iterative process since the sum of the log-mean and half the log-variance equals the logarithm of the object mean. The skewness and kurtosis of the logarithms should each be near zero. Skewness is a measure of asymetry. Kurtosis is a measure of non-normality.

Some computational formulas for skewness and kurtosis are:

$$\text{Skewness} = \frac{\sum(x^3) - 3\sum(x^2)\,(\sum x)/n + 2\,(\sum x)^3/n^2}{n\left[\frac{\sum(x^2) - (\sum X)^2/n}{n-1}\right]^{3/2}}$$

$$\text{Kurtosis}=\frac{\sum(x^4)-4\sum(x^3)(\sum x)/n+6\sum(x^2)(\sum x)^2/n^2-3(\sum x)^4/n^3}{n\left[\frac{\sum(x)^2-(\sum x)^2/n}{n-1}\right]^2}-3$$

If your data has fewer than 1000 paid claims, these computations may be more easily done on a programmable calculator than by convincing busy data processing people to give you sums, sums of squares, sums of cubes, and sums of fourth powers of both paid claim amounts and their logarithms with double precision.

Very likely, no theoretical distribution will fit observed insurance claim size data for several reasons such as these: 1) A popular premise is that small claims are overpaid and large claims underpaid. 2) Some groundless claims are paid for amounts less than probable defense costs(nuisance claims). 3) Many claims cluster about certain "target values" due to the need to approximate uncertain costs. 4) Economic factors operating over the occurrence, reporting, or closure period will "blur" the distributions. The latter effect may hinder analysis of accident-year claims as well as closure-year claims. Each is likely to be the sum of a continuum of log-normal distributions which will not be log-normal. (Products of log-normal distributions may be log-normal, but not sums.)

These departures from theoretical patterns can be simulated on programmable calculators. The log-normal distribution, for instance, can easily be generated from standard random normal number generating routines. Such efforts can be tedious but necessary if the density function cannot be modified or successfully integrated to find theoretical means, variances, and skewnesses to be compared with descriptive statistics from actual data.

The Gini Index

The Gini Index of Concentration is another interesting statistic for comparing distributions. Named for its Italian inventor, the Gini index is a tool used in economics and demographics to measure inequality of distribution.

The associated Lorenz curve, L(x), in our application, represents the fraction of total claim costs which relates to claims closed for $x or less. The Gini index, G, is the ratio of the area between L(x) and an equal distribution curve (a 45-degree line when L is plotted against percentiles) to the total area beneath such curve. If the range of claim sizes is divided into k intervals by a sequence 0, d(1), d(2), ..., d(k) and p(t) is the percentage of paid claims at d(t) or less, then a standard method for calculating the Gini index is:

$$G = 1 - \sum_{t=o}^{k-1} (p(t+1) - p(t)) \left[L(d(t+1)) + L(d(t)) \right] .$$

Since the indexes are linearly related, very likely a great many dramatically different distributions could have the same Gini index (9). Nonetheless, the Gini index can be appropriately applied to a set of distributions understood to be substantially similar. Changes in the coefficient of variance, skewness, kurtosis, modality, or any feature in the shape of a distribution will affect the Gini index.

The Gini index comes from a class of statistics with asymptotically normal distributions, known as "U-statistics". Tests for significance of differences exist beyond the scope and allowed length of this paper (8). Multiplying a distribution by a constant does not change the Gini index, but the index is sensitive to the number and selection of data points used in its construction (11). Our application concerns substantially similar distributions with the same means, so the use of a static set of data points should not distort comparisons of Gini indexes.

The following exhibit illustrates a comparison of claim distributions according to descriptive statistics and Gini indexes. The pattern shown first is for 421 paid claims closed in the second half of 1977 arising from occurrences in the first half of 1974 as reported

COMPARISON OF CLAIM DISTRIBUTIONS

Size of Claim	NAIC Malpractice Claims from first half of 1974 closed during the second half of 1977		Random generation from a log-normal distribution		Random generation from log-normal with small claims overpaid and large claims underpaid	
	Number of claims	Amount of claims	Number of claims	Amount of claims	Number of claims	Amount of claims
$ 1 to $1,999	80	$ 74,080	1,571	$ 1,654,718	226	$ 211,673
2,000 to 4,999	101	308,656	1,804	6,071,217	191	626,524
5,000 to 9,999	64	427,712	1,785	12,976,943	170	1,230,519
10,000 to 19,999	55	716,925	1,698	24,339,768	157	2,212,814
20,000 to 49,999	59	1,671,057	1,738	55,477,202	141	4,461,099
50,000 to 99,999	32	2,029,632	818	57,184,521	63	4,481,277
100,000 to 199,999	19	2,634,312	378	52,642,581	29	4,199,820
200,000 to 499,999	7	1,690,535	169	49,257,642	12	3,063,956
500,000 to 999,999	3	1,713,312	26	17,331,377	10	7,315,137
1,000,000 and over	1	1,400,000	13	22,441,297	1	1,251,388
	421	$12,666,221	10,000	$299,377,266	1000	$29,054,207
Mean		$30,086		$29,938		$29,054
Standard Deviation		92,263		87,727		91,541
Log-Mean		8.89		9.15		8.84
Log-Standard Deviation		1.64		1.54		1.67
Gini Index		.763		.704		.765

by the NAIC (5). A log-normal distribution with the same mean and standard deviation should have a log-mean of 9.14 and a log-standard deviation of 1.53. a random generation of 10,000 log-normal numbers with approximately the same sample mean and standard deviation is shown for comparison. Several smaller simulations strongly suggest the differences in these statistics and in the Gini indexes is more than random.

The higher Gini index for the NAIC data suggests a greater peakedness which might plausibly be explained by the hypothesis that small claims are overpaid and large ones are underpaid. A random generation of 1000 claims from a population with log-mean 8.83 and log-standard deviation 1.70 produced the third pattern shown in the exhibit after claim amounts less than $1000 were amplified by a factor which increased from 1.00 to 2.00 as the amount decreased from $1000 to $0 and the excess portion of claims over $500,000 was multiplied by .75. The differences between this third pattern and that of the NAIC data are well within the bounds of random variation.

DIFFERENCES BETWEEN RISKS

The greatest value of closed claim data is as factual information on the costs of insuring certain risks. If our fears about temporal alignments can be overcome, we should be anxious to use closed claim data for determining the attributes or classification schemes which distinguish individual risks. Recent accident-year data refined by class or attribute may not disclose differences, or may display false differences, since development factors are typically based on aggregated data.

One way to examine the claims costs effect of a certain attribute, such as smoking for drivers or board certification for physicians, would be to construct separate algebraic models for each data group. But that would be subject to the weaknesses of all the trending methods and would be a laborious task, especially if the attribute has several values.

Multivariate statistical methods can be straight forward and enable the researcher to either control or manipulate several variables at once. Hence, if temporal alignment is feared to be influencing comparative observations from closed claim data, then a sensible remedy should be to include time values in a multivariate analysis. Such methods are able to recognize the interactions of factors, so the earlier criticism of simple trending methods which assume time affects costs independently of other factors becomes less worrysome.

Multivariate statistical methods have an advantage of precision. The significance of differences observed for any any variable is measured by comparing those differences to measures of "error", "unexplained", or "random" variance. When other factors are present which "explain" additional portions of the variance, then the error variance is reduced. Seen this way, multivariate methods are indispensible for reviewing closed claim data.

In the remaining pages attention is given to analysis of variance and multiple regression. Thorough and understandable discussions of these methods can be found in the references (12 and 14). Brief mention will be made of more advanced methods and their application.

Analysis of Variance

The NAIC studies (4 and 5) have a great many illustrations of analysis of variance. The basic concept is very elementary. Variation between group or cell means is compared against residual or random variation "within" groups. For several factors and several groups, it is analogous to tests using the standard t-statistic for two group means.

The F-test for significance assumes the populations are normally distributed. The computation process assumes homogeneity of the variances within groups and that the dependent variable has continuous measure with equal intervals. The latter assumption is not of concern in our applications. Non-homogeneity of variance increases the residual variance and makes the F-test more conservative. Simulation models with pure premium distributions - products of non-normal frequency and amount distributions - on a programmable calculator have found non-normality to also reduce F-ratios. Nevertheless, the NAIC analysis (4 and 5) have found several large F-ratios. The conclusion is that analysis of variance with standard F-tests is very robust. Nonparametric analysis of variance methods based on rankings may be used to verify results, but are less powerful for detecting false hypotheses.

The following tables and calculations illustrate the concept of analysis of variance with two independent variables. The illustrated analysis seeks to determine whether the average cost of physicians' malpractice claims differs by type of practice and uses the year of claim disposition as a "control" variable. Occurrence year may be a more natural control or possibly both occurrence year and time required for disposition could

PHYSICIANS AND SURGEONS
MALPRACTICE CLAIMS BY TYPE OF PRACTICE AND YEAR OF CLAIM DISPOSITION

Type of Practice		Closure Year 1975	1976	1977	1978	Total
Institutional	Claims	88	108	86	89	371
	Indemnity	1,420,734	1,313,063	1,607,970	2,695,430	7,037,197
Prof. Corp. or Ptnship	Claims	873	1,370	1,333	1,585	5,161
	Indemnity	18,646,181	27,808,980	33,798,124	54,212,985	134,466,270
Self-Employed	Claims	1,775	2,674	2,457	2,730	9,636
	Indemnity	37,207,032	60,277,675	57,968,550	89,203,413	244,656,670
Employed	Claims	150	201	293	219	863
	Indemnity	2,705,521	3,188,934	7,433,108	5,423,285	18,750,848
Resident	Claims	3	10	8	13	34
	Indemnity	194,333	128,705	116,875	100,411	540,324
Total	Claims	2,889	4,363	4,177	4,636	16,065
	Indemnity	60,173,801	92,717,357	100,924,627	151,635,524	405,451,309

Total sum of squares of raw amounts: 68,739,329,480,645

Source: <u>NAIC Malpractice Claims</u>, Vol. 2, Number 2 (1980).

Computations for Analysis of Variance

Correction from raw amounts to deviations from the mean $= \frac{(405,451,309)^2}{16,065} = 10,232,851,796,051$

Total Sum of Squares $=$

$$\begin{array}{r} 68,739,329,480,645 \\ \underline{-10,232,851,796,051} \\ 58,506,477,684,594 \end{array}$$

Between All Groups Sum of Squares $= \sum_{\text{types}} \sum_{\text{years}} \frac{(\text{indemnity})^2}{\text{claims}} - C$

$$\begin{array}{r} = 10,684,688,432,954 \\ \underline{-10,232,851,796,051} \\ 451,836,636,903 \end{array}$$

Between Types of Practice Sum of Squares $= \sum_{\text{types}} \frac{(\sum \text{indemnity})^2}{\sum \text{claims}} - C$

$$\begin{array}{r} = 10,264,702,339,317 \\ \underline{-10,232,851,796,051} \\ 31,850,543,266 \end{array}$$

Between closure years Sum of Squares $= \sum_{\text{years}} \frac{(\sum \text{indemnity})^2}{\sum \text{claims}} - C$

$$\begin{array}{r} = 10,621,930,858,946 \\ \underline{-10,232,851,796,051} \\ 389,079,062,895 \end{array}$$

Interaction Sum of Squares

$$\begin{array}{r} = 451,836,636,903 \\ - 31,850,543,366 \\ \underline{-389,079,062,895} \\ 30,907,030,742 \end{array}$$

Residual Sum of Squares

$$\begin{array}{r} = 58,506,477,684,594 \\ \underline{- \quad 451,836,636,903} \\ 58,054,641,047,691 \end{array}$$

FINAL ANALYSIS OF VARIANCE TABLE

	Degrees of Freedom	Mean Squares	F-ratio	level of significance
Between types of practice	4	7,962,635,817	2.203	.066
Between Years	3	129,693,020,965	35.887	.000
Interaction	12	2,575,585,895	.713	.740
Residual	16,045	3,613,959,229		
TOTAL	16,064			

be used in a three-way analysis of variance. Closure year was selected for ease of illustration.

In the final Analysis of Variance Table, the sums of squared deviations from the means ("sum of squares") are divided by the statistical degrees of freedom to achieve "mean squares", which are the estimates of variance used in this process. Each of these is compared with the residual mean squares to determine the level of significance. Note that if closure years had not been included in the analysis, the residual mean squares would have been greater, the F-ratio for types of practice would have been lower, and the level of significance would have been greater. (The level of significance is the probability observed differences could occur randomly.)

Multiple Regression

Multiple regression estimates the magnitude of relations between factors and has more general capability then analysis of variance. Control variables can be included more naturally. If the number of observations in the groups or cells are unequal, multiple regression is preferred. There is an R^2 statistic to describe the portion of total variance "explained" by the set of independent variables and an F-ratio for significance.

However, the calculations are much more extensive. Desk calculators are impractical beyond four or five independent variables.

The desired expression is of the form

$$Y=A+B(1)X(1)+B(2)X(2)+\ldots+B(N)X(N)$$

where Y is the dependent variable, the set of X's is the set of independent variables, and the coefficients A and B(1) to B(N) are found to minimize the squared deviations from the predicted values. The process requires solution of a set of equations:

$$\begin{array}{l} r(1,1)\,b(1)+\ldots+r(1,N)b(N)=r(y,1) \\ \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ r(N,1)b(1)+\ldots+r(N,N)b(N)=r(y,N) \end{array}$$

where r(i,j) is the correlation between X(i) and X(j), r(y,i) is the correlation between Y and X(i), and b(i) is the standardized regression coefficient.

The importance of a single factor X(k) is usually evaluated by the significance of the contribution it makes to R^2. Multiple regression strategies are a modern art form, admirably discussed by Cohen and Cohen (12). For our applications, the preferred strategy apparently is to first include the necessary control variables such as time of occurrence, determine an R^2

for this limited set of independent variables, then add the particular variable of interest such as age, gender, or marital status, and redetermine R^2.

Because of random fluctuations, any variable added to the set of independent variables will always increase R^2. Most researchers prefer to use a corrected or "shrunken" R^2 which is a better estimate of the population R^2. With k independent variables and sample size n, the corrected value is:

$$R_c^2 = 1 - (1-R^2)\ \frac{n-1}{n-k-1}.$$

In the simple case of two independent variables, we can define the simipartial correlation, sr, of Y and X(2) to be the correlation between Y and X(2) not related to X(1). Then,

$$R^2=\frac{(r^2(y,1)+r^2(y,2)-2r(y,1)r(y,2)r(1,2))}{1-r^2(1,2)}$$

$$sr^2=(r(y,2)-r(y,1)r(1,2))^2/(1-r^2(1,2))$$

$$F(X(2))=sr^2(n-3)/(1-R^2).$$

Independent variables must be discrete or nominal for analysis of variance, but may be continuous for multiple regression. Continuous variables usually provide the greatest information value.

Discriminant Analysis

A set of independent variables may be used to estimate group membership as the dependent variable.

Most applications discriminate between two groups, but discriminant analysis can be adapted to three or more groups. Some natural applications for insurance are classification of risks and answering claim management questions: Which claims will be paid? Which claims will include law suits? What will be the outcome of arbitration? Which claims will reopen?

Discriminant analysis is becoming recognized as a highly sophisticated risk management tool. As soon as any untoward incident occurs, the particulars may be fed into a discriminant function at a computer terminal and the likelihood of a compensable event is rapidly determined. The risk manager can promptly act to contain the costs. This technology is being introduced at hospitals in various parts of the country. The basic data is from closed claims. The exclusion of incidents which have not produced claims may not seriously reduce the predictive accuracy in many instances. Even if extensive incident data is available, insurance claim costs are clearly necessary in corresponding detail.

The selection of predictive variables is another modern art form. Separate discriminant functions should be constructed for nominal variables such as gender, marital status, and medical specialty. Astonishingly, negative or highly positive correlations between independent variables increases discriminatory power (13).

After discrimant analysis has been used to predict which claims will result in payment, a natural step is to use multiple regression to estimate the amount of payment for each - a loss reserving method.

A genuine time problem may result if such techniques are based on internal data sources only. For instance, if a hospital constructed and periodically revised its discriminant function for compensable events based on its own data, then its own success at decreasing costs would also decrease the predictive accuracy of the predictive variables. Costs would then increase again until predictive power is reestablished. The insurance industry has cycles like that.

Factor Analysis

Factor analysis is an extremely complex computational methodology for discovering natural dimensions behind a number of simple quantitative measures. Psychological tests, for example, measure qualities by asking a great many questions. Most often researchers are not aware of the fundamental dimensions and must seek to learn these from many simple measures.

This analytic technique may eventually be used to find comparatively few important complex dimensions represented by the several hundred variables in closed claim data collection instruments of recent studies.

DESIGNING CLOSED CLAIM DATA

Underwriting, pricing, loss reserving, claim management, and loss prevention are only separated by brief steps and perspectives. The fact that claim files typically contain only sufficient information to establish coverage, establish defenses, and compute payments should not validly prevent us from seeking information important for other functions.

The data should be designed to answer important questions or test important theories. If the task is so well defined, then questions can be easily imagined relevant to the hypothesis and sample sizes determined from formulas in the statistics books.

Unfortunately, the examples of closed claim studies in the public domain have arisen from crises in various kinds of liability insurance where there have been low frequencies, phenomenal variances, myriads of socio-underwriting theories pointing in all directions and often conflicting, no deadlines for new theories, unresponsive rating systems, and simplistic ratemaking methods.

Because of these situations, few assumptions could be made about expected patterns or variances from them and no specific lists of hypothesis could be prescribed in advance. The first purpose of the closed claim studies has been to provide an understanding of the statistical dimensions and to evaluate the importance of hypotheses as they become expressed. Classical research designs and statistical analysis have had to come second.

The closed claim data collection instruments have had to be comprehensive. With no reliable knowledge of what factors may be importantly related to claim occurrences or costs, or of the nature of those relations, or of the variances from such, no sampling techniques could be intelligently chosen and no data item could be dismissed. Hence, the forms have been designed to describe as completely as possible the insured, the claimant, the relations of the insured and the claimant, the incident, the relations of the insured and the claimant to the incident, other persons and factors related to the incident, the loss endured by the claimant, the paths taken to final disposition, and the resulting indemnities and expenses. Then hopes have been expressed that the forms were not so formal as to preclude other significant factors from being discovered.

Actuaries bent on substituting fact for impression should understand this background and learn from it how closed claim data can be an imaginative source for designing responsive rating systems, observing trends, and answering important questions before crisis situations occur.

REFERENCES

CLOSED CLAIM STUDIES

1. All-Industry Committee, Special Malpractice Review: 1974 Closed Claim Survey - Technical Analysis of Survey Results, Insurance Services Office, New York, November 1976.
2. McWhorter, Archer, Jr., "Drawing Inferences from Medical Malpractice Closed Claim Studies", Journal of Risk and Insurance, March 1978, pp. 79-94.
3. NAIC Malpractice Claims, National Association of Insurance Commissioners, Vol. 1, Number 4, May 1977.
4. NAIC Malpractice Claims, National Association of Insurance Commissioners, Vol. 2, Number 1, December 1978.
5. NAIC Malpractice Claims, National Association of Insurance Commissioners, Vol. 2, Number 2, 1980.
6. Product Liability Closed Claim Survey: A Technical Analysis of Survey Results, Insurance Services Office, New York 1977.

GINI INDEX

7. Gastwirth, Joseph L., "The Estimation of the Lorenz Curve and Gini Index", The Review of Economics and Statistics, August 1972, pp. 306-316.
8. Hoeffding, Wassily, "A Class of Statistics with Asymptotically Normal Distribution", The Annals of Mathematical Statistics, Vol. XIX, 1948, pp. 293-325.
9. Morrow, James S., "Toward A More Normative Assessment of Maldistribution: The Gini Index", Inquiry, Vol. XIV, September 1977, pp. 278-292.
10. Paglin, Morton, "The Measurement and Trend of Inequality: A Basic Revision", The American Economics Review, Vol. LXV, Number 4 (September 1975), pp. 598-609.
11. Petersen, Hans-Georg, "Effects of Growing Incomes on Classified Income Distributions, the Derived Lorenz Curves, and Gini Indices", Econometrica, Vol. 47, Number 1 (January 1979).

ANALYTICAL METHODS

12. Cohen, Jacob and Patricia Cohen, Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, John Wiley & Sons, New York, 1975.

13. Eisenbeis, Robert A., "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics", The Journal of Finance, Vol. XXXII, Number 3 (June 1977), pp. 875-900.
14. Kerlinger, Fred N., Foundations of Behavioral Research (Second Edition), Holt, Rinehart, and Winston, New York, 1973.

USES OF CLOSED CLAIM DATA FOR PRICING
BY R. MICHAEL LAMB

Reviewed by RICHARD S. BIONDI

I looked forward to reviewing this paper because, first, I was intrigued by the title (I couldn't understand how closed claim survey data could possibly be used for insurance pricing) and also because I thought I might learn something about closed claim surveys. It also seemed interesting to review a paper written by someone from the regulatory ranks since, as a rating organization employee, I would probably find many areas of healthy disagreement. Anyhow, here goes:

It seems that the first and most fundamental question about this paper is "how can closed claim data be used for insurance pricing without exposures?" The answer, of course, is that it can't, at least not by itself. Nobody has yet figured out a way to calculate rates without exposures (or at least premiums) and, even more important and difficult, the earned exposures must correspond to the same population and to the same time period as the losses. Perhaps it should be mentioned that from a data processing point of view, processing of exposures data is much more difficult and expensive than processing of losses, both closed and open, because there are so many more exposure transactions.

Mr. Lamb recognizes that exposures are important on page 223 when he says ... "unless the insured population is stable and your data source is universal you must have exposure indices for each ocurrence period". Actually, looking at things from the viewpoint of my employer, ISO, the population is never completely "stable" since we rarely have the the entire population and since some companies affiliate and disaffiliate with ISO each year, it is necessary to take great pains to make sure that comparable losses and exposures are present at all times. This is difficult because, for example, health care facilities may self insure one year and insure with an ISO carrier the next.

Runoff on losses must be reported long after reporting of premiums has stopped. All insurers, even doctor owned JUA's, must be made to understand their reporting obligations. For reasons such as these, even the NAIC should not blithely assume they have statistics for the entire population.

Anyhow, closed claim surveys don't capture exposures. What good are they then for ratemaking? Mr. Lamb mention several uses which I wish to discuss below:

1) TREND

Data can be used to measure severity trends. For ratemaking, two types of trend factors required are severity and frequency. Frequency trends can't be measured with closed claim survey data because no exposures are available. Severity trend, however can be measured. Some problems still exist however.

First, a long time period is required. Three or four years is probably not enough time to really measure severity trend for malpractice. The reason is that the malpractice severity for one claim can be represented mathematically as a random variable having an extremely large variance and, even if thousands of claims are collected, the variance of the mean is still rather substantial. Hence, at any given time, the average severity will fluctuate randomly about the expected value which would result from an infinite sample. Illustration: In Mr. Lamb's data, the standard deviation of the claim size distribution approximately equals \$60,000[1]. Hence, if 5,000 claims are included per year in the trend data, the average claim cost can be represented as a random variable having a standard deviation of $\$60{,}000/\sqrt{5{,}000} = \$848$.) Several years of data (7 or 8) are required to average out random fluctuations.

Second, closed claim trends can be distorted by changes in settlement patterns. Assume, for example, that small claims tend to be settled quickly, and large claims are settled more slowly and that insurers change their practices so that they fight more large claims rather than settle quickly. If this happens then initially the average closed claim severity will drop, only to rise back up again several years later. If they decide to fight more small claims, the same situation will occur in reverse.

Another distortion arises if, for example, the number of incurred claims suddenly rises appreciably (because of either increasing claim frequency or just an expanding data base). In this instance, since small claims are closed first, the closed claim data suddenly shows a temporary large influx of small claims along with a big drop in average severity.

Finally, the use of total limits data distorts trends since if insureds purchase higher policy limits over time, average claim costs will rise, all other factors held constant. To properly take this into account for ratemaking, it is necessary to know the policy limit corresponding to each claim.

On page 221, Mr. Lamb says, "using paid (i.e. closed) costs instead of incurred cost is more objective, but disregards all information about open reported claims. There is a tradeoff of advantages to be considered". I'm sure there is a tradeoff of advantages (there always is) and I also believe that use of incurred costs to determine ratemaking trends is better for slow developing lines because information on open claims is used. I'm not sure that use of paid costs is more "objective", especially in situations where reserving methods remain unchanged, but where payout patterns change. It would be nice to compare incurred trend data to closed claim trend data to check whether any of the various possible distortions are really significant. This has been done in my Attachment 1 where it is found that the two sets of data compare well in spite of the possible distortions mentioned above as well as the facts that the two sets of data were collected by two different organizations for different populations of claims.

2) LOSS DEVELOPMENT

Closed claim data can theoretically be used to measure paid loss development on either a policy year or accident year basis. Although, this is possible theoretically, in practice it would be necessary to accumulate closed claim data for over 10 years to really obtain loss development factors since malpractice claims tend to develop over long intervals. Furthermore, incurred loss development has generally been more valuable than paid loss development for medical malpractice insurance pricing since incurred loses develop much more rapidly.

[1]On page 224 , residual mean squares = 3.614×10^9 = variance of claim size distribution for an individual class and year. Standard deviation = $\sqrt{3.614 \times 10^9}$ = \$60,117

3) CLASSIFICATION DEFINITIONS

This may be the greatest application of closed claim survey data for pricing since, in a closed claim survey, much detail is requested that, for cost reasons, is not requested in statistical plans. This additional detail can be used profitably to define rate classifictions which have statistically significant differences in experience. The only problem is that the only data available is severity data, not frequency data. This is too bad since the ISO statistics tend to show that the greatest pure premium differences between the various medical malpractice rate classifications tend to be caused by frequency differences, rather than severity differences. Still, the author carries through an interesting example in which he tests whether the observed differences in average severity between four risk classifications could reasonably be due to random variation.

In the example, begining on page 240, the author reviews data consisting of losses and claims separately grouped by class within closure year. Five classes were included differentiating types of practice, i.,e. "Institutional", "Professional Corporation or Partnership", "Self-Employed", "Employed" and "Resident". Four years of data are shown. From the data, calculations were made of the overall variance of the claim distribution, and the variances of the class and years groupings (i.e. the variance "explained" by the classifications versus the "unexplained" variance).

An F-test was then performed to determine whether the variation in the data from class to class could reasonably be explained as random or whether the odds were overwhelming that real differences existed between the classes. The year to year variation (undoubtedly caused by inflation) could not possibly be attributed to random fluctuation. The class-to-class variation could be random, even though the size of the data base was large, i.e. $405,000,000.

One statement that the author makes on page 245 seemed misleading i.e. "Note that if closure years had not been included in the analysis, the residual mean squares would have been greater, the F-ratio for types of practice would have been lower and the level of significance (of the data variation on the classes) would have been greater". Although this statement is true the effect is negligibly small and in fact, no matter how the test is performed the classes can't be conclusively shown to be statistically different based upon the data provided.

CLAIM SIZE DISTRIBUTIONS

Closed claim survey data can be used to obtain claim size distributions useful for the determination of increased limits factors. As usual, there are advantages and disadvantages to using closed claim data over occurrence data for this purpose. If, for example, companies bulk reserve at all for malpractice claims, the occurrence claim size distribution would tend to be artificially distorted at the lower claim size intervals. On the other hand, since large closed claims tend to be very old, much trending is necessary to adjust closed claim distributions to a present cost level. Since any trend procedure carries with it many judgemental assumptions, the claim size distribution, based upon closed claims only, becomes largely a function of whatever trend assumptions were used.

The author discusses in depth the use of the long-normal curve as a best-fit approximation to actual claim size distributions. He finally concludes that the log-normal curve does not provide a particularly good fit by using a number of formulas for skewness and kurtosis (whatever that is) and also by calculating a "Gini Index of Concentration". One thing that surprised me about all this was that I think he could have proved the same thing (perhaps better) by doing a chi-square test on the data.

The chi-square test is much more widely known and is, in fact, included in the material which must be mastered to pass the Part 2 CAS exam. I have attached to this review (Attachment 2) a chi-square test of the closed claim data which seems to show that a log-normal curve does not fit malpractice claim data well. If the methods discussed by the author were actually superior, I would have liked to see a little more explanation of this.

The author also briefly discusses some of the mathematics underlying multiple regression techniques. This is a subject that we have some experience with at ISO since we use the multiple regression approach to develop insurance trend models, making use of forecasts of data indices external to insurance. The work that ISO is doing is difficult, not so much because of the mathematics involved, but because judgement and experience is required to properly select the external indices (if they exist at all) and to judiciously interpret the results. The author doesn't really discuss any specific applications of multiple regression techniques which relate to insurance pricing.

Two other mathematical techniques that the author briefly mentions are Discriminant Analysis and Factor Analysis. I didn't know anything about either of these before I read the paper and don't know much now, other than that both techniques are "highly sophisticated" and "extremely complex." Apparently these methods can be used to, for example, evaluate the likelihood of a claim being paid given many details about the claim and about other claims paid previously. One statement I found puzzling was that "the exclusion of incidents which have not produced claims may not seriously reduce the predictive accuracy in many instances."

Overall, I found the paper interesting and informative, to some extent because the author often expresses a point of view different from that which I normally hear at a rating organization. I certainly share the author's conclusion that we should strive to find better ways to evaluate the confusing array of data (all data including closed claim data) that we are paid to work with.

Attachment 1

COMPARISON OF AVERAGE
CLOSED CLAIM SEVERITY DATA WITH
POLICY YEAR AVERAGE INCURRED SEVERITY DATA

PHYSICIANS, SURGEONS AND DENTISTS

(1) Policy Year Ended	(2) Average Incurred Severity (000's)	(3) Closure Year	(4) Average Paid Severity (000's)
12/31/70	$16.8		
12/31/71	20.8		
12/31/72	19.7	1975	$20.8
12/31/73	22.9	1976	21.3
12/31/74	24.2	1977	24.2
12/31/75	34.0	1978	32.7

Explanation:

Data in column (2) is ISO total limits increased severity data for all physicians, surgeons and dentists classifications reported to ISO. Losses include all allocated loss adjustment expenses, are evaluated as of March 3, 1978 and are developed to 135 months of maturity.

Data in column (4) is taken from page 242 of Mr. Lamb's paper. Losses are divided by claims for each of the four years shown for all of the classes combined.

DETERMINATION OF WHETHER LOG-NORMAL CURVE
PROVIDES REASONABLE FIT TO CLAIM SIZE DISTRIBUTION DATA
USING CHI - SQUARE TEST

(1) Size of Claim	(2) Ln of Size	(3) Number of Claims	(4) Expected Number of Claims	(5) $\frac{((3)-(4))^2}{(4)}$
$ 1 to $ 1,999	0 to 7.60	80	90.6	1.240
2,000 to 4,999	7.60 to 8.52	101	81.9	4.454
5,000 to 9,999	8.52 to 9.21	64	70.0	.514
10,000 to 19,999	9.21 to 9.90	55	64.6	1.427
20,000 to 49,999	9.90 to 10.82	59	62.7	.218
50,000 to 99,999	10.82 to 11.51	32	27.1	.886
100,000 to 199,999	11.51 to 12.21	19	14.1	1.703
200,000 and over	12.21 and over	10	9.0	.111
				10.553

Explanation:

Columns (1) and (3) are taken from the closed claim survey data on page 237 of the paper. The numbers in column (4) are determined by assuming that the number of claims per unit log-interval is normally distributed with mean 8.89 and standard deviation 1.64 (taken from paper). The Chi-square test simply assumes that:

(a) Within a given interval the actual number of claims in a sample can be represented by a random variable with a Poisson distribution where the expected number of claims equals the variance. (This should be a good assumption for malpractice claims where multiple related claims are uncommon). When the expected number of claims exceeds 10, the distribution essentially becomes Normal.

(b) The numbers listed in column (5) are each Chi-square distributed, each with order 1. The sum of the numbers should be Chi-square distributed with order 8-3 = 5. 3 degrees of freedom should be subtracted because the mean and standard deviation are taken from the data sample and the number of claims in the 8th interval is automatically determined by subtracting the first 7 from the total.

Conclusion:

The total of column (5) equals 10.553. For a Chi-square distribution with 5 degrees of freedom, the statistic should be less than 9.236 90% of the time. Hence it is concluded that the log-normal curve does not fit the data well.

TITLE: RATING CLAIMS-MADE INSURANCE POLICIES

AUTHOR: Mr. Joseph Marker

Mr. Marker is Actuary with the Westfield Companies. He received his FCAS in 1977 and holds a B. A. from the University of Michigan as well as an M. A. from the University of Minnesota.

Mr. F. James Mohl

Mr. Mohl is Senior Actuarial Officer for the St. Paul Fire and Marine. He received his FCAS in 1975 and serves on the CAS Education Committee.

REVIEWER: Mr. Michael F. McManus

Mr. McManus is Associate Actuary with Chubb Insurance Company. He received his B. S. from St. Peter's College in New Jersey and earned his FCAS in 1977. Mike is a member of the American Academy of Actuaries and is a member of the CAS Examination Committee.

I. Introduction

In this paper, we propose to discuss the claims-made approach to pricing Medical/Professional Liability insurance. We will begin with a brief summary of the historic context which lead the largest medical malpractice writer in the country (St. Paul Fire and Marine) to switch its book of business to claims-made. Then we will describe in depth the claims-made concept itself: how it works, how it differs from traditional occurrence coverage, what its inherent advantages are, and what special problems it presents and how these might be resolved. In particular, we will compare the accuracy of claims-made and occurrence ratemaking under varying assumptions about a changing claim environment. We will outline special features of The St. Paul filings which distinguish them from previous claims-made filings by other carriers. Finally, we will highlight special analytic tools which were developed to price claims-made coverages, and will show how these same tools can aid the actuary in pricing and reserving occurrence coverage as well. Let us look back at the time before claims-made to see how the decision to offer this coverage evolved.

II. Claims-Made: A Historic Perspective

The 1950's were an era of steady growth and moderate inflation. Insurance companies generally did well. And "malpractice insurance", as it was then called, was particularly favorable, although it was such a small part of most companies' total book of business that they did not bother to distinguish it from other general (non-automobile) liability lines. Rates were very low and stable--or even falling--throughout the period. Suing a doctor was almost unheard of; suing and winning was even more unusual. Medical Liability insurance was regarded "peace of mind" coverage, if it was thought of at all.

In the 1960's the situation began to change, as inflation gradually accelerated throughout the decade. Moreover, "social inflation"--a term coined to describe the inflation in value of a tort in the minds of plaintiffs, attorneys, judges and juries--consistently ran at a higher rate than economic inflation, adversely affecting claim severity in all liability lines. Compounding the increase in severity was an increase in frequency, brought about in part by a "psychology of entitlement"--a feeling that an injured party should be compensated even if negligence had not been proven. This took the form of the erosion of traditional tort defenses, especially in the malpractice area. Still, at the end of the decade, Medical Liability did not appear to have deteriorated as much as some other liability lines. Appearances were deceiving, however, since Medical Liability insurers simply failed to recognize the impact of claims that had been Incurred But Not Reported.

In the early 1970's the insurance industry was hit with a triple whammy: severe recession resulting in the steepest plunge in the stock market since the 1930's, soaring economic inflation, and price controls which held back rate increases while doing nothing about social inflation. The combination of inflation and price controls lead to inadequate rates on current business. Worse yet, reserves on prior years--particularly reserves for Incurred But Not Reported claims--had to be increased at the same time. Only much later did it become apparent how unprofitable results from the late 1960's and early 1970's really were. This placed further pressure on the companies' surplus, already dwindling due to the stock market collapse. There were some who charged that insurers were trying to make up for stock market losses by raising rates for their policyholders. All lines were affected by these conditions to some degree but the "long tail" Medical/Professional Liability lines were particularly susceptible due to their high ratio of reserves to premium.

Some Medical Liability insurers responded to the "malpractice crisis" by seeking astronomical rate increases. However, company actuaries had difficulty in estimating IBNR and justifying it to regulators. Where rate increases were granted, cries of "unaffordability" could be heard. Where they were not granted, carriers pulled out of the market or, in at least one case, went bankrupt. Availability at any price became a real concern.

For The St. Paul the situation was critical, even assuming we could obtain any rate increase sought--a highly questionable assumption. Costs were spiralling at such a rate that what had once been a minor line was now large enough to place the entire company in jeopardy if we guessed wrong about the IBNR. Something had to be done to cut the exposure presented by the tail: either get out of the business or find a way to expose ourselves to that risk a year at a time and price it a year at a time. Out of that idea grew the decision to try a claims-made approach.

III. Claims-Made Coverage Concepts

The basic idea of claims-made coverage is simple: a claims-made policy covers claims reported ("made") during the policy period, regardless of when the underlying accident occurred. This contrasts with an occurrence policy which covers claims occurring during the policy period. "Claims-made" is not a new concept. Insurers have traditionally written some professional liability lines and many bonds on a claims-made basis.

The St. Paul has modified the claims-made policy concept to meet the specific needs of professional liability insureds. First, an insured who is in his first year of professional practice does not need coverage for acts which occurred prior to his beginning practice. The same is true of an insured who begins claims-made coverage after letting an "occurrence basis" policy expire. These insureds need coverage restricted to accidents occurring on or after the date that they first began insuring on a claims-made policy basis. This need is met by placing a "retroactive date" on a claims-made policy and restricting policy coverage to accidents occurring on or after that date. Second, the insured needs coverage for claims reported after he retires from his occupation -- "tail coverage" policies provide the necessary coverage. This coverage is also needed in case of death, disability, or simply changing insurance carriers.

At this point, we will define some of the coverage terms which will appear throughout the remainder of the paper. A convenient way to explain the coverages is to define the occurrences covered in terms of accident period covered and reported period covered. This turns out to be convenient in pricing because insurance loss data contains both accident date and reported date. If we can define the occurrences covered in terms of these dates, then we can price the policies using insurance loss data, even though that data may be collected under "occurrence basis" rather than claims-made policies.

The matrix of losses by accident year and reported year that we use is indexed differently than the traditional system of using reported year for one dimension and accident year for the other. For convenience we replaced accident year by "accident year lag" which is computed:

Accident Year Lag = Reported Year - Accident Year

We visualize losses in the Figure 1 matrix:

REPORT YEAR (J)

LAG (I)		1	2	3	4	5	6	7
	0	$L_{0,1}$	$L_{0,2}$	$L_{0,3}$	$L_{0,4}$	$L_{0,5}$	$L_{0,6}$	$L_{0,7}$
	1	$L_{1,1}$	$L_{1,2}$	$L_{1,3}$	$L_{1,4}$	$L_{1,5}$	$L_{1,6}$	$L_{1,7}$
	2	$L_{2,1}$	$L_{2,2}$	$L_{2,3}$	$L_{2,4}$	$L_{2,5}$	$L_{2,6}$	$L_{2,7}$
	3	$L_{3,1}$	$L_{3,2}$	$L_{3,3}$	$L_{3,4}$	$L_{3,5}$	$L_{3,6}$	$L_{3,7}$
	4	$L_{4,1}$	$L_{4,2}$	$L_{4,3}$	$L_{4,4}$	$L_{4,5}$	$L_{4,6}$	$L_{4,7}$

FIGURE 1

NOTE: For convenience, we do not show lags greater than four years in the matrix. In practice, we group all losses with longer lags in the LAG 4 row.

With this configuration $L_{i,j}$ is the loss reported in year j with accident year lag i, (the accident year is j-i).

Notice that a complete accident year consists of a Northwest-Southeast diagonal in the matrix. For example:

Accident year 1

= (Acc. year 1, report year 1) + (Acc. year 1, report year 2) + ...

= (Report year 1, lag 0) + (Report year 2, lag 1)

+ (Report year 3, lag 2) + ...

= $L_{0,1} + L_{1,2} + L_{2,3} + L_{3,4} + \cdots$

We are now in a position to describe some of the coverage concepts in terms of the matrix above (in the examples which follow, all policies are assumed written at the beginning of year j for a one-year term):

"Mature" claims-made policy. A policy which covers claims reported during the policy period, regardless of accident date. Such a policy written at the beginning of year j will cover the jth column of matrix L in Figure 1.

First-year claims-made policy. A policy which covers only the "lag 0" row of the jth column of L. An insured in his first-year of a claims-made insurance program would purchase this coverage.

Second-year claims-made policy. A policy which covers the "lag 0" and "lag 1" row of the jth column of L. An insured in his second year of the claims-made insurance program would purchase this coverage.

Occurrence policy. A policy which covers claims arising from accidents occurring during the policy period. Such a policy would cover a Northwest-to-Southeast diagonal of matrix L. This is the traditional form of coverage in most liability lines.

Tail policy. A policy written for an insured who leaves the claims-made program. It covers losses whose accident date lies in the period during which the claims-made coverage was in force, and whose reported date is after the insured's last claims-made policy expired.

Retroactive date. The earliest accident date for which coverage is provided under a claims-made policy. Normally this would be the date on which an insured's first claims-made policy commences. Only claims with accident date subsequent to the retroactive date are covered by any subsequent claims-made or tail policy.

We will illustrate these coverages with the example of a hypothetical insured who begins practice at the start of year 1 and retires four years later. He buys an occurrence policy to cover his first year of practice, then switches to the claims-made program. He purchases first-year, second-year, and third-year claims-made policies for years 2, 3, 4 respectively. At the end of year 4, he retires and purchases a tail policy. The policies cover the Figure 1 loss matrix in the manner shown in Figure 2.

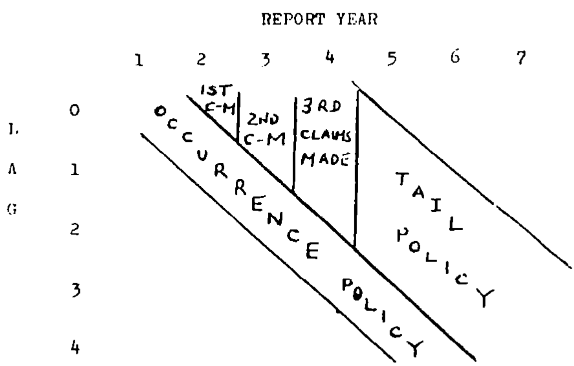


FIGURE 2

An important point to note is that the coverage above is equivalent to the coverage provided under 4 occurrence policies, as shown in Figure 3 below:

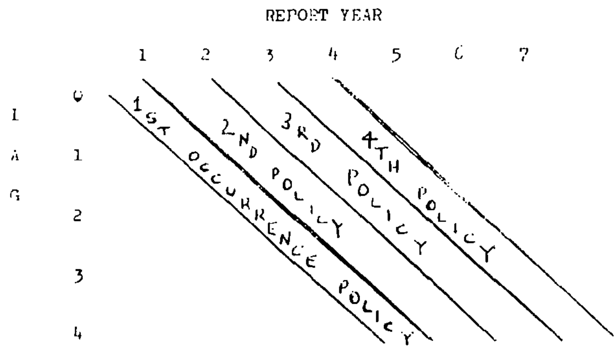


FIGURE 3

Although the coverage to the insured is the same under the claims-made system as under the traditional occurrence system, there is an important difference in the timing of the premium determination. To illustrate, the losses for report year 4 lag 2 are covered by the occurrence policy written at (and priced no later than) the beginning of year 2. For our claims-made insured, these losses would be covered by the third-year claims-made policy written at the beginning of year 4. The claims-made system allowed the insurer an extra two years to price this "lag 2" loss element.

IV. Claims-Made Ratemaking Principles

As noted previously, the major differences between the claims-made and the occurrence policy lies not in the coverage provided, but in the timing of pricing decisions affecting that coverage. Under claims-made we are always pricing next year's claims. Under occurrence pricing we must take into account claims to be reported many years in the future. The accuracy of any forecast is a direct function of how far beyond the data the projection is to be carried. A series of simple examples will illustrate this principle as it applies to claims-made and occurrence policies.

Let reported year J=0 represent the last year of history, J=1 represent the claims-made year we are pricing, and let (0,1), (1,2), (2,3), (3,4), (4,5) represent the components of the occurrence year we are simultaneously pricing.* In terms of the diagram we have:

	Future (Projections)				
History	1	2	3	4	5
$L_{0,0}$	$L_{0,1}$				
$L_{1,0}$	$L_{1,1}$	$L_{1,2}$			
$L_{2,0}$	$L_{2,1}$		$L_{2,3}$		
$L_{3,0}$	$L_{3,1}$			$L_{3,4}$	
$L_{4,0}$	$L_{4,1}$				$L_{4,5}$
	(C-M)				(Occ)

FIGURE 4

* Actually it might be more accurate to state that reported year J=-1 is the last year of history, since 1) there is typically a six month lag between the end of the experience period and the effective date of a filing, and 2) an additional six months between the effective date and the average date when the new price is in effect. Since this would have the same effect on pricing both occurrence and claims-made policies, it will be ignored for the sake of simplicity here.

Now let us assume $L_{0,0} = L_{1,0} = L_{2,0} = L_{3,0} = L_{4,0} = \$200$; that is, losses reported in the last year were produced in equal proportions from occurrences in the last five years. Also let's say we forecast that losses will increase at a rate of $20 per year for each lag. Then our diagram becomes:

	Future (Projections)				
History	1	2	3	4	5
$200	$220				
$200	$220	$240			
$200	$220		$260		
$200	$220			$280	
$200	$220				$300
$1,000	$1,100 (C-M)				$1,300 (Occ)

FIGURE 5

It is immediately apparent that next year's occurrence policy is more expensive than next year's claims-made policy, in this case by $1,300 - $1,100 = $200. The First Principle of Claims-Made Ratemaking states: <u>A claims-made policy should always cost less than an occurrence policy, as long as claim costs are increasing</u>. Furthermore, the greater the trend, the greater the difference will be. For example, suppose we underestimated inflation by $10 per year per lag. Then our diagram would become:

	Future (Assuming Change in Trend)				
History	1	2	3	4	5
$200	$230				
$200	$230	$260			
$200	$230		$290		
$200	$230			$320	
$200	$230				$350
$1,000	$1,150 (C-M)				$1,450 (Occ)

FIGURE 6

Now the difference is \$1,450 - \$1,150 = \$300. But consider what happened to the relative rate levels. The claims-made rate level proved to be inadequate by \$1,150 - \$1,100 = \$50 or 4.5%. The occurrence rate level turned out to be inadequate by \$1,450 - \$1,300 = \$150 or 11.5%.

The result is obvious when you think about it. But it is fundamental to understanding the difference between claims-made and occurrence ratemaking. In fact, it deserves restating as the Second Principle of Claims-Made Ratemaking: <u>Whenever there is a sudden, unpredictable change in the underlying trend, claims-made policies priced on the basis of the prior trend will be closer to the correct price than occurrence policies priced in the same way</u>. Stated another way, the confidence interval about the projected losses for a claims-made policy is narrower than for an occurrence policy priced at the same time.

In addition to a sudden unexpected change in the underlying trend there is another type of change that plagues actuaries pricing long-tailed lines: a sudden unexpected shift in the reporting pattern. Let us see how this would affect pricing accuracy under the two types of policies. First, recall our projections by referring back to Figure 5.

Now let's see what happens if we have a \$20/per year/per lag shift toward later reportings; that is, \$20 of what would normally be reported in lag 0 is not reported until lag 1, \$20 from lag 1 moves to lag 2, etc. (Note that only the first and last lags are affected since the others have the same dollars shifting in and out, and the same total dollars are reported.) Then our example looks like this:

History	Future (Assuming Change in Reporting Pattern) 1	2	3	4	5
$200	$200	$200	$200	$200	$200
$200	$220	$240			
$200	$220		$260		
$200	$220			$280	
$200	$240	$280	$320	$360	$400
$1,000	$1,100				$1,380

FIGURE 7

Under these circumstances, the mature claims-made policy is still priced correctly (as we would expect since the total dollars reported is unchanged), although a first year claims-made policy would have been slightly over-priced. But the occurrence policy is under-priced by $1,380 - $1,300 = $80, or 6.2%. The Third Principle of Claims-Made Ratemaking states: Whenever there is a sudden unexpected shift in the reporting pattern, the cost of mature claims-made coverage will be affected very little if at all relative to occurrence coverage.

If we put the two types of errors together, the result is even more dramatic.

History	Future (Assuming Change in Trend & Shift in Reporting Pattern) 1	2	3	4	5
$200	$210	$220	$230	$240	$250
$200	$230	$260			
$200	$230		$290		
$200	$230			$320	
$200	$250	____	____	____	$400
$1,000	$1,150				$1,530

FIGURE 8

The claims-made policy is under-priced by \$50 or 4.5% as before. But the occurrence policy is under-priced by \$1,530 - \$1,300 = \$230 or 17.7%. By now, it should be obvious that claims-made rates are both more accurate (because of a shorter forecast period) and more responsive to changing conditions (because external changes affect losses as they are reported). Two other points deserve emphasis. First, claims-made policies incur no liability for IBNR claims so the risk of reserve inadequacy is greatly reduced. (Principle Number Four). For example, a company writing occurrence policies for five years at the end of the period marked "history" in Figure 5 would carry an IBNR reserve of 4 x \$220 + 3 x \$240 + 2 x \$260 + 1 x \$280 = \$2,400. A company writing claims-made for the same period would have an IBNR reserve of \$0. The occurrence IBNR reserve needed under varying assumptions would be \$2,600 (Figures 6 and 7) or \$2,800 (Figure 8), so either of the two unfavorable developments would result in an IBNR reserve inadequacy of 8.33% for the occurrence policy. The IBNR needed for the claims-made policy is always 0.

The final point follows directly from the above. Because there is no need for IBNR, the time lapse between the collection of premiums and the payment of claims is greatly reduced. Consequently, the investment income earned from claims-made policies is substantially less than under occurrence policies. (Principle Number 5). The longer the reporting lag, or the shorter the settlement lag, the greater the difference will be.* The point is, as we reduce risk of inadequate

---

* Algebraically, the reduction may be expressed as $R/(R+S+1/2)$, where R is the mean reporting lag in years, S is the mean settlement lag and 1/2 represents the 1/2 year lag between payment of premium and the occurrence of a claim on average. Of course integrals rather than averages should really be used, but this approach produces a reasonably accurate answer given the uncertainties about R and S.

rates and insufficient reserves by switching to claims-made coverage, we pay for it with reduced investment income. On the other hand, the reduced risk should allow us to write more policies for a given amount of capacity, thus making up for the reduction in expected profitability per policy.

Summarizing the five Principles of Claims-Made Ratemaking discussed in this section:

1) A claims-made policy should always cost less than an occurrence policy, as long as claim costs are increasing.
2) Whenever there is a sudden, unpredictable change in the underlying trend, claims-made policies priced on the basis of the prior trend will be closer to the correct price than occurrence policies priced in the same way.
3) Whenever there is a sudden unexpected shift in the reporting pattern, the cost of mature claims-made coverage will be affected very little if at all relative to occurrence coverage.
4) Claims-made policies incur no liability for IBNR claims so the risk of reserve inadequacy is greatly reduced.
5) The investment income earned from claims-made policies is substantially less than under occurrence policies.

Now that the advantages of the claims-made approach are apparent, we will discuss how pure premium data for claims-made pricing is compiled, even where claims-made coverage has never been written.

V. Historical Pure Premium Collection

As explained above, our approach to ratemaking requires that we compute historical pure premiums by reported period and lag. To do this we collect the loss data and the exposure data and form the quotient.

Collection of Losses. It is easy enough to categorize losses by reported period and lag using the coded reported date and accident date. Since we use pure premiums on an "ultimate value" basis, development factors are applied to the most recent loss valuations. The development factors used in our approach have these features:

1. They are a function of report period only.
2. The development factors are applied only to the case reserve portion of the loss, not to the paid component.
3. The factors are determined through a "backward recursive" formula, described in Appendix A.

Because the factors develop reported period losses to ultimate value, they provide for anticipated shortages or redundancies in case reserves, but they do not provide for IBNR (Incurred But Not Reported) losses. There is no need for IBNR losses in the claims-made ratemaking process, since the primary focus is on losses by reported period.

Collection of Exposure. Determining the number of exposures for each reported period and lag is more difficult than tabulating the losses. This is especially the case when the data base consists of a mixture of occurrence and claims-made policies. The best way to see the difficulty is to look at hypothetical premium transactions and see how much each transaction contributes "earned exposure" to each report period by lag combination. Keep in mind that the goal in processing exposure data is to provide a suitable denominator for the pure premium calculations.

In the examples we will use the report period by lag matrix indexing system developed in Section III. We will call $E_{i,j}$ the exposure to loss reported in period j, with accident year lag i. (See Figure 9 below)

		REPORT YEAR (J)						
		1	2	3	4	5	6	7
	0	$E_{0,1}$	$E_{0,2}$	$E_{0,3}$	$E_{0,4}$	$E_{0,5}$	$E_{0,6}$	$E_{0,7}$
L	1	$E_{1,1}$	$E_{1,2}$	$E_{1,3}$	$E_{1,4}$	$E_{1,5}$	$E_{1,6}$	$E_{1,7}$
A	2	$E_{2,1}$	$E_{2,2}$	$E_{2,3}$	$E_{2,4}$	$E_{2,5}$	$E_{2,6}$	$E_{2,7}$
G (I)	3	$E_{3,1}$	$E_{3,2}$	$E_{3,3}$	$E_{3,4}$	$E_{3,5}$	$E_{3,6}$	$E_{3,7}$
	4	$E_{4,1}$	$E_{4,2}$	$E_{4,3}$	$E_{4,4}$	$E_{4,5}$	$E_{4,6}$	$E_{4,7}$

FIGURE 9

Example 1: A "mature" claims-made policy on one insured written at the beginning of year j. This contributes one exposure to all matrix elements with report year = j (i.e., the jth column of the matrix). This is because the policy covers all losses reported in year j, regardless of the lag.

Example 2: An occurrence policy written at the beginning of year j. This policy contributes one exposure to the following matrix elements:

$E_{i,j+i}$ for i = 0, 1, 2, ...

Example 3: A mature claims-made policy written 1/3 of the way through year j. This policy contributes:

2/3 exposure to $E_{i,j}$ for $i = 0, 1, 2, \ldots$

1/3 exposure to $E_{i,j+1}$ for $i = 0, 1, 2, \ldots$

This is the familiar "uniform earning" which also characterizes occurrence policies and most other policies in property and casualty insurance.

Example 4: A second-year claims-made policy written at the beginning of year j. This policy generates one exposure for only lag 0 and 1 portions of reported year j, (i.e., $E_{0,j}$ and $E_{1,j}$).

Example 5: A first-year claims-made policy written 1/3 of the way (i.e., May 1) through year 1. Before jumping to any conclusions about the amount of exposure look at Figure 10 on the next page. In Figure 10, we introduce the term "difference" as the difference between reported date and accident date. (This is in contrast to "lag", which is the difference between reported year and accident year.)

In Figure 10, the solid lines delineate regions represented by report year-lag combinations. These are parallelograms except for the "lag 0" region, which is a right triangle. The shaded area is the triangular region covered by the policy of Example 5. We can see that the policy covers the following

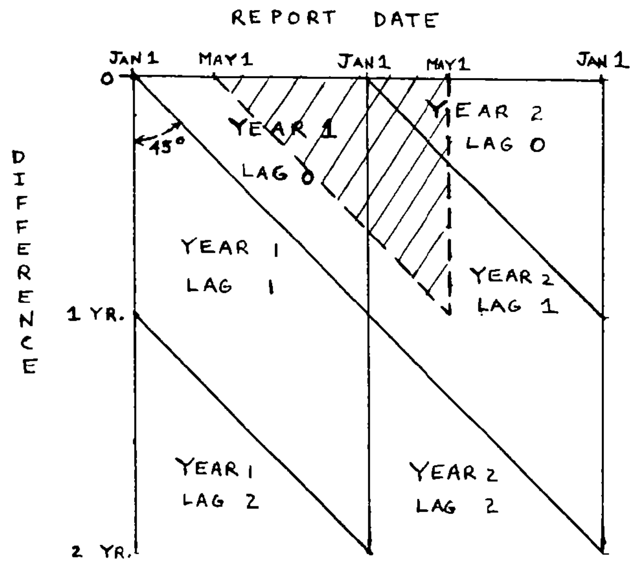


FIGURE 10

THIS FIGURE ILLUSTRATES THE COVERAGE OF A FIRST-YEAR CLAIMS-MADE POLICY WRITTEN ON MAY 1 OF YEAR 1.

SHADED AREA REPRESENTS THE COVERAGE OF THIS POLICY.

SOLID LINES REPRESENT BOUNDARIES OF "REPORTED YEAR - LAG" CELLS.

"DIFFERENCE" (VERTICAL AXIS) REPRESENTS THE TIME DIFFERENCE BETWEEN DATE OF ACCIDENT AND DATE OF REPORTING.

"LAG" IS REPORTED YEAR MINUS ACCIDENT YEAR.

proportion of these regions:

$2/3 \times 2/3 = 4/9$ of report year 1, lag 0;

$1/3 \times 1/3 = 1/9$ of report year 2, lag 0;

$1/3 \times 2/3 = 2/9$ of report year 2, lag 1.

These proportions are the earned exposure contributions to $E_{0,1}$, $E_{0,2}$, and $E_{1,2}$ respectively.

We can see, from Example 5, that the determination of exposure by report year and lag can be a fairly complex problem. This is especially so for "non-mature" claims-made policies and tail policies. However, the graphical technique used in Figure 10 is general and can be applied to any type of policy.

Before going on to a discussion of pure premium projection we will make some observations about the earned exposure calculations. We have concentrated on the general theory of how to make the exposure calculation, given the "maturity" of a claims-made policy transaction and the actual commencement date. In reality one may have to make these calculations using only summarized written premiums and earned premiums by type of policy and by time period (rather than using detailed transaction data). If this is this case, accuracy is greatly improved if the time periods are as fine as possible. Another problem which arises is the actual determination of the "maturity" of a claims-made policy. This requires the coding of the date on which an insured first purchases claims-made coverage (the "retroactive date"). This date is crucial and must be accurately recorded.

A simplification was made in the exposure calculation argument based on Figure 10. Using the proportional areas of the figure was equivalent to assuming a "uniform claim potential" within each year-lag parallelogram or triangle in the figure.

Summary. In this section we have attempted to describe the process by which historical pure premiums (quotient of loss and exposure) by reported period and lag can be computed. The tabulation of loss is straightforward, since insurance loss data contains reported date and accident date. The tabulation of exposure is much more complex, since different claims-made and occurrence policies contribute to different report period-lag exposure "cells".

Once these historical pure premiums are computed, the actuary can begin the projection of future pure premiums.

VI. Future Pure Premium Projection

Once historic pure premiums have been calculated, future pure premium projection proceeds in two steps. First, the future "mature" claims-made pure premium is determined. Second, the total pure premium is distributed back to lags and hence to policies at different levels of maturity. We will discuss each of these steps in turn.

In its simplest form, mature pure premium projection consists of nothing more than polynomial or exponential regression, using time as the independent variable.* This is suitable for countrywide data and perhaps for a few high volume states. It is not suitable for most states, however, as random fluctuation and the distortion of changes in legal or social climate can produce very poor fits and unreliable estimates. The actuary must be careful to check for this every time he does his analysis, even for the largest states, since a sudden surge or drop in claims being reported can occur in a single state at any time, destroying a previously stable trend. Normally, these events average out when countrywide data is used, although the evidence of recent years indicates that the experience in all states is becoming more highly correlated with one another.

If the actuary decides a particular state's trend is <u>not</u> sufficiently stable or reliable to use for projecting its future mature pure premiums, he may project them through a two-stage process. First,

---

* Other curves, such as log or power functions, have been proposed as alternatives. Unfortunately, the results derived from fitting these functions are highly dependent on the time index chosen, since the regression is done against the log of the index rather than the index itself.

the actuary generates countrywide fitted pure premium using polynomial regression as described above. Second, he applies linear regression or "regression through the origin"*, with state pure premium as the dependent variable and countrywide fitted pure premium as the independent variable. This approach assumes that, in the long run, a consistent relationship exists between state and countrywide pure premium. In other words, it assumes that the state will have the same percentage change from one year to the next that all states do, while using the individual state's own experience to determine its "relativity" to the countrywide rate. Linear regression is similar, but adds a constant term which allows partial recognition of the state's own apparent trend. Of course, if linear regression is used in both stages one and two, the result is the same as using linear regression against time directly.

Two points merit emphasis about the above procedure. First, as always, it is the actuary's task to strike the delicate balance between stability and responsiveness. This is done directly through his choice of a projection method, rather than indirectly through the choice of a credibility formula. The question to be kept constantly in mind is : How reliable is this data as an indicator of the claim process in this state? Fortunately, a wealth of information about the quality of the regression is available to help answer that question. Second, all projections are done on the experience itself. No "outside" frequency or severity trend information is superimposed on the data, thus avoiding the problem of explaining two or three sets of data and reconciling them with one another. There is no reason why this procedure should be limited to claims-made; its advantages apply equally well to any type of coverage.

---

* See Appendix B for a technical description.

Once the future mature pure premium has been determined, the problem of distributing it to lags may be approached in several ways. The original approach taken was to regress individual lag pure premiums ("a row" in the pure premium matrix) against time in the same way we regressed the total of all lags. As might be imagined from the above discussion, this method is highly sensitive, so much so that some lags will shoot upward at high rates while others are trending downward, in some cases even projecting <u>negative</u> values. Even if such trends were accurate reflections of what was going on in the real world, they would be undesirable for projecting pure premiums and rates since a smooth transition between rates for policies at succeeding maturities is very important in helping insureds understand the steps in claims-made coverage. A less-sensitive method was clearly needed. One simple approach we tried was to calculate the historic proportion for each lag, as follows:

$$(1)\quad b_i = \sum_j X_{i,j} \Big/ \sum_{i,j} X_{i,j},$$

where $X_{i,j}$ is the pure premium for report period j, accident period lag i.

The problem with this approach is that it does not recognize trends in relative pure premiums between lags. It was decided that a weighted proportion - with greater weight going to the larger, and presumably more recent, observations - would be a better representation. Surprisingly, it turned out that regression through the origin was the answer again. In this case, the historic proportion for lag i turns out to be:

$$(2)\quad b_i = \sum_j X_{i,j}\,\hat{X}_j \Big/ \sum_j (\hat{X}_j)^2,$$

where $\hat{X}_j$ is the fitted report period total pure premium.

Let's see how this compares to the historic proportion calculated above. Note that:

$$X_j \equiv \sum_i X_{i,j} \quad \text{and} \quad \sum_j \hat{X}_j = \sum_j X_j = \sum_{i,j} X_{i,j}.^*$$

Therefore, (1) can be re-written as follows:

$$b_i = \sum_j X_{i,j} \Big/ \sum_{i,j} X_{i,j}$$

$$b_i = \sum_j X_{i,j} \Big/ \sum_j X_j$$

$$b_i = \sum_j X_{i,j} \Big/ \sum_j \hat{X}_i$$

Thus, we see the difference between (1) and (2) is simply that the $\hat{X}_{j's}$ are used as weighting factors to place greater weight on larger pure premiums. It is important to note that $\sum_i b_i = 1$ since the $b_{i's}$ are the fractions of the total pure premium associated with each lag.

Summing up, the projection of pure premiums may be viewed as a two-step process. First, project the total ("mature") pure premium ignoring lags. Second, distribute the total pure premium back to lags. Several methods for carrying out each step have been suggested in this section.

There is no one "right" method for all circumstances. In fact, once the data is collected into a historic pure premium matrix, the possibilities for projection methods are limited only by the actuary's imagination and the flexibility of his statistical software package. For example, both econometrics and time series analysis merit exploration since the pure premium data by reported year seems to indicate distinct cycles about the long term trend line roughly corresponding

---

* This is true, if and only if, the $\hat{X}_{j's}$ were arrived at through linear regression. For regression through the origin, the residuals do not sum to zero.

to the economic cycle. This is logical since the incidence of malpractice should vary only with the utilization of medical services, while the reporting of a claim has a lot to do with how the claimant feels about his own economic situation. In any case, we suggest a "simulation" approach be used as a means of sensitivity analysis.

At the St. Paul, we divide states into four categories:

A - States, with highly stable patterns, where we use regression on their total pure premiums to determine their own trend;

B - States, where we use regression through the origin for both the total and lag pure premium;

C - States, where we use regression through the origin for the total pure premium but use the countrywide lag pattern and;

D - States, with very thin data, where we use a judgmental relativity to the countrywide pure premium.

As noted earlier, the categorization of each state must be reviewed each year to make sure changes in claim environment have not materially altered the data's reliability. Sensitivity analysis provides valuable insights in this process as well.

The ability to project pure premiums allows the actuary to determine more than prices for claims-made policies. Specifically, it can also be used to price occurrence policies, and to predict IBNR emergence and reserve adequacy. We will have more to say on this in Section VIII. But first, we will briefly discuss some special features of The St. Paul filings which distinguish them from those of other claims-made writers.

VII. Special Features of St. Paul Claims-Made Filings

Not all claims-made policies are alike in coverage. The St. Paul claims-made form contained several unique (at the time) coverage features which presented the actuaries with special pricing problems. Also, we chose several pricing techniques which were not traditional to facilitate the process of claims-made ratemaking from occurrence data. We will briefly discuss each of these special features in turn.

Several unique features of The St. Paul filings have already been discussed in Section III. One of these was the retroactive date; i.e., the earliest accident date for which coverage is provided under the claims-made policy. Previous claims-made or "discovery" policies treated all insureds the same, even if they had no prior exposure (e.g., just coming out of medical school) or if they were previously covered by an occurrence policy.

Another concept mentioned earlier was "tail coverage". When we introduced claims-made we felt it was the wave of the future. Someday all insureds would be using it and insureds would move from one carrier to another, carrying their retroactive dates with them. However, we recognized that this might not occur for many years and decided that we would have to offer our claims-made insureds guaranteed coverage for the "tail"; i.e., for claims which occurred while the insured was covered by claims-made but were not reported until after the last claims-made policy had expired. This was considered a rather dangerous step, since it in effect gave the insured the right to convert his coverage from claims-made to occurrence at any time. Were we leaving ourselves open to the same pricing problems we had had under occurrence? We argued that the risk could be greatly reduced by selling the tail coverage in three annual installments, or reporting endorsements, reserving the right to price them one at a time. The first reporting endorsement would be just like the insured's next claims-made policy, providing coverage for claims reported during that year only, except that accidents occurring in that year would be excluded. The second reporting endorsement would be similar, except that accidents occurring in the two year period after expiration of the last claims-made policy would not be covered. Only the third reporting endorsement would provide the kind of perpetual coverage that the occurrence policy did, with similar pricing hazards. It was argued that that hazard was acceptable since 1) we would be pricing it at least three years later than we would have priced the comparable coverage under an occurrence policy; 2) each insured would buy this "occurrence" coverage only once instead of every year, while the great majority of insureds were still buying claims-made; and 3) by the time we reach the third reporting endorsements the proportion of claims remaining to be reported is fairly small, so even a large percentage error in the rate would not result in a large dollar loss.

As we discussed the claims-made concept with our insureds, it became apparent that the three-pay reporting endorsement concept was acceptable to the majority but could work a real hardship for a few. So we added an additional option: We would sell a single-payment reporting endorsement to any insured terminating coverage due to death, disability or retirement.

The pricing of reporting endorsements - both three-pay and single-pay - poses no special problems. It merely requires trending the projected pure premiums further into the future. In fact, since the policy is essentially selling IBNR coverage, the pricing of reporting endorsements is equivalent to the determination of IBNR reserves, which will be discussed in Section VIII. Before proceeding to that discussion, however, we will briefly mention three special features of The St. Paul rate filings not directly linked to the coverage provided.

The St. Paul's claims-made policy is on an annual basis. But semi-annual reporting periods and lags were used in calculating and projecting pure premiums. The advantages of using this approach are twofold:

1. Less distortion calculating the earned exposures by "cell" (See Section IV), and
2. More data points for use in the regression.

Underlying the whole idea of pricing claims-made coverage from occurrence data is the implied assumption that the same body of claims would be reported at the same time under either policy. The more we thought about it, the less reasonable this assumption seemed. We decided that two changes were likely to occur at the transition point, both due to insureds understanding that coverage for a particular claim would not commence until the claim had been reported.

First, we assumed that, on average, claims would be reported sooner. Specifically, we assumed a two-month "shift" forward in claim reporting; algebraically,

$$L'_{0,1} = L_{0,1} + 1/6\ L_{1,2}$$

$$L'_{1,2} = L_{1,2} - 1/6\ L_{1,2} + 1/6\ L_{2,3} \text{ etc.}$$

Second, we assumed that there would be some additional reporting of incidents that would never have come in under the occurrence policy. Few, if any, of these incidents would result in loss payment, but they would require investigation and hence loss expense payment. Specifically, we assumed 5% additional claim dollars would be reported, and that all of this additional activity would come in at Lag 0. Algebraically,

$$L''_{0,1} = L'_{0,1} + .05 \times \quad L_{i,1}$$

We can never know what would have been reported had we continued with the occurrence policy, so it is impossible to test whether or not the "shift" and the additional incident reporting actually occurred. Now that several years of claims-made experience is contained in the data base, the need for this special adjustment no longer exists and it has been dropped from the filing.

The final special feature of the St. Paul filings involves the treatment of company expense. It was obvious that the pure premium and hence the rate for a first-year claims-made policy would be much less than for a mature policy. It did not follow that all expenses would be proportionately lower. In fact, most company expenses are probably fixed: i.e., they do not vary with the size of the premium. This was recognized by splitting the expense dollar into two parts: fixed and variable. This affected not only the relativities between different policy maturities, but between different classes of risk as well: the higher the rate, the lower the expense ratio. Algebraically, the rate calculation changed from:

$$R = PP/(1-E-P)$$

to $$R = (PP + FE)/(1-VE-P)$$

where R is the rate, PP is the pure premium, P is the profit allowance and E = FE + VE is the expense, broken down into its fixed and variable components. The following example will illustrate how this early instance of "expense flattening" works.

	Pure Premium (Relativity)	Fixed Expenses (% of Rate)	Variable Expenses and Profit (25% of Rate)	(Relativity)
Class 1 Physician, First-Year Claims-Made Policy	\$100 (1.00)	\$35 (19.4%)	\$45	\$180 (1.00)
Class 1 Physician, Mature Claims-Made Policy	\$500 (5.00)	\$35 (4.9%)	\$178	\$713 (3.96)
Class 7 Surgeon, First-Year Claims-Made Policy	\$800 (8.00)	\$35 (3.1%)	\$278	\$1,113 (6.19)
Class 7 Surgeon, Mature Claims-Made Policy	\$4,000 (40.00)	\$35 (0.56%)	\$1,345	\$5,380 (29.89)

FIGURE 11

VIII. Other Uses of Analytical Tools Developed

The techniques discussed in this article were developed specifically to price the claims-made coverages. However, once we develop a method to project pure premiums by future report year and lag, we have developed a tool which we can use to solve a variety of insurance problems. We can price occurrence coverages by adding up the appropriate elements from Figure 1 of Section III. For example, the projected pure premium for an occurrence policy commencing at the beginning of year 3 is straightforward:

$$X_{0,3} \quad + \quad X_{1,4} \quad + \quad X_{2,5} \quad + \quad X_{3,6} \quad + \quad \ldots,$$

where $X_{i,j}$ is the projected pure premium for reported year j, lag i.

Another area where the methods have application is in loss reserve determination. The "pure IBNR" (Incurred But Not Reported loss) for a company writing occurrence policies falls out of the projected loss calculation. For example, the IBNR reserve at the end of year 2 is the following area from the Figure 12 loss matrix below:

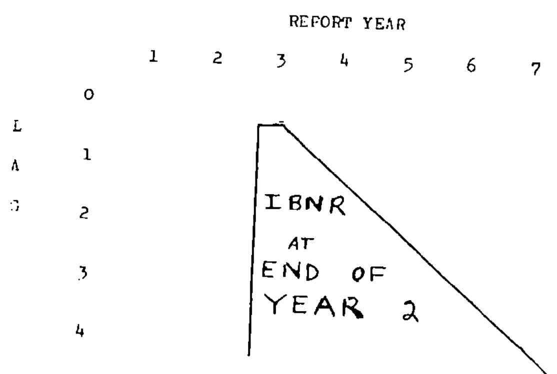


FIGURE 12

That is, the IBNR as of the end of year 2 on occurrence policies is the sum of all losses with reported year greater than 2 and accident year less than 3.

Determining "pure IBNR" is only half the problem of determining a total loss reserve. One must also project the additional development to be incurred on case reserves (reserves on losses already reported). This latter problem must be solved before we even begin to project the pure premiums in Section VI. It turns out that the case reserve development is easier to project once the loss data is collected into the report year by lag format of Figure 1. Appendix A discusses the precise method by which we project this case development.

Thus, the method of analyzing data which was developed to price claims-made policy gives us a convenient way of separating loss development into its two major components and projecting each separately:

Anticipated loss development = IBNR + Case Development.

Moreover, the method also projects an emergence pattern for the IBNR Loss.

IX. Summary

We began this paper by discussing the historical situation which led to the decision to write medical malpractice on a claims-made basis (Section II). Next we translated the problem of pricing the claims-made coverages into the problem of determining pure premiums by report period and accident period "lag" (Section III). Section IV presented a discussion of claims-made ratemaking principles.

Next followed a technical discussion of how to calculate historical pure premiums by report period and lag given insurance loss and exposure data (Section V). Once this is accomplished there are a variety of techniques available to project future pure premiums, and hence rates (Section VI). The St. Paul claims-made program and pricing techniques have several unique features (Section VII). Finally, the analytical tools used in claims-made ratemaking can also be applied to the general problem of IBNR determination for occurrence policies (Section VIII).

Appendix A

The Backward Recursive Reserve Development Method

In claims-made ratemaking the losses for each reported period must be developed to their ultimate value. We used a "Backward Recursive" reserve development method to accomplish this.

This method requires that loss data be available by reported period and "age". (Age 0 means the valuation as of the end of the reported period, age 1 is the valuation one period later, etc.) It also requires that the losses be separated into paid and case reserve components.

The "Backward Recursive" method calculates development factors which are applied to the reserve component of loss only. The determination of these factors proceeds in two steps:

1. "One-step" factors are calculated to develop losses as of each age to the next age. Two factors are calculated for each age k. P_k is the proportion of reserves of age k which will be paid by age k+1. R_k is the ratio of reserves at age k+1 to reserve at age k.
2. Ultimate factors are generated from the "one-step" factors. These factors apply to the reserves at age k to bring them to ultimate valuation.

The calculation of the "one-step" factors is a straightforward tabulation of the data. The factors are simply the following:

P_k = (Paid as of age k+1 - Paid as of age k)/(Age k reserve)

R_k = (Reserve as of age k+1)/(Reserve as of age k).

In order to generate the development factors to take reserves to their ultimate valuation, we need to assume an "ultimate" age, that is, an age N after which no further development occurs. The calculation proceeds in a "backward" fashion from the ultimate age in the following order:

$D_{N-1,N} = P_{N-1} + R_{N-1}$

$D_{N-2,N} = P_{N-2} + R_{N-2} \times D_{N-1,N}$

$D_{N-3,N} = P_{N-3} + R_{N-3} \times D_{N-2,N}$

*
*
*

$D_{0,N} \quad = P_0 + R_0 \times D_{1,N}$

Here $D_{n,N}$ is the development factor which brings reserves at age n to their valuation at age N. Each equation above merely says that the ultimate development on reserves of age n is the sum of payments during the time period n+1 and the ultimate value of reserves of age n+1. The first equation says that if N is assumed to be the ultimate age then reserves of age N-1 are either paid within one time period or remain outstanding at age N.

An example will illustrate the principles. Suppose that the following "single period" development factors have been determined and that age 3 is "ultimate".

Age (k)	0	1	2
P_k	.300	.500	.400
R_k	.800	.500	.500

Recall the meaning of these factors. For example, of all the reserves at age 0, 30% will be paid by age 1 and 80% will remain as reserve. The compound factors to apply to <u>reserves</u> are calculated in "backward" fashion:

$D_{2,3} = .400 + .500 = .900$

$D_{1,3} = .500 + .500 \times (.900) = .950$

$D_{0,3} = .300 + .800 \times (.950) = 1.060$

Appendix B

Regression Through the Origin

"Regression through the Origin" is a least-squares statistical technique similar to linear regression, except that the line of best fit is constrained to pass through the "Origin". Like linear regression this technique uses a line of "best fit" to fit a set of observations of some dependent variable to a set of observations of an independent variable. Unlike linear regression the line of best fit is constrained so that, when the value of the independent variable is zero, the fitted value of the dependent variable is also zero. The criterion for "best fit" is the same for both techniques: the line of best fit is chosen to minimize the sum of the squares of the differences between the observed and fitted values of the dependent variable.

There are two situations where Regression through the Origin might be substituted for linear regression. The first is the case where, a priori, the value of the dependent variable must be zero when the independent variable is zero. The second situation is one where linear regression has been run, but the "intercept" is not significantly different from zero, so that it can be dropped without hurting the accuracy of the model.

An example of a problem where Regression through the Origin might be used is the problem of projecting one company's output as a function of an industry's total output, given historical annual figures. If the second variable takes on a value of zero, the first must also.

The structure and mechanics of Regression through the Origin are similar to linear regression. The modeler has at hand observed values of a dependent variable ($Y_1, Y_2, \ldots Y_N$) and observations of an independent variable ($X_{1,2}, \ldots X_N$). The task is to calculate a parameter b such that

$$\hat{Y} = b\,X$$

gives the expected value of the variable Y given any observed value of the independent variable X. (Recall that in linear regression we look for parameters a,b to use in an expression $\hat{Y} = a + bX$.)

The parameter b is chosen so that the sum of squares $\sum_{i=1}^{N} (Y_i - bX_i)^2$ is minimized. The formula for b is given by

$$b = \frac{\sum X_i Y_i}{\sum X_i^2}$$

The statistic b has the property that

$$\sum (Y_i - kX_i)^2 = \sum (Y_i - bX_i)^2 + (b-k)^2 \sum X_i^2$$

for any constant k. We can see that the last expression is minimized when k = b, so that b is optimal in the "least squares" sense. For a fuller discussion of the statistical properties of the model, consult John Neter and William Wasserman's <u>Applied Linear Statistical Models</u> (Richard D. Irwin, Inc., 1974).

Although it appears that the Regression through the Origin model is a special case of linear regression, the reverse is actually true! This is because any linear regression model $\hat{Y} = a+bX$ can be rewritten

$$(\hat{Y} - \bar{Y}) = b\,(X - \bar{X}),$$

where $\bar{X}$ and $\bar{Y}$ are the sample means of X and Y respectively.

With this formulation we can see that any linear regression model is a special case of Regression through the Origin in which each variable has zero mean.

RATING CLAIMS-MADE INSURANCE POLICIES

by Joseph O. Marker and F. James Mohl

REVIEWED BY Michael F. McManus

The introduction of the claims-made policy as a vehicle for providing Medical Professional Liability insurance coverage in the mid-1970's clearly marked a turning point in the nature of the insurance market for this volatile line of business. Since the time when the St. Paul Fire and Marine converted their entire Medical Liability book of business from an occurrence form to a claims-made form in 1975, a significant portion of the market, especially the so-called "medical mutuals", has also shifted to providing coverage on this basis. At long last we have a comprehensive, actuarial perspective on the various advantages of this form of coverage and a technical discusssion of one approach that can be used to price it. Messrs. Marker and Mohl have filled a significant gap in the actuarial literature in this regard.

This review consists of a series of comments on the various sections of the paper and then a brief outline of another approach that has been used to price claims-made coverage.

Historic Perspective

Marker and Mohl present a brief but reasonably complete history of the developments that led to the St. Paul's decision to offer Medical Liability coverage only on a claims-made basis. Since no other major writer offered Medical Liability coverage on this basis at the time, this decision presented major challenges as respects the determination of proper rate levels, obtaining approval of the concept from regulatory officials and convincing policyholders that this change was in their long term best interest.

As noted by the authors, most companies responded to the "malpractice crisis" by either pulling out of the market or seeking large rate increases. St. Paul, however, decided to take the more innovative approach of limiting coverage to only those claims reported within the policy period, and thus eliminated the "pure" IBNR projection problem from ratemaking and reserving. There was considerable negative reaction at first from the medical community, largely because of their concern about the availability and price of "tail" coverage. The fact that many of the medical mutuals have subsequently adopted this form of coverage speaks well for their eventual understanding of the actuarial benefits of claims-made, as respects pricing and reserving.

The reader should be aware, however, that there were a few states in which St. Paul did not succeed in getting their claims-made program approved because of opposition from either the medical community or the state insurance department, and this resulted in their pulling out of these states. The fact that St. Paul's program was endorsed by the state medical societies in many states considerably aided their efforts in getting approval of the claims-made concept.

Coverage Concepts

In this section of their paper, Marker and Mohl present a clear explanation of the coverage terms applicable to claims-made by referring to a matrix of losses by report year and accident year lag. While the explanations are clear, it should be pointed out that several of the coverage "modifications" made by St. Paul were virtually mandated by the market situation at the time. For example, a "retroactive date" was absolutely necessary to avoid duplication of coverage and policy limits, since virtually all insureds previously had occurrence coverage. Of course, this situation allowed the first year claims-made rate to be about half of what an adequate occurrence rate would have been, and this temporary premium reduction helped sell the concept. Similarly, the availability of tail coverage had to be guaranteed in most states to obtain regulatory approval.

As pointed out by the authors, a major benefit of the report year/lag matrix approach is that occurrence data can be used to price claims-made policies; however, this approach is only feasible if report date has historically been captured accurately. Since this date was not required for experience reported to ISO, most companies did not begin capturing it until 1976, when ISO began requiring it on all Medical Liability experience. St. Paul was fortunate that it had been collecting this data element on its own experience.

The interrelationships between occurrence coverage, the various types of claims-made policies (first year, second year, etc.) and the role of tail coverage are well demonstrated by Figures 2 and 3, as well as the timing advantage in pricing claims-made.

Ratemaking Principles

The authors make good use of several simplified examples in this section to develop various "Principles of Claims-Made Ratemaking."

The manner in which data in Figure 5 is manipulated implies that pure premiums, and not losses, are being used. Thus the First Principle should in fact state that a claims-made policy would cost less than an occurrence policy, as long as pure premiums are increasing.

The remaining Principles demonstrate the intuitive benefits of pricing and reserving for mature claims-made policies versus occurrence policies in an unstable claim climate which has been and probably will continue to be the essence of Medical Liability. One minor complaint is that, in discussing reserving, it would have been helpful to indicate that the "pure" definition of IBNR, i.e., excluding case development, was being used.

Historical Pure Premium Collection

The approach outlined by Marker and Mohl in this section to develop reported losses to ultimate is notable for two reasons: 1) only the case reserve portion of the incurred amount is developed, and 2) a "backward recursive" formula is used to compute development factors.

The first feature is well suited to the development of report year losses because of the lack of IBNR claims. Thus in developing losses to ultimate, any significant shifts in settlement rates are recognized since paid losses are not developed. In statewide ratemaking, differences in settlement rates by state can also be recognized. A second favorable aspect is that, assuming that development factors are calculated on a countrywide basis (which is not indicated), loss development is solely a function of reserve adequacy, and therefore the use of countrywide factors should be much less distortive than for occurrence experience where different claim reporting patterns by state make countrywide factors much less appropriate.

The backward recursive approach is a logical one, given the desire to develop case reserves only. Since the procedure starts by assuming an "ultimate" age, N, after which no further development occurs, the stability of developments between ages N -1 and N would appear to be fairly critical. Unfortunately, no comments are offered by the authors on this point or in general on what procedure is used to develop "one-step" factors--simple averaging, weighted average, etc.

The authors' treatment of the determination of appropriate exposures for the report year/lag matrix is interesting in that what is a straightforward process for standard occurrence ratemaking becomes a very tricky calculation for claims-made. While the graphical solution to this problem illustrated in Figure 10 is clear and understandable, one wonders about the difficulty of describing it in a computer program so that it can be automated and the quality of the results controlled.

One other comment in this regard concerns the simplifying assumption that all exposures have a "uniform claim potential" within each report year/lag combination. While this would generally not be a problem, it clearly is not very accurate for first year (and even second year) policies and also for tail policies.

An interesting result of this treatment of exposures is that the number of exposure counts generated by a claims-made policy is equal to the number of report year lags covered by the policy, e.g., a mature policy generates five exposures. As the authors point out, accurate coding of the retroactive date is critical in determing the maturity and therefore the number of exposures of a claims-made policy.

Future Pure Premium Projection

In this section, Marker and Mohl describe their approach for first projecting future mature claims-made pure premiums from historic pure premiums, and then distributing the total pure premium back to each lag period.

The procedure used to accomplish the first goal is noteworthy for potential application to other lines of business. The usual question of determining the credibility of statewide experience instead becomes a question of what method should be used to project statewide pure premiums into the future. If the state's actual experience is not sufficiently stable, then statewide pure premiums (dependent variable) are regressed against countrywide fitted pure premiums (independent variable) using either linear regression or regression through the origin. With this approach, various measures of the fit of the regression are available to evaluate the "credibility" of the statewide experience.

The development of the procedure used by the authors to distribute these mature pure premiums back to each lag period is indicative of the evolutionary process that ratemaking frequently goes through. The original approach of simply regressing each row of the pure premium matrix against time proved to be highly sensitive to random fluctuations. On the other hand, simply averaging the historic proportion for each lag was totally insensitive to real trends in relative pure premiums between lags. The approach finally selected was a weighted average, with the fitted report year total pure premiums used as weights. Because of a normally positive pure premium trend, this resulted in greater weight effectively being applied to more recent observations, which was one of the authors' goals.

Special Feature of St. Paul Filings

This section of the paper describes the approach used to price several unique features of the St. Paul claims-made filings, including tail coverage and expense flattening.

Because tail coverage had to be made available to all insureds, St. Paul was concerned because this gave everyone the theoretical option of converting to occurrence coverage at any time. The pricing uncertainty for this coverage was considerably minimized by the decision to sell such coverage in three annual installments, so that an occurrence type projection would only have to be made for the third installment, which would only cover a small percentage of the claims covered by the tail policy. This was a creative solution to a pricing "problem".

In this section the authors also describe two adjustments that were made to historic occurrence data, because of their concern that prior claim reporting patterns would be impacted by 1) an accelleration in the reporting of claims, and 2) an increase in the frequency rate, because of the insured's concern about having claims-made coverage. While the adjustments made were reasonable judgements and were only required during the transition from occurrence to claims-made, it would have been interesting to at least see a comparison of the claim reporting lag before and after claims-made, recognizing that the actual effect on the frequency rate cannot be isolated by itself.

The approach used to flatten expenses is similar to that currently being used in other filings. It is interesting to note that St. Paul decided to flatten expenses by class, as well as by year of claims-made coverage. Prior to this time, very few occurrence carriers had flattened expenses by class, basically because of their concern about its impact on their distribution of insureds by class, when other carriers had not yet made such a change.

Other Uses of Analytical Tools

In this final section, Marker and Mohl validly point out that the pricing approach they have developed is not limited in its application to claims-made ratemaking. Their backward recursive approach to quantifying case development can be generally adopted to developing report year losses to ultimate.

In utilizing the report year/lag matrix, however, the reader should be aware of the possible instability that may result from refining experience in this matter. Attention should also be paid to the accuracy of report date coding on non-claims-made coverage.

An Alternative Approach to Claims-Made Ratemaking

When St. Paul announced their intention in 1975 to convert their book of Medical Liability business from occurrence to claims-made coverage, there was considerable concern on the part of other leading writers as to the effect this would have on the nature of the market. Presumably this could have been (and in fact turned out to be) the start of a general change in the way in which Medical Liability coverage was provided.

In order to be prepared to possibly compete with St. Paul, several major writers asked Insurance Services Office (ISO) to develop a claims-made program (forms and rates) that could be adopted by a member company who wished to offer claims-made coverage. Accordingly a special ISO committee of actuaries and underwriters was formed to develop such a program.

Since claim report date had never been required in ISO statistical reports, an immediate problem faced by this committee was the lack of suitable industrywide compilations of experience that could be used to price claims-made coverage. Fortunately, several individual companies were able to provide their own data in reasonably appropriate formats, but a sophisticated analysis was not possible given the data and time constraints in effect.

The approach that was developed by this committee was to come up with a series of countrywide claims-made "multipliers" that would be applied to an adequate occurrence rate by state to develop claims-made rates by year of coverage. Thus, after making adjustments for expected changes in claim reporting patterns similar to those described by Marker and Mohl, it was decided that the rate for first year claims-made coverage should be 50% of an adequate occurrence rate. Increasing percentages were selected for the other years of claims-made coverage, reaching 90% for the fifth year; this was in recognition of the fact that a mature claims-made rate should always be less than an adequate occurrence rate. While it was expected that claim reporting patterns would be somewhat different by state, adequate data by state was not available to examine this assumption, and so, countrywide factors were adopted.

This special committee was also asked to recommend the necessary statistical plan changes so that actual claims-made experience written by St. Paul and other companies could be identified and appropriately adjusted for industrywide ratemaking purposes. Thus, claim report date and "Date of Entry into Claims-Made Coverage" (retroactive date) were added to ISO statistical requirements.

While the program developed by this committee was made available to interested carriers, they did not really address how actual claims-made data would be used in ratemaking. This issue was later addressed by ISO's General Liability Actuarial Subcommittee, which had responsibility for Professional Liability ratemaking procedures at the time, and subsequently by the Professional Liability Actuarial Subcommittee.

The GLAS first addressed the general question of whether claims-made and occurrence data should be combined for ratemaking or reviewed separately. Since both types of coverage were being provided to a significant portion of the market and given the credibility problems inherent in making rates by state for Medical Liability sublines, the GLAS felt there was no real choice on this point: claims-made and occurrence data should be combined for ratemaking.

Because many years of experience are needed in Professional Liability ratemaking and since the available occurrence experience was summarized on a policy year basis, the GLAS next decided that, for purposes of statewide ratemaking, claims-made exposures would be extended by current claims-made rates (appropriate claims-made multiplier times current occurrence rate) and occurrence exposures, of course, by current occurrence rates. Thus the actual review of claims-made multipliers would be separately addressed, just as classification relativities are separately reviewed.

In the area of loss development, historic policy year occurrence loss development factors were obviously inappropriate to apply to claims-made losses. Since St. Paul's experience showed that report year case development was minimal, it was temporarily decided to assume that claims-made losses would not be subject to any further development, since they were largely comprised of St. Paul experience. Occurrence losses were developed using historic policy year occurrence loss development factors.

With these adjustments it was felt that the standard ISO ratemaking procedure could then be followed to review the adequacy of occurrence rate levels. When sufficient data was available, the adequacy of claims-made multipliers could be addressed by reviewing accident year/report year tabulations that are being compiled. It was subsequently determined, however, that one further adjustment had to be made in the area of trend. The ISO procedure at the time was an exponential projection of countrywide policy year average incurred claim costs and claim frequencies. When claims-made experience entered these calculations, however, a distortion in the frequency calculation resulted because, for claims-made years, ultimate claim counts (occurrence plus claims made) were being compared to the unadjusted number of insureds.

The eventual solution to this problem was to revise the trend procedure so that policy year ultimate incurred loss ratios are exponentially projected, instead of severity and frequency separately. Since the premium at present rates in the denominator of this calculation reflected the extension of claims-made exposures by claims-made rates, the distortion noted above was eliminated.

While the ISO procedure is clearly not as sophisticated as the St. Paul approach described by Marker and Mohl, it is a reasonably sound technique, given the industrywide data available at the time. It was described here merely to indicate that the St. Paul approach is not the only way to price claims-made coverage.

Conclusion

In closing this review, there were a few areas not addressed by the authors that would have been of interest to many actuaries. These include:

1. An analysis of increased limits requirements by year of claims-made coverage. It would seem that the severity of claims reported in the first year of claims-made coverage would be less than that of claims covered by a mature policy.

2. A discussion of any differences in claim reporting patterns by class. The original St. Paul filings recognized faster claim reporting for at least two classes: anesthesiologists and neurosurgeons. It

would be interesting to see if their expectations have been met and if any other classes have shown significant differences from the average.

3. A summary of the adjustments that have been made to the Experience Rating Plan for Hospital Professional Liability in order to allow the inclusion of claims-made experience, including its impact on loss development factors, credibility values, D-ratios, etc.

4. A discussion of changes required to accounting procedures to accomodate claims-made coverage, especially the manner in which premium is earned.

These minor omissions do not take anything away from the value of the authors' paper, which represents a significant contribution in an area that has not been adequately addressed in the past.

TITLE: AN ANALYSIS OF RETROSPECTIVE RATING

AUTHOR: Mr. Glenn Meyers

Mr. Meyers is Senior Actuarial Analyst with CNA of Chicago. He received a PhD in Mathematics from the State University of New York in 1972 and became an ACAS in 1979. Glenn has authored articles appearing in Studia Mathematica and Proceedings of the American Mathematical Society.

REVIEWER: Mr. James F. Golz

Mr. Golz is Associate Actuary and Assistant Manager with Employers Insurance of Wausau. Jim received his FCAS in 1974 and is a member of the American Academy of Actuaries since 1976. He has served on the CAS Editorial Committee and Education Committee.

I. INTRODUCTION

The purpose of this paper is to address the following question. Should the present retrospective rating formula be modified to account for the claim severity of the risk being insured, and for the loss limit chosen for the plan? It will be shown that there are significant differences in premium adequacy that can attributed to the above mentioned factors. Alternatives to the present formula will be proposed.

The Present Retrospective Rating Formula

The premium for an insured written under a retrospective rating plan is given by the following formula. This formula is generally used in Workers' Compensation insurance.

$$R = [\ (P \times b) + (P \times c \times e) + (c \times A)\ ] \times t$$

subject to a minimum of $h \times P$ and a maximum of $g \times P$.

Where:

- R = Retrospective Premium,
- P = Standard Premium,
- b = Basic Premium Factor,
- c = Loss Conversion Factor,
- e = Excess Loss Premium Factor,
- A = Actual Limited Losses,
- t = Tax Multiplier,
- h = Minimum Premium Factor and
- g = Maximum Premium Factor.

In some plans, losses arising out of a single accident are limited to a specified amount before entering the retrospective premium calculation. The excess loss premium factor provides for the cost of this loss limit.

The basic premium factor can be written as follows:

$$b = a + (c \times i).$$

The factor a provides for acquisition expenses, general underwriting expenses and profit. The factor i is called the insurance charge. This factor provides for the net cost of limiting the retrospective premium between the minimum and maximum premiums.

The standard formula for calculating the insurance charge does not take into account the claim severity distribution of the individual insured, nor does it take into account the loss limit selected for the plan.[1] In other words, the insurance charge, as calculated by the standard formula, will be the same no matter what claim severity distribution applies to the insured, or what loss limit is used.

The loss experience will be more volatile for a high severity, low frequency insured than for a low severity, high frequency insured. Since a high severity, low frequency insured will "break the maximum" more often, he should have a higher insurance charge than an otherwise comparable low severity, high frequency insured.

---

1. National Council of Compensation Insurance, Retrospective Rating Plan D

The insurance charge includes a provision for the portion of the losses which exceed any potential loss limit. But, in a plan which has a loss limit, these losses are provided for by the excess loss premium factor. Thus, a plan with a loss limit should have a lower insurance charge than a plan with no loss limit.

It has long been recognized that these factors can significantly affect the adequacy of the retrospective premium. Perhaps the main reason the rating formula has not been modified is that it would involve making an already complex rating formula even more complex. According to one account, it could require 200,000 pages of tables to properly calculate the insurance charge.[2]

Another problem is inherent in the way data has been gathered under the present formula. The distribution of loss ratios is tabulated by direct observation. This allows one observation per insured each year. If one were to create categories of insureds and tabulate the experience for each of the categories, he might well find that the experience is not credible.

---

2. An excellent discussion of these issues can be found in "The California Table L", PCAS LXI, by David Skurnick, and the ensuing discussions by Frank Harwayne and Richard H. Snader.

The general approach taken by this paper will be to build a mathematical model of the loss process. This model will be used to generate annual losses for different kinds of insureds. We will then quantify differences in premium adequacy that can be attributed to the factors mentioned above. Next we will explore modifications to the current formula which can more adequately price a retrospective rating plan.

II. THE MODEL

The Generalized Poisson Distribution

The Generalized Poisson distribution will be used to model the loss process.[3] This model is based on the following assumptions.

1. The number of claims has a Poisson distribution.
2. Claim severity is independent of claim frequency.

Three claim severity distributions have been selected. These distributions will represent a standard insured, a high severity insured and a low severity insured. The distributions are given in Exhibit I. These distributions are hypothetical ones selected by the author.

The following information is needed to generate a distribution of annual losses: (1) the expected losses; (2) the claim severity distribution; and (3) the loss limit. Sample values for the distribution are calculated by the following steps.

---

3. R. E. Beard, T. Pentikainen and E. Pesonen, Risk Theory, Chapman and Hall Ltd. (1977), Ch.3.

1. Calculate the average claim size from the claim severity distribution.

2. Calculate the parameter, λ , for the Poisson distribution.

$$\lambda = \frac{\text{Expected Losses}}{\text{Average Claim Size}}$$

3. For each sample do the following.

   3.1 Randomly select the number of claims, n, from the Poisson distribution.

   3.2 Do the following n times.

      3.2.1 Randomly select a claim amount from the claim severity distribution.

      3.2.2 Adjust the claim amount for the loss limit.

   3.3 The sample loss amount is the sum of all claim amounts generated by step 3.2.

The annual loss distributions used in this paper are "empirical" ones consisting of 10,000 samples.

The use of the Poisson distribution for the number of claims deserves some comment. The author chose this distribution because of its widespread use in the actuarial literature. The author has no evidence that the Poisson distribution is the most appropriate. However, if some other distribution is chosen, one should expect only a slight increase in the variance of the annual loss distribution.[4] Thus the results of this paper should hold even if this assumption is changed.

The major results of this paper will be based on the difference between insureds represented by the claim severity distributions in Exhibit I. No attempt has been made to fit this model to live data.

However, using Exhibits IIa and III, one can compare the results of this model with the present retrospective rating formula. Exhibit IIa provides the excess loss premium factors derived from the claim severity distributions in Exhibit I. Exhibit III gives the insurance charges calculated using the standard formula, and by a method (to be described below) using the claim severity distribution for the standard insured.

Adequacy of the Retrospective Premium

When given the parameters of the retrospective rating plan and the 10,000 loss samples generated by the model, it is possible to calculate the average retrospective premium generated by the plan. Similarly, one can calculate the average premium that would be generated by a "cost-plus" rating plan (i.e. a retrospective rating plan with no minimum or maximum premium). The premium for a "cost-plus" rating plan is given by the following formula:

$$CP = [\ (P \times a) + (P \times c \times e') + (c \times A)\ ] \times t,$$

where e' is the "correct" excess loss premium factor as derived from the claim severity distribution.

The retrospective premium adequacy of a plan (RPA) can be defined as follows:

$$RPA = \frac{\text{Average "Cost-Plus" Premium}}{\text{Average Retrospective Premium}}$$

The retrospective premium adequacy of plan is a measure of its profitability. If the retrospective premium adequacy is less than 1.00, the insurer should expect to make more than the budgeted profit. *Conversely, if the retrospective premium adequacy is greater than 1.00,* the insurer should expect to make less than the budgeted profit.

If all the parameters of a retrospective rating plan are given except the insurance charge, the retrospective premium adequacy can be thought of as a function of the insurance charge. To use the model to find the insurance charge one solves the following equation.

$$RPA(i) = 1$$

This equation can be solved by standard numerical methods.[5] It should be pointed out that solving this equation by hand would be extremely difficult due to the large number of terms involved. However, solving this equation by computer has proved to be very speedy and reliable. It should also be pointed out that this method of finding the insurance charge can easily be adapted to other kinds of retrospective rating formulas.

---

5. The author used the Modified Regula Falsi method, which is described in Elementary Numerical Analysis: An Algorithmic Approach, McGraw Hill Inc. (1972), by S.D. Conte and Carl de Boor.

4. R.S. Miccolis, "On the Theory of Increased Limits and Excess of Loss Pricing", PCAS LXIV, p 43.

III. AN ANALYSIS OF THE CURRENT FORMULA

Like it or not, we already have a formula for retrospective rating in use. With some minor exceptions, this formula is used on a countrywide basis for Workers' Compensation.

Since the price of a retrospective rating plan is fixed, the problem becomes one of risk selection. This section seeks to identify those insureds which can profitably be written under a retrospective rating plan.

Another particularly troublesome problem with the current formula is that many people feel that the excess loss premium factors currently in use are inadequate. This section will show how to quantify the effect of such an inadequacy.

A Model of the Current Procedure

Ideally, the current retrospective rating formula can be described in the following manner. A single loss distribution is chosen to represent all insureds with a given expected loss amount. The insurance charge is calculated from this loss distribution on the assumption that no loss limit will be used. This insurance charge is used whether or not a loss limit is actually used in the plan.

The current formula will thus be modeled in the following manner. The standard claim severity distribution will be used to calculate insurance charges. They are given in the last column of Exhibit III. These insurance charges will be used to evaluate the retrospective premium adequacy of a plan no matter what claim severity distribution represents the insured, and no matter what loss limit is selected.

Exhibit V shows the retrospective premium adequacy for the high and low severity insureds when there is no loss limit. As can be seen from this exhibit, there are substantial differences in the retrospective premium adequacy that can be attributed to differences in claim severity. Clearly it is not desirable for the insurer to write a high-severity insured on such a retrospective rating plan.

Exhibit VI shows the retrospective premium adequacy for plans which have a loss limit. As can be seen from the exhibit, the overlap between the excess loss premium factor and the insurance charge results in a very favorable retrospective premium adequacy from the viewpoint of the insurer. This is true even for the high severity insureds which fared poorly when there were no loss limits.

The Effect of Inadequate Excess Loss Premium Factors

After examining Exhibit VI, one might conclude that an insurer should require loss limits on all retrospective rating plans. However, there are some problems with this strategy. In talking with various actuaries and underwriters who work in Workers' Compensation, the author has found a strong consensus that the excess loss premium factors currently in use are inadequate. To get some idea of the effect of inadequate excess loss premium factors, the author calculated the retrospective premium adequacy of plans with the excess loss premium factors cut in half. The results are shown in Exhibit VII.

The results of these calculations show that, in some cases, it still may be more profitable to write an insured with a loss limit. The profitability of a plan depends upon the balance between the amount of inadequacy in the excess loss premium factors and the redundancy in the insurance charge. This balance is more favorable to the insurer in plans with a low maximum premium. It should also be noted that this balance works against the insurer for the larger premium sizes.

If an underwriter is concerned about inadequate excess loss premium factors, he should encourage the insured to take a plan with a high maximum premium and no loss limit, or a plan with a low maximum premium and a loss limit. The author has discussed this underwriting strategy with both underwriting and marketing personel. They both thought that neither of these programs are marketable. It should be clear why a plan with a high maximum would not sell. The marketability of the low maximum plan with a loss limit deserves some comment.

When deciding whether or not to purchase a plan with a loss limit, the insured will look at his past experience and see what he would have paid under each plan. Exhibits VIII and IX provide such a price comparison based on the 10,000 samples generated by the loss model. These exhibits show calculations of retrospective premium at various percentiles. It should be noted that the insured in this example is paying $25,062 in excess premium in the plan with a $30,000 loss limit. In examining these exhibits one can see that the insured would be paying a greater than or equal premium for the plan with a loss limit at every percentile. The only time there is equality is when both plans pay the maximum premium.

The underwriter went on to say that he would be extremely suspicious of any insured that would be willing to accept a plan with a loss limit. Such a plan would be acceptable to an insured who has experienced a severe loss and is afraid of another one.

The possibility of adverse selection in plans with a loss limit is something that could be tested. What is required is a comparison between claim severity distributions for insureds who have, and who have not purchased a plan with a loss limit. The author has not seen such a comparison.

Adverse selection could provide an explanation for inadequate excess loss premium factors.

IV. OTHER RETROSPECTIVE RATING FORMULAS

Insurance Charges Which Reflect Claim Severity and Loss Limits

Given the differences in the retrospective premium adequacy of the various plans mentioned above, it is natural to ask what should the insurance charge be in order to accurately reflect differences due to claim severity and loss limits. Exhibits X and XI provide the proper insurance charges.

The taking into account of differences due to claim severity presents the problem of rating different exposures which are under the same retrospective plan. To do this, one can simply sum the losses incurred by each separate exposure and then proceed as usual. Exhibit XVa provides calculations of insurance charges for an insured with standard premiums of $150,000 in a class represented by the high severity distribution and $50,000 in each of two classes represented by the low severity distribution and the standard distribution. This method can easily be generalized to cases where the expense factors and loss limits are different for each class.

While this method of calculating the insurance charge does not require an excessive number of tables; it does require a great deal of computer time. The overwhelming majority of the computer time is consumed by generating the distribution of annual losses. The author is aware of quicker ways to generate losses, which deserve serious consideration.[6]

---

6. R.E. Beard, T. Pentikainen and E. Pesonen, <u>op. cit.</u>, Ch.7.

Retrospective Rating Plans Which Require a Loss Limit

In his observations of Exhibit XI, the reader may have already noticed that the insurance charges for plans with the same standard premium and loss limit are nearly equal.[7] The difference in the price for insureds with different claim severity distributions can be attributed almost entirely to the excess loss premium factor. This is true because we are substituting a fixed excess premium for the most volatile part of the actual losses.

This observation suggests that, when using a fixed loss limit, one can devise a retrospective rating formula for which the differences in the insurance charges due to claim severity can be kept to an acceptable minimum. This plan would simply use the insurance charge calculated for the standard insured, as the insurance charge for all insureds. Each insured would still use the appropriate excess loss premium factor. The retrospective premium adequacies for various insureds under such a plan are given in Exhbits XII and XVb.

The author would also propose that the insured not be given a choice of loss limits. This would minimize the number of tables needed to calculate the insurance charge. The loss limit would be determined by the total expected losses of the insured. Furthermore, if it is determined that adverse selection is a cause of inadequate excess loss premium factors, it may be necessary to require that all insureds have the same loss limit.

---

7. The reader should note the different definitions of the insurance charge that are in the literature. Skurnick's insurance charge provides for both the excess losses on individual claims and the effect of limiting the retrospective premium. Harwayne suggests reducing the excess loss premium factor to account for the overlap.

If we are to require that a specific loss limit be used for a given insured, we should try to choose a loss limit that will be acceptable to a majority of the insureds. It may be desirable to calculate excess losses by the following formula.

Let L be the total loss arising out of a single accident.

If $L \leq A$ $\quad \begin{cases} \text{Primary Loss} = L \\ \text{Excess Loss} = 0 \end{cases}$

If $L > A$ $\quad \begin{cases} \text{Primary Loss} = \dfrac{L \times B}{L+B-A} \\ \text{Excess Loss} = L - \text{Primary Loss} \end{cases}$

In this case we say the loss limit is (A:B).

One can see that primary portion of the loss will be between A and B when the loss is greater than A. This formula is similar to the one used in multi-split experience rating for Workers' Compensation.

Exhibits XIII and XIV show calculations of the insurance charge and the retrospective premium adequacy for plans with a dual loss limit. It should be noted that a more restrictive loss limit allows less variance in the retrospective premium adequacy. The selection of a required loss limit will depend upon what will be acceptable to a majority of insureds and upon how much variance in the retrospective premium adequacy the insurer is willing to tolerate.

IV. CONCLUSION

This paper discusses three options which can be taken with regard to the retrospective rating formula.

The first option is to leave the present formula unchanged. If this option is elected, a retrospective rating plan will produce premium deficiencies for high severity insured, while it may produce premium redundancies for plans which have a loss limit. Such plans are not appropriate for high severity insureds.

The second option is to replace the present formula with one that properly accounts for claim severity and loss limits. This option would allow complete freedom in choosing the kind of plan to be used. The main drawback to this option is the large amount of computer time needed to calculate the insurance charge. It will be necessary to develop a more efficient loss generation program before this option can be implemented.

The third option is to restrict the number of plans available to the insured. This provides an immediate reduction in the number of tables needed. If we require that all retrospective rating plans have a loss limit, it turns out that the claim severity of an insured has only a slight effect on the insurance charge. Because of this it should not be necessary to have separate tables for each claim severity group in order to calculate the insurance charge. If a single loss limit is required, the resulting procedure should be no more complex than the present one. A single loss distribution and loss limit could be chosen to represent all insureds with a given expected loss amount.

This paper attempts to quantify the effect of each of these options.

The author prefers a flexible formula like that mentioned in option two. Should this approach prove unworkable at the present time, the author would then choose option three. The present retrospective rating formula discards accuracy in order to maintain flexibility. The proposed formula discards flexibility in order to maintain accuracy.

This paper bases its conclusions on computer simulation using hypothetical data. These techniques permitted a vast amount of experimentation with various retrospective rating plans. These conclusions are the results of this experimentation. Any concrete proposal for changing the current procedure must look at real data. The modification of the current procedure will be a very expensive and time consuming undertaking. It is hoped that this paper will convince the reader that such an undertaking is worth the effort.

The ideas expressed in this paper are the result of conversations the author had with many people at his company. The author would like to thank these people for their contributions.

Exhibit I Claim Severity Distributions

Claim Amount	Probability that a claim will be less than Column 1		
(1)	(2)	(3)	(4)
50	0.4310	0.3692	0.2464
100	0.5781	0.5147	0.4385
250	0.8561	0.8419	0.6195
500	0.8994	0.8835	0.8474
750	0.9175	0.9040	0.8684
1,000	0.9291	0.9155	0.8862
1,500	0.9455	0.9310	0.9050
2,500	0.9628	0.9495	0.9225
3,500	0.9718	0.9606	0.9348
5,000	0.9788	0.9704	0.9468
7,500	0.9846	0.9780	0.9592
10,000	0.9886	0.9824	0.9665
15,000	0.9935	0.9878	0.9748
25,000	0.9969	0.9936	0.9823
35,000	0.9982	0.9961	0.9862
50,000	0.9990	0.9977	0.9903
75,000	0.9995	0.9988	0.9941
100,000	0.9997	0.9992	0.9961
150,000	0.9998	0.9996	0.9977
250,000	1.0000	0.9998	0.9989
350,000	-	0.9999	0.9993
500,000	-	1.0000	1.0000

Column 2 - Low Severity Insured
Column 3 - Standard Insured
Column 4 - High Severity Insured

It is assumed that the claim severity distribution is uniform between any two consecutive amounts in Column 1.

Exhibit IIa

Loss Limit	Excess Loss Premium Factor* Low Severity Insured	Standard Insured	High Severity Insured
10,000	0.191	0.270	0.391
15,000	0.146	0.222	0.353
20,000	0.118	0.187	0.322
25,000	0.098	0.162	0.296
30,000	0.084	0.143	0.274
40,000	0.064	0.116	0.237
50,000	0.052	0.098	0.208
75,000	0.033	0.070	0.156
100,000	0.023	0.053	0.124
150,000	0.010	0.034	0.083
200,000	0.003	0.023	0.056
250,000	-	0.015	0.038

Exhibit IIb

Loss Limit**	Excess Loss Premium Factor* Low Severity Insured	Standard Insured	High Severity Insured
(2,000:20,000)	0.206	0.272	0.380
(5,000:60,000)	0.114	0.170	0.276
(10,000:100,000)	0.075	0.124	0.220
(10,000:20,000)	0.155	0.228	0.350
(30,000:60,000)	0.064	0.114	0.227
(50,000:100,000)	0.038	0.076	0.166

* Expected Loss Ratio = .600

**Excess losses for a dual loss limit (A:B) are given by the following formula.

Let L be the total loss arising out of a single accident.

$$\text{If } L \leq A \quad \begin{cases} \text{Primary Loss} = L \\ \text{Excess Loss} = 0 \end{cases}$$

$$\text{If } L > A \quad \begin{cases} \text{Primary Loss} = \dfrac{L \times B}{L+B-A} \\ \text{Excess Loss} = L - \text{Primary Loss} \end{cases}$$

Exhibit III   Comparison of insurance-charges indicated by the model, and the standard formula using Table M.

Standard Premium = 50,000
No Loss Limit

		Insurance Charge*	
Min.	Max.	Standard Formula	Model
BxTM	1.00	0.267	0.300
BxTM	1.20	0.173	0.219
BxTM	1.40	0.122	0.174
BxTM	1.60	0.090	0.144
BxTM	1.80	0.068	0.123
0.60	1.00	0.254	0.299
0.60	1.20	0.117	0.195
0.60	1.40	0.038	0.124
0.60	1.60	-0.016	0.071
0.60	1.80	-0.052	0.029

Standard Premium = 150,000
No Loss Limit

		Insurance Charge*	
Min.	Max.	Standard Formula	Model
BxTM	1.00	0.173	0.179
BxTM	1.20	0.092	0.112
BxTM	1.40	0.059	0.079
BxTM	1.60	0.044	0.060
BxTM	1.80	0.029	0.047
0.60	1.00	0.150	0.171
0.60	1.20	0.047	0.087
0.60	1.40	0.000	0.043
0.60	1.60	-0.025	0.014
0.60	1.80	-0.042	-0.005

Standard Premium = 250,000
No Loss Limit

		Insurance Charge*	
Min.	Max.	Standard Formula	Model
BxTM	1.00	0.130	0.128
BxTM	1.20	0.060	0.073
BxTM	1.40	0.033	0.048
BxTM	1.60	0.025	0.033
BxTM	1.80	0.015	0.023
0.60	1.00	0.099	0.119
0.60	1.20	0.012	0.054
0.60	1.40	-0.016	0.021
0.60	1.60	-0.032	0.001
0.60	1.80	-0.040	-0.004

* The parameters for the plans are given in Exhibit IV.

Exhibit IV Parameters for Retrospective Rating Plans

	Total Standard Premium		
	50,000	150,000	250,000
Expected Losses	30,000	90,000	150,000
Loss Conversion Factor (c)	1.125	1.125	1.125
Expense in Basic Premium Factor (a)	0.149	0.139	0.134
Tax Multiplier (t)	1.040	1.040	1.040

Exhibit V Retrospective Premium Adequacy for Plans without a Loss Limit

Standard Premium = 50,000
No Loss Limit

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.951	1.000	1.127
BxTM	1.20	0.936	1.000	1.161
BxTM	1.40	0.935	1.000	1.170
BxTM	1.60	0.937	1.000	1.170
BxTM	1.80	0.940	1.000	1.163
0.60	1.00	0.951	1.000	1.112
0.60	1.20	0.951	1.000	1.103
0.60	1.40	0.962	1.000	1.084
0.60	1.60	0.974	1.000	1.066
0.60	1.80	0.984	1.000	1.049

Standard Premium = 150,000
No Loss Limit

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.951	1.000	1.119
BxTM	1.20	0.947	1.000	1.123
BxTM	1.40	0.953	1.000	1.113
BxTM	1.60	0.958	1.000	1.098
BxTM	1.30	0.962	1.000	1.085
0.60	1.00	0.956	1.000	1.078
0.60	1.20	0.964	1.000	1.052
0.60	1.40	0.976	1.000	1.028
0.60	1.60	0.987	1.000	1.008
0.60	1.80	0.994	1.000	0.992

Standard Premium = 250,000
No Loss Limit

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.961	1.000	1.102
BxTM	1.20	0.961	1.000	1.095
BxTM	1.40	0.966	1.000	1.077
BxTM	1.60	0.972	1.000	1.061
BxTM	1.30	0.977	1.000	1.048
0.60	1.00	0.967	1.000	1.061
0.60	1.20	0.975	1.000	1.031
0.60	1.40	0.987	1.000	1.007
0.60	1.60	0.996	1.000	0.988
0.60	1.80	1.004	1.000	0.974

* The parameters for the plans are given in Exhibits III and IV.

Exhibit VI Retrospective Premium Adequacy for Plans with a Loss Limit

Standard Premium = 50,000
Loss Limit = 10,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.868	0.865	0.855
BxTM	1.20	0.814	0.811	0.800
BxTM	1.40	0.819	0.818	0.813
BxTM	1.60	0.838	0.838	0.836
BxTM	1.80	0.857	0.856	0.856
0.60	1.00	0.868	0.865	0.855
0.60	1.20	0.829	0.827	0.816
0.60	1.40	0.864	0.863	0.859
0.60	1.60	0.912	0.913	0.912
0.60	1.80	0.958	0.961	0.962

Standard Premium = 150,000
Loss Limit = 30,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.904	0.908	0.901
BxTM	1.20	0.889	0.894	0.889
BxTM	1.40	0.906	0.909	0.907
BxTM	1.60	0.924	0.925	0.924
BxTM	1.80	0.939	0.939	0.939
0.60	1.00	0.908	0.912	0.905
0.60	1.20	0.912	0.916	0.914
0.60	1.40	0.944	0.945	0.947
0.60	1.60	0.974	0.973	0.977
0.60	1.80	0.995	0.994	0.999

Standard Premium = 250,000
Loss Limit = 50,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.925	0.931	0.937
BxTM	1.20	0.923	0.927	0.931
BxTM	1.40	0.940	0.941	0.944
BxTM	1.60	0.957	0.957	0.958
BxTM	1.80	0.969	0.969	0.969
0.60	1.00	0.931	0.936	0.943
0.60	1.20	0.942	0.944	0.948
0.60	1.40	0.970	0.967	0.969
0.60	1.60	0.992	0.988	0.987
0.60	1.80	1.010	1.005	1.003

* The parameters for the plans are given in Exhibits IIa, III and IV.

Exhibit VII Retrospective Premium Adequacy for Plans with a Loss Limit and Inadequate Excess Loss Premium Factors

Standard Premium = 50,000
Loss Limit = 10,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.899	0.914	0.936
BxTM	1.20	0.884	0.919	0.978
BxTM	1.40	0.910	0.955	1.031
BxTM	1.60	0.939	0.989	1.076
BxTM	1.80	0.964	1.017	1.110
0.60	1.00	0.899	0.914	0.937
0.60	1.20	0.906	0.944	1.009
0.60	1.40	0.963	1.013	1.102
0.60	1.60	1.021	1.073	1.166
0.60	1.80	1.069	1.121	1.213

Standard Premium = 150,000
Loss Limit = 30,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.928	0.952	1.003
BxTM	1.20	0.930	0.967	1.048
BxTM	1.40	0.955	0.994	1.089
BxTM	1.60	0.976	1.016	1.120
BxTM	1.80	0.993	1.034	1.142
0.60	1.00	0.933	0.957	1.009
0.60	1.20	0.954	0.988	1.062
0.60	1.40	0.991	1.024	1.103
0.60	1.60	1.022	1.054	1.135
0.60	1.80	1.045	1.076	1.156

Standard Premium = 250,000
Loss Limit = 50,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.943	0.968	1.028
BxTM	1.20	0.952	0.982	1.060
BxTM	1.40	0.972	1.004	1.088
BxTM	1.60	0.990	1.023	1.110
BxTM	1.80	1.004	1.038	1.127
0.60	1.00	0.950	0.974	1.027
0.60	1.20	0.970	0.996	1.056
0.60	1.40	1.000	1.024	1.083
0.60	1.60	1.023	1.045	1.102
0.60	1.80	1.042	1.063	1.118

* The parameters for the plans are given in Exhibits IIa, III and IV. The Excess Loss Premium Factors in Exhibit IIa are multiplied by .5.

Exhibit VIII Distribution of Retrospective Premium with 30,000 Loss Limit - Standard Insured

1.	Standard Premium	150000
2.	Basic Premium (Excl Ins Chg But Incl. Tax)	21684
3.	Basic Premium (Incl 0.179 Ins Chg and Tax)	53098
4.	Excess Premium Generated by E.L.P.F. (Inc Tax)	25062
5.	Needed Excess Premium (Inc Tax)	25062
6.	Minimum Premium (= Line 3)	53098
7.	Maximum Premium (Line 1 x 1.000)	150000

A	B	C	D	E
Probability that Subject Losses Are < = Col B *	Losses Subject To Retro Rating *	Retrospective Premium **	Cost Plus Premium***	Difference C - D
Min	10659	88819	57405	31414
.005	18287	96447	65033	31414
.010	20942	99102	67688	31414
.050	30342	108502	77088	31414
.100	37238	115398	83984	31414
.200	48255	126415	95001	31414
.300	57966	136126	104712	31414
.400	66673	144833	113419	31414
.500	75372	150000	122118	27882
.600	84315	150000	131061	18939
.700	95106	150000	141852	8148
.800	108743	150000	155489	-5489
.900	129005	150000	175751	-25751
.950	147786	150000	194532	-44532
.990	184776	150000	231522	-81522
.995	200951	150000	247697	-97697
Max	283075	150000	329821	-179821

Notes

* Subject Losses are adjusted to include L.A.E. and Taxes

** Retrospective Premium = Line 3 + Line 4 + Col B
Subject to Minimum and Maximum Premium

*** Cost Plus Premium = Line 2 + Line 5 + Col B

Exhibit IX Distribution of Retrospective Premium with No Loss Limit - Standard Insured

1.	Standard Premium	150000
2.	Basic Premium (Excl Ins Chg But Incl. Tax)	21684
3.	Basic Premium (Incl 0.179 Ins Chg and Tax)	53098
4.	Excess Premium Generated by E.L.P.F. (Inc Tax)	0
5.	Needed Excess Premium (Inc Tax)	0
6.	Minimum Premium (= Line 3)	53098
7.	Maximum Premium (Line 1 x 1.000)	150000

A	B	C	D	E
Probability that Subject Losses Are < = Col B *	Losses Subject To Retro Rating *	Retrospective Premium **	Cost Plus Premium***	Difference C - D
Min	10659	63757	32343	31414
.005	18287	71385	39971	31414
.010	20942	74040	42626	31414
.050	30342	83440	52026	31414
.100	37238	90336	58922	31414
.200	48273	101371	69957	31414
.300	58668	111766	80352	31414
.400	69178	122276	90862	31414
.500	81194	134292	102878	31414
.600	94581	147679	116265	31414
.700	112488	150000	134172	15828
.800	140164	150000	161848	-11848
.900	190628	150000	212312	-62312
.950	258305	150000	279989	-129989
.990	532459	150000	554143	-404143
.995	615667	150000	637351	-487351
Max	938677	150000	960361	-810361

Notes

* Subject Losses are adjusted to include L.A.E. and Taxes

** Retrospective Premium = Line 3 + Line 4 + Col B
Subject to Minimum and Maximum Premium

*** Cost Plus Premium = Line 2 + Line 5 + Col B

Exhibit X Indicated Insurance Charges

Standard Premium = 50,000
No Loss Limit

Min.	Max.	Insurance Charge* Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.230	0.300	0.424
BxTM	1.20	0.153	0.219	0.351
BxTM	1.40	0.113	0.174	0.305
BxTM	1.60	0.089	0.144	0.269
BxTM	1.80	0.072	0.123	0.241
0.60	1.00	0.226	0.299	0.424
0.60	1.20	0.129	0.195	0.351
0.60	1.40	0.071	0.124	0.289
0.60	1.60	0.034	0.071	0.224
0.60	1.80	0.006	0.029	0.159

Standard Premium = 150,000
No Loss Limit

Min.	Max.	Insurance Charge* Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.118	0.179	0.303
BxTM	1.20	0.063	0.112	0.217
BxTM	1.40	0.039	0.079	0.168
BxTM	1.60	0.026	0.060	0.135
BxTM	1.80	0.018	0.047	0.110
0.60	1.00	0.111	0.171	0.300
0.60	1.20	0.046	0.087	0.181
0.60	1.40	0.017	0.043	0.096
0.60	1.60	-0.000	0.014	0.031
0.60	1.80	-0.012	-0.005	-0.021

Standard Premium = 250,000
No Loss Limit

Min.	Max.	Insurance Charge* Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.083	0.128	0.234
BxTM	1.20	0.039	0.073	0.154
BxTM	1.40	0.021	0.048	0.109
BxTM	1.60	0.011	0.033	0.080
BxTM	1.80	0.005	0.023	0.060
0.60	1.00	0.079	0.119	0.222
0.60	1.20	0.030	0.054	0.107
0.60	1.40	0.009	0.021	0.033
0.60	1.60	-0.003	0.001	-0.021
0.60	1.80	-0.010	-0.014	-0.061

* The parameters for the plan are given in Exhibit IV.

Exhibit XI Indicated Insurance Charges

Standard Premium = 50,000
Loss Limit = 10,000

		Insurance Charge*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.054	0.049	0.032
BxTM	1.20	0.013	0.012	0.006
BxTM	1.40	0.003	0.003	0.001
BxTM	1.60	0.001	0.001	0.000
BxTM	1.80	0.000	0.000	0.000
0.60	1.00	0.052	0.049	0.032
0.60	1.20	0.008	0.009	0.006
0.60	1.40	-0.004	0.000	0.001
0.60	1.60	-0.006	-0.003	0.000
0.60	1.80	-0.007	-0.004	0.000

Standard Premium = 150,000
Loss Limit = 30,000

		Insurance Charge*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.046	0.052	0.045
BxTM	1.20	0.010	0.013	0.011
BxTM	1.40	0.002	0.004	0.003
BxTM	1.60	0.000	0.001	0.001
BxTM	1.80	0.000	0.000	0.000
0.60	1.00	0.041	0.047	0.044
0.60	1.20	0.002	0.004	0.007
0.60	1.40	-0.006	-0.006	-0.003
0.60	1.60	-0.008	-0.009	-0.005
0.60	1.80	-0.009	-0.010	-0.006

Standard Premium = 250,000
Loss Limit = 50,000

		Insurance Charge*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.038	0.044	0.052
BxTM	1.20	0.007	0.010	0.013
BxTM	1.40	0.001	0.002	0.004
BxTM	1.60	0.000	0.000	0.001
BxTM	1.80	0.000	0.000	0.000
0.60	1.00	0.035	0.039	0.047
0.60	1.20	0.002	0.001	0.003
0.60	1.40	-0.004	-0.007	-0.007
0.60	1.60	-0.006	-0.009	-0.011
0.60	1.80	-0.006	-0.010	-0.011

* The parameters for the plan are given in Exhibits IIa and IV.

Exhibit XII Retrospective Premium Adequacy for Alternate Plan #1

Standard Premium = 50,000
Loss Limit = 10,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	1.004	1.000	0.983
BxTM	1.20	1.002	1.000	0.993
BxTM	1.40	1.001	1.000	0.998
BxTM	1.60	1.000	1.000	0.999
BxTM	1.80	1.000	1.000	1.000
0.60	1.00	1.002	1.000	0.983
0.60	1.20	0.998	1.000	0.997
0.60	1.40	0.996	1.000	1.002
0.60	1.60	0.996	1.000	1.004
0.60	1.80	0.996	1.000	1.006

Standard Premium = 150,000
Loss Limit = 30,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
[illegible]	1.00	0.994	1.000	0.994
BxTM	1.20	0.996	1.000	0.997
BxTM	1.40	0.998	1.000	0.998
BxTM	1.60	0.999	1.000	0.999
BxTM	1.80	1.000	1.000	1.000
0.60	1.00	0.995	1.000	0.998
0.60	1.20	0.998	1.000	1.003
0.60	1.40	0.999	1.000	1.003
0.60	1.60	1.001	1.000	1.004
0.60	1.80	1.001	1.000	1.005

Standard Premium = 250,000
Loss Limit = 50,000

		Retrospective Premium Adequacy*		
Min.	Max.	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.994	1.000	1.008
BxTM	1.20	0.997	1.000	1.004
BxTM	1.40	0.999	1.000	1.002
BxTM	1.60	1.000	1.000	1.001
BxTM	1.80	1.000	1.000	1.000
0.60	1.00	0.996	1.000	1.007
0.60	1.20	1.001	1.000	1.003
0.60	1.40	1.003	1.000	1.000
0.60	1.60	1.004	1.000	0.998
0.60	1.80	1.005	1.000	0.998

* The insurance charges used are those of the Standard Insured in Exhibit XI. The parameters for the plan are given in Exhibits IIa and IV.

Exhibit XIII Retrospective Premium Adequacy for Alternate Plan #2

Standard Premium = 50,000
Loss Limit = (2,000:20,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.055	0.999	1.000	0.992
BxTM	1.20	0.015	0.999	1.000	0.997
BxTM	1.40	0.005	0.999	1.000	0.998
BxTM	1.60	0.001	1.000	1.000	1.000
BxTM	1.80	0.000	1.001	1.000	1.000
0.60	1.00	0.055	0.998	1.000	0.992
0.60	1.20	0.014	0.996	1.000	0.998
0.60	1.40	0.002	0.997	1.000	1.002
0.60	1.60	-0.002	0.998	1.000	1.004
0.60	1.80	-0.003	0.998	1.000	1.004

Standard Premium = 150,000
Loss Limit = (5,000:60,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.046	0.992	1.000	1.007
BxTM	1.20	0.012	0.994	1.000	1.003
BxTM	1.40	0.003	0.998	1.000	1.002
BxTM	1.60	0.001	0.999	1.000	1.000
BxTM	1.80	0.000	1.000	1.000	1.000
0.60	1.00	0.043	0.993	1.000	1.008
0.60	1.20	0.006	0.997	1.000	1.006
0.60	1.40	-0.003	1.000	1.000	1.003
0.60	1.60	-0.005	1.000	1.000	1.001
0.60	1.80	-0.006	1.001	1.000	1.001

Standard Premium = 250,000
Loss Limit = (10,000:100,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.039	0.993	1.000	1.014
BxTM	1.20	0.008	0.997	1.000	1.009
BxTM	1.40	0.002	0.999	1.000	1.003
BxTM	1.60	0.000	1.000	1.000	1.002
BxTM	1.80	0.040	1.000	1.000	1.000
0.60	1.00	0.036	0.994	1.000	1.013
0.60	1.20	0.003	1.000	1.000	1.005
0.60	1.40	-0.004	1.002	1.000	1.000
0.60	1.60	-0.006	1.003	1.000	0.998
0.60	1.80	-0.006	1.003	1.000	0.997

* The parameters for the plan are given in Exhibits IIb and IV.

Exhibit XIV Retrospective Premium Adequacy for Alternate Plan #3

Standard Premium = 50,000
Loss Limit = (10,000:20,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.029	1.000	1.000	0.987
BxTM	1.20	0.078	0.999	1.000	0.992
BxTM	1.40	0.026	0.998	1.000	0.995
BxTM	1.60	0.010	0.999	1.000	0.997
BxTM	1.80	0.004	1.000	1.000	1.000
0.60	1.00	0.077	0.998	1.000	0.988
0.60	1.20	0.019	0.995	1.000	1.000
0.60	1.40	-0.001	0.995	1.000	1.009
0.60	1.60	-0.008	0.996	1.000	1.012
0.60	1.80	-0.011	0.996	1.000	1.014

Standard Premium = 150,000
Loss Limit = (30,000:60,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.071	0.989	1.000	1.004
BxTM	1.20	0.022	0.992	1.000	1.004
BxTM	1.40	0.008	0.995	1.000	1.001
BxTM	1.60	0.003	0.998	1.000	1.000
BxTM	1.80	0.001	0.999	1.000	1.000
0.60	1.00	0.064	0.991	1.000	1.007
0.60	1.20	0.008	0.997	1.000	1.004
0.60	1.40	-0.009	1.001	1.000	1.002
0.60	1.60	-0.014	1.003	1.000	0.999
0.60	1.80	-0.016	1.004	1.000	0.999

Standard Premium = 250,000
Loss Limit = (50,000:100,000)

			Retrospective Premium Adequacy*		
Min.	Max.	Insurance Charge*	Low Severity Insured	Standard Insured	High Severity Insured
BxTM	1.00	0.058	0.980	1.000	1.019
BxTM	1.20	0.016	0.995	1.000	1.013
BxTM	1.40	0.005	0.997	1.000	1.006
BxTM	1.60	0.001	1.000	1.000	1.003
BxTM	1.80	0.000	1.000	1.000	1.002
0.60	1.00	0.051	0.994	1.000	1.014
0.60	1.20	0.004	1.001	1.000	1.003
0.60	1.40	-0.009	1.005	1.000	0.996
0.60	1.60	-0.013	1.007	1.000	0.992
0.60	1.80	-0.014	1.007	1.000	0.991

* The parameters for the plan are given in Exhibits IIb and IV.

Exhibit XVa Multi-Exposure Insured

Standard Premium for:	High Severity Insured	=	150,000
	Standard Insured	=	50,000
	Low Severity Insured	=	50,000
Total			250,000

		Indicated Insurance Charge*	
Min.	Max.	No Loss Limit	50,000 Loss Limit
BxTM	1.00	0.183	0.047
BxTM	1.20	0.115	0.011
BxTM	1.40	0.080	0.002
BxTM	1.60	0.057	0.000
BxTM	1.80	0.042	0.000
0.60	1.00	0.175	0.044
0.60	1.20	0.086	0.003
0.60	1.40	0.033	-0.006
0.60	1.60	-0.002	-0.009
0.60	1.80	-0.028	-0.009

Exhibit XVb Multi-Exposure Insured

Loss Limit = 50,000

Min.	Max.	Insurance Charge**	Retrospective Premium Adequacy*
BxTM	1.00	0.044	1.001
BxTM	1.20	0.010	1.000
BxTM	1.40	0.002	1.000
BxTM	1.60	0.000	1.000
BxTM	1.80	0.000	0.999
0.60	1.00	0.039	1.001
0.60	1.20	0.001	1.001
0.60	1.40	-0.007	0.999
0.60	1.60	-0.009	0.999
0.60	1.80	-0.010	0.999

* The parameters for the plan are given in Exhibits IIa and IV.

** From Exhibit XI.

AN ANALYSIS OF RETROSPECTIVE RATING

by Glenn Meyers

Reviewed by James F. Golz

Glenn Meyers has written a fine, concise paper. He begins with hypothetical loss distributions representing low, standard, and high workers' compensation severities. Combining these with a Poisson frequency distribution, he demonstrates how our present retrospective rating procedure fails to react properly to severity differences and how it overcharges (at least theoretically) when loss limits are selected. Meyers notes that a complete computer modeling of the interaction among excess loss premium factors, insurance charges, and frequency and severity distributions is still a lengthy process. However, he observes that once excess loss premium factors are properly accounted for, the remaining insurance charge is approximately equal regardless of the severity distribution. Thus, for practical reasons, he suggests that either a limited number of loss limits or a single mandatory loss limit be imposed.

Our first task, as Meyers observes, should be to confirm his conclusions using actual data. The adjustment of reported workers' compensation data to a suitable level for analysis will be an interesting task in itself. The claims will, of course, have to be put on the level of the current law. The late recognition of many severe claims and the present valuing of pension claims will complicate the development of individual claims to ultimate cost levels.

The frequency and severity distributions may even have to be adjusted to reflect changing benefit utilization patterns.

There may be a practical difficulty in implementing Meyers's suggestion to restrict the number of loss limits available. Our retrospective rating plans have grown to their current level of detail in response to the competitiveness within our industry and the demands of the marketplace. I suspect that many underwriters would rather have some protection from even an unbalanced retrospective rating plan than not be able to use retro at all as a defensive tool.

A further complication arises from the fact that retrospective rating plans are employed in audited lines of insurance. We face not only the problem of combining states, lines and severities for quotation purposes, but also of determining the rating parameters at adjustment time. Although we live in an age of computers, not everyone has access to them. Therefore, our retrospective rating plans should be simple enough for manual calculation and adjustment. Recently the National Council on Compensation Insurance considered proposing that insurance charges be determined using expected number of claims (which would vary by hazard group) rather than expected loss dollars. This would have been a reflection of risk severity, much as Meyers proposes. However, the matter was eventually dropped, apparently for fear that the retrospective rating process would become too complex.

Meyers defines the basic premium factor to include all expenses other than taxes and loss adjustment expenses. This is a useful simplification often adopted for educational purposes. However, the expenses remaining in the basic premium are actually a function of the loss conversion factor, and many plans are sold with loss conversion factors higher or lower than the actual relationship of loss adjustment expenses to losses. It should be noted that such plans can be analyzed by Meyers's methods by using the actual relationship of loss adjustment expenses to losses in the "cost-plus" formula and the selected loss conversion factor in the "retrospective premium" formula.

It will be interesting to see whether Meyers's suggestions can be implemented for all retrospective rating. In any event, I expect the wise company will soon apply severity analyses of this sort to the pricing of large accounts.

TITLE: ESTIMATING AGGREGATE LOSS PROBABILITY AND INCREASED LIMIT FACTOR

AUTHOR: Dr. Shaw Mong

Dr. Mong is an Actuarial Consultant for the Financial Research and Planning Division of Fred. S. James & Co. He received his PhD degree in Mathematics from John Hopkins University. Shaw holds a visiting membership in the Institute for Advanced Study and has written several articles on mathematical and statistical subjects.

REVIEWER: A review from the original reviewer was not received. It is intended that a review by an alternate reviewer be distributed at the meeting (May 11-14).

Aggregate loss probability is an effective tool in actuarial rate making, risk charging, and retention analysis for both primary and secondary insurance companies. A noticeable trend over recent years indicates that it also is becoming an indispensable element in the risk management operations of many manufacturing and commercial firms. Some major insurance brokerage houses in the U.S., in step with the trend, already employ this technique routinely in selecting a retention plan for their clients. In its broadest form, the application extends beyond the actuarial domain into the broader area of corporate financial planning.

Most existing procedures for estimating aggregate loss probability distributions have significant disadvantages. Most often, these disadvantages are associated with inadequate treatment of skewed data. The purpose of this paper is to present a recently developed technique which seems to handle the aggregate loss estimation problem more effectively.

The first section presents a brief review of the strength and weakness of most popular techniques currently in use. This is followed by a brief description of the newly developed technique. Next, the results of a comparative study of the cost and effectiveness of these alternative procedures are reported. Finally, we illustrate the impact of improved aggregate loss estimation on the pricing of reinsurance. An appendix contains the mathematical derivation for those who would like to verify our results.

Standard Aggregate Loss Estimation Procedures

In dealing with the estimation of aggregate loss probability, there are three fundamental approaches commonly in use. They are analytical, approximation, and simulation models. Each is distinguished from the others by its own characteristics, advantages and disadvantages. The pure analytical model[1] generally is the most accurate. The handicap is that it can be applied to only a few distribution types. A frequently used approximation model is the Normal Power approximation. This is easy to implement but yields disturbingly large approximation errors when applied to highly skewed data.[2] Another less well-known approximation technique is the Gamma approximation,[3] which seems more accurate than the NP approximation in most occasions.[4] The only weakness of the Gamma approximation is that, like the NP approximation, it does not respond to the sensitive choice of frequency distribution. Simulation modeling is perhaps the most widely used technique in the field of management sciences; however, like the other techniques, it has disadvantages, too. First, since the error brought on from simulation is statistical rather than mathematical, it can be reduced significantly only by increasing considerably the number of iterations.[5] This would be an unfavorable element should the consideration of computing time and cost become crucial. Secondly, simulation is a brute force technique and offers limited insight into how a system works. Thus, any sensitivity analysis or optimization drawn from a simulation model is virtually a trial and error process and can not be justified mathematically.

---

[1] See Appendix B for a summary.

[2] Reports compiled from experiments decline to recommend the use of NP approximation on data of skewness exceeding 1 or 2, see [2] and [13].

[3] See Appendix C for background materials on this technique.

[4] There is a controversy in the literature [16] and [18] concerning which approximation is superior. In our study, we found out that at least for the distributions listed in this article, the result for the gamma approximation is much better than that from the NP approximation.

[5] See Table 7.1 given in [2], p. 93 for relation between the degree of error and number of iterations.

The aim of this paper is to introduce a new model which is designed to meet the dual requirements of accuracy and simplicity in implementation. Our approach is a blend of the analytical and approximation models. It is approximate, because the answer is not the exact, and analytical primarily because the formula is derived from the fundamental characteristics of collective risk theory. To demonstrate the precision of our model, apart from the mathematical deduction attached as an appendix, we compare the results of the new model with those where the exact probability can be calculated directly using the analytical method.

A New Model (Modified Gamma Approximation)

Aggregate loss, occurring as a random process, is compiled from two variables: one is identified as the number of claims experienced in a given time span (normally one year) and denominated as the "frequency of loss." The other is the size of an individual claim and is termed the "severity of loss." Jointly, frequency and severity determine total or aggregate loss from all claims in the given time span. The most often used frequency distributions are poisson and negative binomial.[6] For severity distributions, experience[7] indicates that normal, gamma, inverse normal, pareto, log-normal and log-gamma[8] are appropriate for casualty and property insurance.

---

[6] Some authors also recommend a third type, the generalized Waring distribution, for details please see [19] .

[7] See [3] , [8] , [10] , [11] , and [18] .

[8] *A summary of these distributions can be found in Appendix A.*

For convenience, we shall adopt the term generalized poisson model for the aggregate loss distribution which uses the poisson distribution as the frequency function and leaves the choice of the severity function open. Similarly, the generalized negative binomial model reflects the application of the negative binomial distribution as the frequency function.

To describe our formula, we need the following statistics which can be estimated[9] from the sample data:

λ	:	frequency mean
σ_f	:	frequency standard deviation
μ_s	:	severity mean
σ_s	:	severity standard deviation

and statistics which can be derived intrinsically:

μ	:	aggregate mean (e.g., the product of λ and μ_s)[10]
σ	:	aggregate standard deviation (e.g., $\lambda(\mu_s^2+\sigma_s^2)$ for the generalized poisson model and $\lambda\sigma_s^2+\mu_s^2\sigma_f^2$ for the generalized $n.b.$ model)[10]
γ_s	:	severity skewness

Our formula states that the probability $F(x)$ of annual aggregate loss less than or equal to x is given by:[11]

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty f(t)\,\frac{\sin\left(xt/\sigma + g(t)\right)}{t}\,dt \qquad (1)$$

---

[9] See [7] for the estimation of these statistics.

[10] See [14] p. 179 for the derivation.

[11] The derivation of formula (1) and the subsequent tables are given in Appendix D.

where functions $f(t)$ and $g(t)$ in the integrand are defined by:

Table 1

Model	$f(t)$	$g(t)$
Generalized Poisson	$\exp(\lambda \ell(t))$	$-\lambda k(t)$
Generalized NB	$\left\{\left(1-\left(\frac{\sigma_\rho^2-\lambda}{\lambda}\right)\ell(t)\right)^2+\left(\left(\frac{\sigma_\rho^2-\lambda}{\lambda}\right)k(t)\right)^2\right\}^{-\frac{\lambda^2}{2(\sigma_\rho^2-\lambda)}}$,	$\left(\frac{\lambda^2}{\sigma_\rho^2-\lambda}\right)\tan^{-1}\left(\frac{(\sigma_\rho^2-\lambda)k(t)}{\lambda-(\sigma_\rho^2-\lambda)\ell(t)}\right)$

and functions $\ell(t)$ and $k(t)$ in both models are given by:

Table 2

$\ell(t)$	$k(t)$	$\delta(t)$	$\theta(t)$
$\delta(t)\cos(\theta(t))-1$	$\delta(t)\sin(\theta(t))$	$\left(1+\left(\frac{t\gamma_\rho\sigma_\rho}{2\sigma}\right)^2\right)^{-\frac{2}{\gamma_\rho^2}}$	$t\left(\frac{\mu_\rho-2\sigma_\rho/\gamma_\rho}{\sigma}\right)+\frac{4}{\gamma_\rho^2}\tan^{-1}\left(\frac{t\gamma_\rho\sigma_\rho}{2\sigma}\right)$

The only quantity which has not been expressed explicitely in the formula is the severity skewness γ_ρ . Since each of the six severity distribution functions has exactly two parameters, each is defined and described completely by the severity sample mean and standard deviation. All the other quantities, including γ_ρ , depend ultimately on the type of the severity distribution chosen; that is on the sample mean and standard deviation. The corresponding severity skewness of the six alternative severity distributions are tabulated as follows:[12]

Table 3

Type	Skewness[13] γ_ρ
Normal	0
Gamma	$2\left(\frac{\sigma_\rho}{\mu_\rho}\right)$

---

12 The derivation of Table 3 is given in Appendix A.

13 If the skewness is zero, replace it with any small number (e.g., 10^{-9}) in the computation, since dividing by zero is prohibited in our formula.

Inverse Normal	$3\left(\frac{\sigma_p}{\mu_p}\right)$
Pareto[14]	$2\left(\frac{\sigma_p}{\mu_p}\right)\left(\frac{3\sigma_p^2-\mu_p^2}{3\mu_p^2-\sigma_p^2}\right)$
Log-normal	$\left(\frac{\sigma_p}{\mu_p}\right)\left(\left(\frac{\sigma_p}{\mu_p}\right)^2+3\right)$
Log-gamma[14]	*[15]

Formula (1) and its consequent computations may seem complex in the form shown above. However, the implementation is quite simple. Any standard numerical integration technique would handle the computation effectively; for example, the extended Simpson's rule is adequate to calculate the integration in (1) and is easy to code in any scientific programming language. A practical discussion on the use of extended Simpson's rule and the truncated range of integration in formula (1) is given in Appendix D.

Effectiveness of the Modified Gamma Approach

From a conceptual point of view, the new model seems to satisfy the objective of increased accuracy at nominal cost. The ultimate test, however, lies in its effectiveness in handling actual loss data.

By combining the poisson or negative binomial (for frequency) with the normal, gamma, or inverse normal (for severity) it is possible to compute an exact aggregate distribution using the pure analytical method (A).

---

14 In the cases of the pareto and log-gamma distributions, the skewness may not always exist; it depends on the relation between the sample mean and standard deviation. Thus, if the following conditions

$3\mu_p^2 > \sigma_p^2$ for pareto ,

$\log(\sigma_p^2+\mu_p^2)/\log\mu_p \leq 2.7095$ for log-gamma

are not met, the new model is not applicable. See Appendix A for details.

15 Since the skewness of log-gamma does not admit any closed form in terms of the sample mean and standard deviation, it is best expressed by its functional parameters, see Appendix A.

This procedure was used to provide a series of control distributions for a comparison of the relative accuracy of the normal power approximation (NP), standard gamma approximation (G), and the new, modified gamma approximation (MG).

In the analysis each of the four methods was used to generate aggregate probability distributions for several sets of hypothetical loss data. The primary variation in the data reflected differences in skewness (from a relatively modest .5 to a substantial skewness factor of 5). The points on the probability distribution were chosen in terms of standardized deviations from the mean rather than in absolute dollar amounts. Calculations were made utilizing both the generalized poisson and the generalized negative binomial models.[16]

Exhibit I, presents three sets of data for the generalized poisson model, as does exhibit II for generalized negative binomial model. In each set, the severity type, severity coefficient of variation and frequency mean are selected (in the case of negative binomial the frequency variance is also required), and the aggregate skewness is calculated by the aid of Table A6 given in appendix C2. Two auxiliary exhibits, labeled by Ia and IIa respectively, display the difference between results obtained from analytic method and the other three methods. At the bottom row, their variances, calculated by summing the squares of the difference dividing by the number of rows, are computed respectively.

---

16 The objective of the analysis was to uncover any systematic bias or approximation errors inherent in the alternative approximation techniques. In normal practice the candidate distributions would be determined by a goodness of fit criterion.

As can be seen from both exhibits I and II, the new model clearly is superior to the other two approximation models in all scenarios. The discrepancy of NP approximation is particularly serious not only on highly skewed data but also on modestly skewed data (e.g., $\gamma_p =$ [illegible]). Also notice that in both the generalized poisson and negative binomial models, the results from the standard gamma and NP approximation are determined ultimately by the skewness, e.g., the differences in the control distributions reflecting the choice of frequency distribution are not captured by either traditional approximation methods. The new model does detect the difference between poisson and negative binomial frequency distributions.

Finally, we want to indicate the degree of sensitivity of the estimated aggregate loss probability to the selection of the type of severity function. Exhibit IV assumes that the frequency distribution is poisson (with mean = 60.383) and the estimated severity coefficient of variation is equal to 4. If the severity function is the inverse normal the aggregate skewness would be 1.5. The same parameter would be 9.02 for log-normal. Also a tail appears in the aggregate picture when the log-normal is selected for the severity. This phenomenon can be explained mathematically by the following observation: given a severity sample mean and variance, the magnitude of the severity skewness, according to Table 3, can be arranged in the following increasing order:

$$\text{normal, gamma, inverse normal,} \quad \left\{ \begin{array}{l} \text{pareto} \\ \text{log-normal} \end{array} \right\} .$$

Since the aggregate skewness varies along with the severity skewness, the selection of log-normal as severity function always yields a larger aggregate skewness than does the selection of inverse normal.

Increased Limits Factors for Stop-Loss Reinsurance

One of the practical applications of estimating aggregate loss probability is its use in excess of loss pricing, aggregate pricing and stop-loss reinsurance. The case of excess of loss pricing has been covered extensively in a recent article by Robert S. Miccolis.[17] We would like to concentrate on the latter two situations here.

Aggregate pricing and stop-loss reinsurance are fundamentally one concept. Stop-loss reinsurance is a process which transfers the risk above an aggregate limit to a reinsurer. Aggregate pricing structure can be envisaged as zero limit stop-loss reinsurance pricing structure, e.g., the reinsurer absorbs all the loss. Thus, as far as the pricing structure is concerned, we can treat aggregate pricing as a special case of stop-loss reinsurance pricing.

If F(x), as before, represents the aggregate loss probability distribution, without an aggregate limit, then let $F_L(x)$ be the truncated distribution where an aggregate limit L is introduced, μ_L and σ_L are respectively the mean and standard deviation of $F_L(x)$. The formula for premium, excluding loss expense, charged for stop-loss coverage of an aggregate limit L is given[18] as follows:

$$P_L = \mu_L + c\,\sigma_L^2 \qquad (2)$$

---

17 See [15] .

18 See [5] p. 85-87. An alternate suggestion for the safety loading in formula (2) is to use the standard deviation σ_L instead of variance σ_L^2, see [1] .

where the loading coefficient C generally is chosen from experience. If L is zero (e.g., this is a full stop-loss coverage for the primary carrier) μ_L and σ_L become aggregate μ and σ as specified before. Suppose a is the loss expense ratio, then the total premium charged for a stop-loss coverage with limit L is $(1+a)P_L$. Then by definition, the increased limit factor, $I(L)$, of a stop-loss policy limit L imposed on a stop-loss basic limit L_0 is

$$I(L) = \frac{\text{Total premium of policy limit } L}{\text{Total premium of basic limit } L_0}$$

$$= \frac{(1+a)P_L}{(1+a)P_{L_0}} = \frac{\mu_L + c\,\sigma_L^2}{\mu_{L_0} + c\,\sigma_{L_0}^2} \qquad (3)$$

A formula is needed to calculate μ_L and σ_L^2. This can be worked out from the truncated distribution $F_L(x)$. Since the aggregate loss of the reinsurer under a stop-loss coverage with a policy limit L is reduced by an amount of L dollars, the probability $F_L(x)$ is given by:

$$F_L(x) = F(x+L) \qquad (4)$$

Hence, the jth moment, $\mu_{L,j}$ accordingly is defined by:

$$\mu_{L,j} = \int_0^\infty x^j \, dF_L(x)$$

$$= \int_0^\infty x^j \, dF(x+L)$$

(replaced variable $x+L$ by x)

$$= \int_L^\infty (x-L)^j \, dF(x)$$

particularly, when j = 1 and 2, we have

$$\mu_L = \int_L^\infty (x-L)\,dF(x) = \int_0^\infty x\,dF(x) - \int_0^L x\,dF(x) - L\int_L^\infty dF(x)$$

$$= \mu - \int_0^L x\,dF(x) - L\,(1-F(L)) \qquad (5)$$

$$\mu_{L,2} = \int_L^\infty (x-L)^2\,dF(x)$$

$$= \int_0^\infty x^2 dF(x) - \int_0^L x^2 dF(x) - 2L\int_L^\infty (x-L)\,dF(x) + L^2\int_L^\infty dF(x)$$

$$= \sigma^2 + \mu^2 - \int_0^L x^2 dF(x) - 2L\mu_L + L^2(1-F(L)) \qquad (6)$$

Notice that $\sigma_L^2 = \mu_{L,2} - \mu_L^2$, thus two more values:

$$\int_0^L x^j \, dF(x), \qquad j = 1, 2 \tag{7}$$

have to be calculated before we compute formula (3). For this, the precise form of $F(x)$ would come into play. Since by (1),

$$\frac{dF(x)}{dx} = \frac{1}{\pi\sigma}\int_0^\infty f(t)\cos\left(xt/\sigma + g(t)\right)dt \tag{8}$$

Substitute $dF(x)$ in (7) by (8). We have

$$\int_0^L x^j dF(x) = \frac{1}{\pi\sigma}\int_0^L x^j dx \int_0^\infty f(t)\cos(xt/\sigma + g(t))dt$$

(exchange the order of integration)

$$= \frac{1}{\pi\sigma}\int_0^\infty f(t)dt \int_0^L x^j \cos(xt/\sigma + g(t))\, dx \tag{9}$$

Now the first integrand $\int_0^L x^j \cos(xt/\sigma + g(t))dx$ (denoted by $w_j(t)$) in (9)

has a closed form, and the desired values are given as follows:

Table 4

$w_1(t)$	$w_2(t)$
$\frac{L\sigma \sin(Lt/\sigma + g(t))}{t} + \frac{\sigma^2\{\cos(Lt/\sigma + g(t)) - \cos(g(t))\}}{t^2}$,	$\frac{L^2\sigma \sin(Lt/\sigma + g(t))}{t} + \frac{2L\sigma^2 \cos(Lt/\sigma + g(t))}{t^2} - \frac{2\sigma^3\{\sin(Lt/\sigma + g(t)) - \sin(g(t))\}}{t^3}$

where g(t) and f(t) are given as before in Tables 1 and 2.

In summary, the increased limits factor I(L) is calculated by formula (3), where $\sigma_L^2 = \mu_{L,2} - \mu_L^2$ with μ_L and $\mu_{L,2}$ given by (5) and (6). Whereas in formula (5) and (6), $\int_0^L x^j dF(x)$, $j=1,2$ is calculated by:

Table 5

$\int_0^L x\,dF(x)$	$\int_0^L x^2 dF(x)$
$\frac{1}{\pi\sigma}\int_0^\infty f(t)w_1(t)dt$	$\frac{1}{\pi\sigma}\int_0^\infty f(t)w_2(t)dt$

The integrations in Table 5 can be handled by any numerical integration technique as discussed before, e.g., extended Simpson's rule, etc. Exhibit III illustrates an increased limits table derived by formula (3) and tables 4 and 5.

Conclusion

The effectiveness of the estimation of aggregate loss probability and the aggregate pricing model introduced in this article will, to a great extent, depend on how consistantly the loss-experience data is treated. In our model, we assume that all the losses have already been adjusted to the present or ultimate level. That is: losses have been developed to the ultimate; IBNR has been adjusted and inflation has been trended to the forecasting year, etc. The reason that we did not discuss those in here is because they are rather standard actuarial techniques practiced in most areas of rate-making and have been covered extensively elsewhere in the literature.[19]

The analysis shown above indicates that for many classes of distributions the new modified gamma approximation is superior in estimation accuracy and poses no significant increase in computation effort or expense. The new technique thus, is potentially valuable in more effective pricing of certain classes of reinsurance.

---

19 An alternate approach is to incorporate those effects into the parameters of distribution as suggested in [12] and [15].

Acknowledgements

The author would like to express his gratitude to Dr. Harry L. Shuford for his encouragement. Without his constant support, this project would not have reached the present stage of development. In addition, his comments on earlier drafts have influenced strongly the structure of this article. A conversation with David J. Grady, a Fellow of the C.A.S., led to the inclusion of the discussion of aggregate pricing as a separate section of the paper. Thanks also goes to the Financial Research and Planning Division of Fred. S. James & Co. and especially Dr. Tom C. Allen for providing the opportunity and resources to conduct the research project described in this paper.

Exhibit I

Generalized Poisson Model

Aggregate Probability $F(x)$ (%)

$z\left(=\frac{x-\mu}{\sigma}\right)$	γ = .5, Sev = Gamma				γ = 1, Sev = Inv. Normal				γ = 5, Sev = Gamma			
	A	MG	G	NP	A	MG	G	NP	A	MG	G	NP
-1.5	4.87	4.87	4.87	5.04	2.49	2.60	1.90	2.28	*	*	*	*
-1	15.58	15.58	15.56	15.87	14.17	14.21	14.29	15.87	*	*	*	*
- .5	32.93	33.09	33.06	33.28	34.69	34.79	35.28	36.16	*	*	*	*
0	53.14	53.33	53.33	53.30	56.33	56.36	56.65	56.45	78.49	78.47	78.52	71.44
.5	71.28	71.32	71.33	71.14	73.52	73.51	73.50	72.76	87.18	87.20	87.20	78.81
1	84.33	84.33	84.35	84.13	85.05	84.99	84.88	84.13	91.44	91.44	91.46	84.13
1.5	92.30	92.30	92.31	92.16	91.99	91.85	91.82	91.29	94.00	94.00	94.01	88.05
2	96.56	96.56	96.56	96.49	95.86	95.80	95.76	95.45	95.68	95.68	95.68	90.97
3	99.44	99.44	99.46	99.45	98.98	98.98	98.97	98.90	97.63	97.63	97.63	94.81
4	99.93	99.93	99.93	99.94	99.77	99.77	99.77	99.78	98.64	98.64	98.64	97.01
5	99.99	99.99	99.99	99.99	99.95	99.95	99.95	99.96	99.20	99.20	99.20	98.27
Frequency Mean	100				77.84				100.5			
Severity Coefficient of Variation	2.5				3				25			

Note: (*) Points below zero dollar limit.

(i) All four models are calculated by a HP-19 programmable calculator.

Exhibit II

Generalized Negative Binomial Model

Aggregate Probability $F(x)$ (%)

$z\left(=\frac{x-\mu}{\sigma}\right)$	$\gamma = 2$, Sev = Normal				$\gamma = 3$, Sev = Normal				$\gamma = 5$, Sev = Gamma			
	A	MG	G	NP	A	MG	G	NP	A	MG	G	NP
-1.5	*	*	*	*	*	*	*	*	*	*	*	*
-1	*	*	*	*	*	*	*	*	*	*	*	*
- .5	39.86	39.32	39.35	42.97	40.43	40.63	41.11	50.00	*	*	*	*
0	63.51	63.39	63.21	61.90	69.33	69.35	69.25	66.06	78.61	78.59	78.52	71.44
.5	78.03	78.10	77.69	75.16	81.57	81.49	81.47	76.79	87.22	87.21	87.20	78.81
1	86.71	86.77	86.47	84.13	88.37	88.42	88.29	84.13	91.46	91.46	91.46	84.13
1.5	91.95	92.01	91.79	90.04	92.47	92.52	92.40	89.18	94.01	94.01	94.01	88.05
2	95.13	95.09	95.02	93.84	95.05	95.08	94.99	92.64	95.67	95.67	95.68	90.97
3	98.33	98.32	98.17	97.72	97.79	97.78	97.75	96.63	97.63	97.63	97.63	94.81
4	99.40	99.40	99.30	99.19	98.99	99.00	98.96	98.47	98.64	98.64	98.64	97.01
5	99.78	99.78	99.75	99.72	99.53	99.53	99.51	99.31	99.20	99.20	99.20	98.27
Frequency Mean	90.25				111				100			
Severity Coefficient of Variation	2				2				25			

Note: (*) Points below zero dollar limit

(1) All four models are calculated by a HP-19 programmable calculator.

Exhibit Ia

Generalized Poisson Model

Variances of Modified Gamma, Gamma and NP vs Analytical Model

$Z(=\frac{x-\mu}{\sigma})$	γ = .5 MG/A	Sev = G/A	Gamma NP/A	γ = 1 MG/A	Sev = Inv.Normal G/A	NP/A	γ = 5 MG/A	Sev G/A	= Gamma NP/A
-1.5	0	0	.17	.11	-.59	-.21	x	x	x
-1	0	-.02	-.01	.04	.12	1.70	x	x	x
-.5	.16	.13	-.65	.10	.59	1.47	x	x	x
0	.19	.19	.16	.03	.32	.12	-.02	.03	-7.05
.5	.04	.05	-.14	-.01	-.02	-.74	.02	.02	-8.37
1	0	.02	-.20	-.06	-.17	-.92	0	.02	-7.31
1.5	0	.01	-.14	-.14	-.17	-.70	0	.01	-5.95
2	0	0	-.07	-.06	-.1	-.41	0	0	-4.71
3	0	.02	.01	0	-.01	-.08	0	0	-2.82
4	0	0	.01	0	0	.01	0	0	-1.53
5	0	0	0	0	0	.01	0	0	-.93
Variance vs Analytical Model	.006	.005	.051	.005	.080	.652	0	0	30.243

Exhibit IIa

Generalized Negative Binomial Model

Variances of Modified Gamma, Gamma and NP vs Analytical Model

$Z(=\frac{x-\mu}{\sigma})$	γ = 2 MG/A	Sev = Normal G/A	NP/A	γ = 3 MG/A	Sev = Normal G/A	NP/A	γ = 5 MG/A	Sev = Gamma G/A	NP/A
-1.5	x	x	x	x	x	x	x	x	x
-1	x	x	x	x	x	x	x	x	x
-.5	-.54	-.51	3.11	.20	.68	9.57	x	x	x
0	-.12	-.30	-1.61	.02	-.08	-3.27	-.02	-.09	-7.17
.5	.07	-.34	-2.87	-.08	-.10	-4.78	-.01	-.02	-8.41
1	.06	-.24	-2.58	.05	-.08	-4.24	0	0	-7.33
1.5	.06	-.16	-1.91	.05	-.07	-3.29	0	0	-5.96
2	-.04	-.11	-1.29	.03	-.06	-2.41	0	.01	-4.70
3	-.01	-.16	-.61	-.01	-.04	-1.16	0	0	-2.81
4	0	-.10	-.21	.01	-.03	-.52	0	0	-1.63
5	0	-.03	-.06	0	-.02	-.22	0	0	-.93
Variance vs Analytical Model	.036	.066	3.654	.006	.055	17.933	.000	.001	30.612

Sensitivity on the Selection of Severity Distribution

Aggregate Probability (%)

Frequency = Poisson

Frequency mean = 60.383

Severity coefficient of variation =4

$Z = \left(\frac{x-\mu}{\sigma}\right)$	Inv. Normal	log-normal
-1.5	0	.01
-1	10.71	.02
-.5	37.34	8.07
0	59.98	74.71
.5	75.66	92.52
1	85.64	94.92
1.5	91.70	96.23
2	95.28	97.09
3	98.52	98.15
4	99.55	98.77
5	99.87	99.15

Severity skewness	12	68
Aggregate skewness	1.5	9.02

<u>Increased Limited Factors</u>

(1)	(2)	(3)	(4)	(5)	(6)	(7)*	(8)*
$\frac{L-\mu}{\sigma}$	$F(L)$	$\frac{1}{\sigma}\int_0^L x\,dF(x)$	$\frac{1}{\sigma^2}\int_0^L x^2\,dF(x)$	$\mu_L/\sigma = (\frac{\mu}{\sigma})-(3)-(\frac{L}{\sigma})(1-(2))$	$\sigma_L^2/\sigma^2 = 1 + (\frac{\mu}{\sigma})^2 - (4) - 2(\frac{L}{\sigma})(5) + (\frac{L}{\sigma})^2(1-(2)) - (5)^2$	(5)+c(6)	$I(L) =$ (7)/.430403
-1.5	.048710	.284267	.186949	1.324071	11.745089	1.353434	3.14457
-1	.155801	.362889	.860435	1.061534	13.363123	1.094942	2.54399
-.5	.330885	.885834	2.425967	.680985	14.543811	.717345	1.66668
0	.533291	1.589031	4.873183	.397018	13.354180	.430403	1.00000
.5	.713208	2.302312	7.704674	.210377	10.460275	.236528	.54955
1	.843333	2.882251	10.292049	.101770	7.101659	.119524	.27770
1.5	.923029	3.276798	12.246940	.045233	4.248274	.055854	.12977
2	.965591	3.508574	13.509924	.018602	2.273409	.024286	.05643
3	.994601	3.685267	14.588247	.002567	.490539	.003793	.00881
4	.999351	3.718841	14.825890	.000284	.077647	.000478	.00111
5	.999937	3.723562	14.863940	.000026	.009605	.000050	.00012

(*) Parameters: Frequency mean = 100.551724, severity coefficient of variation =2.5, aggregate coefficient $\frac{\mu}{\sigma}$ =3.724138, laoding coefficient c= .0025*σ . Aggregate mean is selected as the stop - loss basic policy limit

Note: Columns (2), (3) and (4) are calculated by using extended simpson's rule with integration range [0,10] subdivided into 50 intervals.

References

[1] Ammeter, H. (1951): 'The Calculation of premium rates for excess of loss and stop loss reinsurance treaties' Trans. Int. Congr. Act. Scheveningen

[2] Beard, R.E., Pentikäinen, T. and Pesonen, E. (1969): 'Risk Theory', Methuen & Co., London

[3] Bickerstaff, D.R.: 'Automobile Collision deductibles and repair cost groups: The lognormal model' PCAS LIX p.68

[4] Bohman, H. and Esscher, F. (1963-64): 'Studies in risk theory with numerical illustrations concerning distribution functions and stop loss premiums I - II' Skand. Actu. Tidskr. 46, p. 173, 47, p.1

[5] Bühlmann, H. (1970): 'Mathematical methods in risk theory', Springer-Verlag

[6] Copson, E.T. (1955): 'An introduction of the theory of functions of a complex variable', Oxford Univ. Press.

[7] Cummins, J.D. and Freifelder, R. (1978): 'A comparative analysis of alternative maximum probable yearly aggregate loss estimators', J. of Risk and Insurance, p.27

[8] Dropkin, L.B. (1964): 'Size of loss distribution in workmen's compensation insurance' PCAS LI, p.198

[9] Fisz, M. (1963): 'Probability theory and mathematical statistics', John Wiley & Sons, Inc.

[10] Hewitt, C.C. (1966): 'Distribution by size of risk - A model', PCAS LIII, P. 106

[11] Hewitt, C.C. (1967): 'Loss ratio distribution - A model', PCAS LIV, P.70

[12] Hewitt, C.C. and Lefkowitz, B. (1978): 'Method for fitting distributions to insurance loss data' CAS Fall meeting

[13] Kauppi, L. and Ojantakanen, P. (1969): 'Approximation of the generalized poisson function', Astin Bulletin

[14] Mayerson, A.L., Jones, D.A. and Bowers, N.L. (1968): 'On the credibility of the pure premiums', PCAS LV, p. 175

[15] Miccolis, R.S. (1977): 'On the theory of increased limits and excess of loss pricing', PCAS LXIV, p.27

[16] Pentikäinen, T. (1977): 'On the approximation of the total amount of claims', Austin Bulletin, p. 281

[17] Seal, H.L. (1969): 'Stochartic theory of a risk Business', John Wiley and Sons

[18] Seal, H.L. (1977): 'Approximations to risk theory's by means of the gamma distribution', Austin Bulletin, p.47

[19] Seal, H.L. (1978): 'Survival probabilities', John Wiley and Sons

Appendix

A. Backgrounds on distributions listed in this paper.

A1. Function types

Table A1

Frequency Distributions

Type	Density Function $p(n)$	Range of parameters
Poisson	$e^{-\lambda}\lambda^n/n!$	$\lambda > 0$
Negative Binomial [20]	$(1-q)^{w}\binom{-w}{n}(-q)^{n}$	$w > 0,\ 0 < q < 1$

Table A2

Severity Distributions

Type	Notation	Cumulative Dist. Function	Range of parameters
Normal	$N(x;\mu,\sigma)$	$\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy$	$\mu,\ \sigma > 0$
Gamma	$G(x;b,p)$	$\frac{b^p}{\Gamma(p)}\int_0^x y^{p-1}e^{-by}dy$	$b,\ p > 0$
Inverse Normal	$I(x;a,b)$	$\frac{b}{\sqrt{\pi}}e^{2ab}\int_0^x e^{-a^2y-b^2/y}\,y^{-\frac{3}{2}}dy$	$a,\ b > 0$
Pareto	$P(x;c,m)$	$1-(1+x/c)^{-m}$	$c > 0,\ m \geq 2$
Log-Normal	$LN(x;d,u)$	$\frac{1}{\sqrt{2\pi}\,u}\int_0^x e^{-\frac{1}{2}\left(\frac{\log y-d}{u}\right)^2}y^{-1}dy$	$u,\ d > 0$
Log-Gamma	$LG(x;a,v)$	$\frac{a^v}{\Gamma(v)}\int_1^x(\log y)^{v-1}y^{-(a+1)}dy$	$v > 0,\ a > 2$

A2. Characteristic Functions

A powerful feature in the study of distribution functions and their moments is the characteristc function, $\varphi_F(t)$, associated with a given distribution function F. which is defined as [21]

20. The parameters are given by: $q = (\sigma_p^2-\lambda)/\sigma_p^2,\ w = \lambda^2/(\sigma_p^2-\lambda)$
see [9] p.167

21. See [9] Chap. 4

$$\varphi_F(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

where $i = \sqrt{-1}$ is the imaginary number. The overwhelming advantages of using the characteristic function is evident from the following:

(1) A moment generating function is defined over real number; the characteristic function is its complex analogue. It retains all the desirable properties of the moment generating function and unlike the moment generating function, it always exists;

(2) a standard mathematical technique known as the Laplace transformation (or Fourier transformation) asserts that as long as the characteristic function is known, one can rediscover the associated distribution function. This invertible property (not valid for moment generating function) offers an algorithm to compute the aggregate loss probability directly.

Without using these two features of characteristic functions, the derivation of our formula for the new model would be virtually impossible.

Among the six severity distribution functions listed in Table A2, only the first three have an explicit form for their characteristic function. The last three do not admit any closed form for the characteristic functions. We will derive the characteristic function of the inverse normal distribution here, and leave that of the normal and gamma to the interested reader.[22]

Letting the variable $y = z^{-2}$ in the c.d. f of inverse normal (Table A2), we have

$$I(x; a, b) = \frac{2a}{\sqrt{\pi}} e^{2ab} \int_{\frac{1}{\sqrt{x}}}^{\infty} e^{-b^2 z^2 - a^2/z^2} dz \qquad (10)$$

22. See [9] pp. 147 and 152

Now, observe that $2b\,dz = d(bz+a/z) + d(bz-a/z)$, thus

$$I(x;a,b) = \frac{e^{2ab}}{\sqrt{\pi}}\left\{\int_{1/\sqrt{x}}^{\infty} e^{-b^2z^2-a^2/z^2}\,d(bz+a/z) + \int_{1/\sqrt{x}}^{\infty} e^{-b^2z^2-a^2/z^2}\,d(bz-a/z)\right\}$$

next setting variable $y = bz + a/z$ in the first integral and $y = bz - a/z$ in the second, it follows that

$$I(x;a,b) = \frac{1}{\sqrt{\pi}}\left\{e^{4ab}\int_{b\sqrt{x}+a/\sqrt{x}}^{\infty} e^{-y^2}dy + \int_{b\sqrt{x}-a/\sqrt{x}}^{\infty} e^{-y^2}dy\right\}$$

(change $y = -z/\sqrt{2}$)

$$= \frac{1}{\sqrt{\pi}}\left\{e^{4ab}\int_{-\infty}^{-\sqrt{2x}\,b-a\sqrt{2/x}} + \int_{-\infty}^{a\sqrt{2/x}-b\sqrt{2x}}\right\} e^{-z^2/2}\,dz$$

Thus, we have a practical form for the inverse normal distribution expressed in terms of the normal distribution:

$$I(x;a,b) = e^{4ab}N\left(-b\sqrt{2x}-a\sqrt{2/x};0,1\right) + N\left(a\sqrt{2/x}-b\sqrt{2x};0,1\right) \qquad (11)$$

The calculation of the characteristic function follows closely the approach which led to the derivation of formula (11). In fact, by definition,

$$\varphi_{IN}(t) = \frac{b}{\sqrt{\pi}}e^{2ab}\int_0^{\infty} e^{-(a^2-it)/y - b^2y}\, y^{-3/2}\,dy$$

(change $z = y^{-1/2}$)

$$= \frac{2b}{\sqrt{\pi}}e^{2ab}\int_0^{\infty} e^{-(a^2-it)/z^2 - b^2z^2}\,dz$$

$$= e^{2b(a-\sqrt{a^2-it})}\left(\frac{2b}{\sqrt{\pi}}e^{2\sqrt{a^2-it}\,b}\int_0^{\infty} e^{-(a^2-it)/z^2-b^2z^2}\,dz\right)$$

Compare the form inside the parenthesis with the r.h.s of (10), it is identical to $I(\infty;\sqrt{a^2-it}, b)$[23], taking the fact that cumulative probability is always equal to 1 when the argument tends to infinity, then

$$\varphi_{IN}(t) = \exp\left(2b(a-\sqrt{a^2-it})\right) \qquad (12)$$

The comparable results for the normal and gamma distributions are given by

$$\begin{cases} \varphi_N(t) = \exp(it\mu - \frac{1}{2}\sigma^2t^2), & \text{for normal} \\ \varphi_G(t) = (1-it/b)^{-p}, & \text{for gamma} \end{cases} \qquad (13)$$

23. Since parameters of distribution have to be real numbers, while, here we have complex numbers involved, it is thus confusing to use same notation. However, in this particular case (and except this ambiguous notation) the property of distribution still holds in the extended situation, see M. Abramowitz and I.A. Stegan: "Handbook of Mathematical Functions" National Bureau of Standards, formula 7.4.3 p. 302

A3. Background information on the Inverse Normal Distribution

An immediate consequence of deriving the characteristic function is that one can readily determine the cumulants of a givwn distribution function. Since, by definition, log $\varphi_{..}(t)$ can be formally expanded as follows:[24]

$$\log \varphi_S(t) = \mu_s(it) + \tfrac{1}{2}\sigma_s^2(it)^2 + \tfrac{1}{6}\gamma_s\sigma_s^3(it)^3 + \cdots \qquad (14)$$

where μ_s, σ_s^2 and γ_s are the mean, variance and skewness of S. Now, apply (14) to inverse normal distribution, we have

$$\begin{aligned}\log \varphi_{IN}(t) &= 2ab\,(1-(1-it/a)^{\frac{1}{2}})\\ &= 2ab\left(\tfrac{1}{2}(\tfrac{it}{a}) + \tfrac{1}{8}(\tfrac{it}{a})^2 + \tfrac{1}{16}(\tfrac{it}{a})^3 + \cdots\right)\end{aligned}$$

From a comparison with the right hand side of (14), it can be deduced that

$$\mu_s = b, \quad \sigma_s^2 = \tfrac{1}{2}(\tfrac{b}{a}), \quad \gamma_s\sigma_s^3 = \tfrac{3}{4}(\tfrac{b}{a^2})$$

Next, solving the first two equations for a and b, and placing the results in the last equation, it can be seen that:

$$a = \tfrac{1}{2}(\mu_s/\sigma_s^2), \quad b = \mu_s, \quad \gamma_s = 3(\sigma_s/\mu_s) \qquad (15)$$

A4. The skewness of severity distribution

The last equation of (11) proved the case of inverse normal distribution stated in Table 3. The case of the normal is quite straight forward, since:

$$\log \varphi_N(t) = it\mu - \tfrac{1}{2}\sigma^2 t^2$$

24. See [9] formula (4.3.3) p. 111, note that in [9], the term semi-invariants, instead of cumulants, is used.

thus $\gamma_\Delta \sigma_\Delta^3 = 0$ i.e., $\gamma_\Delta = 0$. As for the gamma distribution:

$$\log \varphi_G(t) = -p \log(1 - it/b) = \frac{p}{b}(it) + \frac{p}{2b^2}(it)^2 + \frac{p}{3b^3}(it)^3 + \cdots$$

or $\mu_\Delta = p/b$, $\sigma_\Delta^2 = p/b^2$ and $\gamma_\Delta \sigma_\Delta^3 = 2p/b^3$. It follows that

$$\gamma_\Delta = 2\left(\frac{\sigma_\Delta}{\mu_\Delta}\right)$$

which completes the case for the gamma distribution.

For the other three types, it is necessary to use an alternative definition of skewness, which (if it exists) is [25]:

$$\gamma_\Delta = E\left[\frac{X - \mu_\Delta}{\sigma_\Delta}\right]^3 = \sigma_\Delta^{-3}\left(E[X^3] - 3\mu_\Delta E[X^2] + 2\mu_\Delta^3\right) \qquad (16)$$

In carrying out the calculation of the first three moments of the pareto, log-normal and log-gamma distribution, we have the following table:

Table A3

Type	$E(X)$	$E(X^2)$	$E(X^3)$
Pareto	$c/(m-1)$	$2c^2/((m-1)(m-2))$	$6c^3/((m-1)(m-2)(m-3))$
log-normal	$\exp(d + \frac{u^2}{2})$	$\exp(2d + 2u^2)$	$\exp(3d + \frac{9}{2}u^2)$
log-gamma	$(1 - 1/a)^{-v}$	$(1 - 2/a)^{-v}$	$(1 - 3/a)^{-v}$

where in order to ensure the existence of the integration in the derivation of the 3rd moment, the following conditions have to be satisfied:

$$\begin{cases} m > 3, & \text{in pareto case} \\ a > 3, & \text{in log-gamma case} \end{cases} \qquad (17)$$

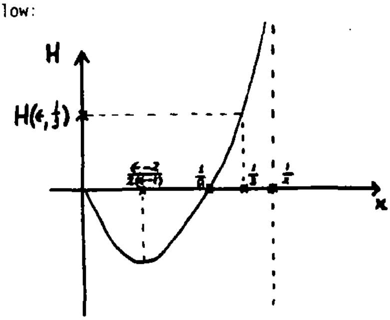
25. See [9] p. 73

Bysolving the first two columns of Table A3 for parameters in each case and substituting the results in the last column and in formula (16), we have the following table:

Table A4

Type	Parameters	Skewness
Pareto	$c=\mu_\rho(\mu_\rho^2+\sigma_\rho^2)/(\sigma_\rho^2-\mu_\rho^2),\ m=2\sigma_\rho^2/(\sigma_\rho^2-\mu_\rho^2)$	$2\left(\frac{\sigma_\rho}{\mu_\rho}\right)\left(\frac{3\sigma_\rho^2-\mu_\rho^2}{3\mu_\rho^2-\sigma_\rho^2}\right)$
log-normal	$u=\sqrt{\log((\sigma_\rho^2+\mu_\rho^2)/\mu_\rho^2)},\ d=\log(\mu_\rho^2/\sqrt{\mu_\rho^2+\sigma_\rho^2})$	$\left(\frac{\sigma_\rho}{\mu_\rho}\right)\left(\left(\frac{\sigma_\rho}{\mu_\rho}\right)^2+3\right)$
log-gamma[26]	$a,\ v=-\log\mu_\rho/\log(1-1/a)$	$\sigma_\rho^{-3}\left((1-3/a)^{-v}-3\mu_\rho\sigma_\rho^2-\mu_\rho^3\right)$

By using an explicit value for m in (17), the assertion of the footnote 14 for the pareto case is established. For the log-gamma case, since $(1/a)$ is the root of $H(\epsilon,x)$ (see footnote 26) and the graph of $H(\epsilon,x)$ can be portrayed as follow:



where $(\epsilon-2)/2(\epsilon-1)$ is a local minimum point of $H(\epsilon,x)$. Thus the requirement of (17) asserts that $(1/a)<\frac{1}{3}$, which is equivalent to $H(\epsilon,1/a)>0$, or

$$\epsilon \leq \frac{\log 3}{\log 3-\log 2} \doteq 2.7095$$

We thus prove the last assertion of footnote 14.

26. Since from the first two columns in Table A3, we have $\log\mu_\rho=-v\log(1-1/a)$, $\log(\sigma_\rho^2+\mu_\rho^2)=-v\log(1-2/a)$, hence $(1/a)$ is the solution of the following equation

$$H(\epsilon,x)=\log(1-2x)-\epsilon\log(1-x)=0, \quad \text{where } \epsilon=\log(\sigma_\rho^2+\mu_\rho^2)/\log\mu_\rho$$

B. Analytical Model

A fundamental equation in collective risk theory demonstrates that the aggregate cumulative distribution function F(z) of annual aggregate loss less than or equal to Z, is given by [27]

$$F(z) = \sum_{n=0}^{\infty} p(n)\, S^{n*}(z) \tag{18}$$

where $S^{n*}(z)$ is the nth convolution of S or, equivalently, the cumulative distribution of exactly n claims with total loss less than or equal to Z and $p(n)$ is the frequency density function as listed in Table A1.

Formula (18) has practical value only when the characteristic function of S has a closed form, so that the precise form of $S^{n*}(z)$ can be derived. Among the severity functions in Table A2, only the first three meet this condition. For the rest three which do not admit a closed form for the characteristic function an alternative numerical technique has to be devised to calculate their characteristic functions, this would cause the whole computation not only time consuming but also, sometimes, very messy.

In the case of normal, gamma and inverse normal, where their characteristic functions are known, it is possible to use the following two fundamental properties of characteristic function:

(i) the characteristic function of the convolution of two functions is the product of their respective characteristic functions;

(ii) if two distribution functions possess identical characteristic functions, then the distribution functions are equal,

we can derive the explicit form for the $S^{n*}(z)$ as shown in the following table:

27. For an expository treatment of collective risk theory, please see [2] and [17].

Type	(1) S(Z)	(2) $\varphi_S(t)$	(3) $\varphi_S^n(t)\left(=\varphi_{S^{n*}}(t)\right)$	(4) $S^{n*}(Z)$
Normal	$N(z;\mu,\sigma)$	$\exp\left(it\mu-\frac{1}{2}\sigma^2t^2\right)$	$\exp\left(itn\mu-\frac{1}{2}n\sigma^2t^2\right)$	$N\left(z;n\mu,\sqrt{n}\sigma\right)$
Gamma	$G(z;b,p)$	$(1-it/b)^{-p}$	$(1-it/b)^{-np}$	$G(z;b,\sqrt{n}p)$
Inverse Normal	$I(z;a,b)$	$\exp\left(2b\left(a-\sqrt{a^2-it}\right)\right)$	$\exp\left(2nb\left(a-\sqrt{a^2-it}\right)\right)$	$I(z;a,nb)$

In column (3), property (i) is used and in column (4), property (ii) is used.

C. Gamma Approximation

C1. The derivation of the gamma approximation

The gamma distribution $G(x;b,p)$ has only two parameters which are determined by the first two sample moments. If we add one more parameter α to the function, $G(x+\alpha;b,p)$, then the third moment is required to estimate the parameters. This procedure is called the gamma approximation.

To specify parameters x, b and p, one sets up three equations for the first two moment and skewness, then solves them for x, b and p. To do this let us first calculate the characteristic function of $G(x+\alpha;b,p)$. Since the density function of $G(x+\alpha;b,p)$ is

$$\frac{d}{dx}G(x+\alpha;b,p) = \frac{b^p}{\Gamma(p)}(x+\alpha)^{p-1}e^{-b(x+\alpha)}, \qquad -\alpha \le x < \infty$$

then the characteristic function is defined by

$$\varphi_S(t) = \frac{b^p}{\Gamma(p)}\int_{-\alpha}^{\infty} e^{itx}(x+\alpha)^{p-1}e^{-b(x+\alpha)}dx$$

(letting y = x + α)

$$= \frac{b^p}{\Gamma(p)}\int_0^{\infty} e^{it(y-\alpha)}y^{p-1}e^{-by}\,dy$$

$$= e^{-it\alpha}(1-it/b)^{-p} \qquad (19)$$

take logrithm both sides and expand the left hand side into power series of (it), we have

$$\log \varphi_S(t) = -\alpha(it) + p\left(\left(\frac{it}{b}\right) + \frac{1}{2}\left(\frac{it}{b}\right)^2 + \frac{1}{3}\left(\frac{it}{b}\right)^3 + \cdots\right),$$

compare the coefficients of the three lowest terms of (it), we set up three equations:

$$\begin{cases} \mu_p = p/b - \alpha \\ \sigma_p^2 = p/b^2 \\ \gamma_p \sigma_p^3 = 2p/b^3 \end{cases}$$

solve them for x, b and p, we have

$$b = 2/(\sigma_p \gamma_p), \quad p = (2/\gamma_p)^2, \quad \alpha = 2/(\sigma_p \gamma_p) - \mu_p \qquad (20)$$

Therefore the gamma approximation is expressed as

$$G(x+\alpha; b, p) = \frac{b^p}{\Gamma(p)} \int_0^{x-\mu_p+2/(\sigma_p\gamma_p)} y^{(\frac{2}{\gamma_p})^2-1} e^{-\left(\frac{2y}{\sigma_p\gamma_p}\right)} dy$$

(letting $z = 2y/(\gamma_p \sigma_p)$)

$$= G\left(\left(\frac{2}{\gamma_p}\right)^2 + \left(\frac{2}{\gamma_p}\right)\left(\frac{x-\mu_p}{\sigma_p}\right); 1, \left(\frac{2}{\gamma_p}\right)^2\right) \qquad (21)$$

C2. Aggregate skewness

In applying NP or gamma approximation, one needs the input of aggregate mean, standard deviation and skewness. In this section we are going to derive the aggregate skewness for both generalized poisson and negative binomial models.

As usual we first calculate the characteristic function of either model. Taking the fact that characteristic function of S^{n*} is the product of n characteristic functions of S, together with formula (18) and the explicit form of p(n) in Table A1, it is not difficult to see that the characteristic function of the aggregate distribution function F is given by.

$$\varphi_F(t) = \begin{cases} \exp(\lambda(\varphi_S(t) - 1)), & \text{for poisson} \\ \left(\dfrac{1 - q\varphi_S(t)}{1-q}\right)^{-w}, & \text{for negative binomial} \end{cases} \qquad (22)$$

Where $\varphi_S(t) = 1 + \mu_\Delta (it) + \mu_{\Delta,2}/2\,(it)^2 + \mu_{\Delta,3}/6\,(it)^3 + \cdots$ is the characteristic function of the severity distribution S. Taking the logrithm both sides, it becomes

$$\log \varphi_F(t) = \begin{cases} \lambda(\varphi_S(t) - 1), & \text{for poisson} \\ -n \log\left(1 - \frac{q}{1-q}(\varphi_S(t) - 1)\right), & \text{for negative binomial} \end{cases} \tag{23}$$

Identifying the coefficients of $(it)^3$ both sides, we can evaluate [28] the aggregate skewness. They are given as follows:

Table A5

Aggregate Skewness

Generalized poisson model	Generalized negative binomial model
$\dfrac{\mu_{\Delta,3}}{\sqrt{\lambda}\,(\mu_{\Delta,2})^{3/2}}$	$\dfrac{2(\sigma_p^2-\lambda)^2\mu_\Delta^3/\lambda + 3(\sigma_p^2-\lambda)\mu_{\Delta,2}\cdot\mu_\Delta + \lambda\mu_{\Delta,3}}{(\sigma_\Delta^2\mu_\Delta^2 + \lambda\,\sigma_\Delta^2)^{3/2}}$

Table A5 is a general formula for aggregate skewness and does not use the precise form of severity statistics. If the individual severity type is incorporated the formula for skewness would bear the following form.

28. This is done by expanding $\varphi_S(t)$ into series of (it), and using footnote 20 for the negative binomial model.

Table A6

Aggregate Skewness

Severity/Frequency	Poisson	Negative binomial
normal	$(1+3c_\Delta^2)/\sqrt{\lambda(1+c_\Delta^2)^3}$	$\{\lambda(1+3c_\Delta^2)+A\}/B$
gamma	$(1+2c_\Delta)^2/\sqrt{\lambda(1+c_\Delta^2)}$	$\{\lambda(1+c_\Delta^2)(1+2c_\Delta)^2+A\}/B$
inverse normal	$(3c_\Delta^4+3c_\Delta^2+1)/\sqrt{\lambda(1+c_\Delta^2)^3}$	$\{\lambda(3c_\Delta^4+3c_\Delta^2+1)+A\}/B$
pareto	$3\sqrt{1+c_\Delta^2}/((3-c_\Delta^2)\sqrt{\lambda})$	$\{3\lambda(1+c_\Delta^2)/(3-c_\Delta^2)+A\}/B$
log-normal	$(1+c_\Delta^2)^{1.5}/\sqrt{\lambda}$	$\{\lambda(1+c_\Delta^2)^3+A\}/B$
log-gamma[29]	$((1-3/a)/(1-2/a))^{-v}/\sqrt{\lambda}$	$\dfrac{\lambda(1-3/a)^{-v}+3(\sigma_p^2-\lambda)\mu_p\mu_{\Delta 2}+2(\sigma_p^2-\lambda)^2\mu_\Delta^3/\lambda}{(\sigma_p^2\mu_\Delta^2+\lambda\sigma_\Delta^2)^{3/2}}$

where $c_\Delta=\sigma_\Delta/\mu_\Delta$ is the severity coefficient of vatiation,

$A=3(\sigma_p^2-\lambda)(1+c_\Delta^2)+2(\sigma_p^2-\lambda)^2/\lambda$ and $B=(\sigma_p^2+\lambda c_\Delta^2)^{3/2}$

D. Modified Gamma Approximation

D1. Derivation fo the new model

The approach we adopt for the new bodel is that, first we use the gamma approximation technique to match the selected severity distribution (one of the six types in Table A2) by identifying the statistics $\mu_\Delta, \sigma_\Delta$ and γ_Δ with those of the slected type, then utilize the following well known formula[30]

$$F(x) = \frac{1}{2} - \frac{1}{2\pi i}\int_{-\infty}^{\infty} \varphi(t)\, e^{-itx}\, t^{-1}\, dt \qquad (24)$$

to invert the aggregate distribution function from its characteristic function. Thus what we have to do here is to substitute $\varphi_F(t)$ in (24) by (22), then simplify it to the form given in (1), Table 1 and Table 2.

29. See footnote 26 for a and v.

30. This formula uses property (ii) discussed in section A. For detailed information on Laplace transformation please see [17].

Now, let $f(t)$ be the modules (or absolute value) of $\varphi_F(t/\sigma)$ and $-g(t)$ be the arguments of $\varphi_F(\frac{t}{\sigma})$, then [31]

$$\varphi_F(t/\sigma) = f(t)\, e^{-i g(t)} \tag{25}$$

where both f(t) and g(t) are real numbers. Replace $s = t\sigma$ in(24) and replace $\varphi_F(\frac{s}{\sigma})$ by (25), it turns out that

$$F(x) = \tfrac{1}{2} - \frac{1}{2\pi i}\left(\int_0^{\infty} + \int_{-\infty}^{0}\right) f(s)\, e^{-i(g(s) + sx/\sigma)} s^{-1}\, ds$$

(Change the 2nd integral by s =-t)

$$= \frac{1}{2} - \frac{1}{2\pi i}\left(\int_0^{\infty} f(t)\left(e^{-i(g(t)+tx/\sigma)} - e^{i(g(t)+tx/\sigma)}\right)\frac{dt}{t}\right)^{32}$$

(taking the fact that $e^{iy} - e^{-iy} = 2i\,\mathrm{Sin}(y)$)

$$= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} f(t)\,\frac{\mathrm{Sin}(tx/\sigma + g(t))}{t}\,dt$$

which is the form given in (1). The next step is to find out f(t) and g(t) in either models. To continue our derivation, let us decompose $\varphi_S(\frac{t}{\sigma})-1$ into two parts:

$$\varphi_S(t/\sigma) - 1 = k(t) + i\,\bar{k}(t) \tag{26}$$

then identify (25) with (22) via (26), we have

$$f(t)e^{-ig(t)} = \begin{cases} e^{\lambda k(t)} \cdot e^{i\lambda \bar{k}(t)}, & \text{for poisson} \\ \left(1 - \left(\frac{q}{1-q}\right)k(t) - i\left(\frac{q}{1-q}\right)\bar{k}(t)\right)^{-\alpha}, & \text{for negative binomial} \end{cases} \tag{27}$$

The case of the generalized poisson model in Table 1 is obvious from (27).

---

31. An complex number can be expressed by its modulus and argument, see [6] p.6

32. Here we take the following fact that

$$\varphi_F(-t) = \int_{-\infty}^{\infty} \cos tx\, dF(x) - i\int_{-\infty}^{\infty} \sin tx\, dF(x) = \overline{\varphi_F(t)}$$ (the conjugate of $\varphi_F(t)$)

to demonstrate that $f(-t) = f(t)$ and $g(-t) = -g(t)$, and ultilize in our derivation, for conjugate number see [6].

For the generalized negative binomial, by the definition of modulus and argument[33], we find that

$$\begin{cases} f(t) = \left\{ \left(1 - \left(\frac{g}{1-g}\right) h(t)\right)^2 + \left(\frac{g}{1-g}\right)^2 h(t)^2 \right\}^{-\frac{w}{2}} \\ g(t) = w \tan^{-1}\left(\frac{\lambda g\, h(t)/(1-g)}{\lambda - \lambda g\, h(t)/(1-g)} \right), \end{cases}$$

last item that has to be verified is Table 2. This is straight-forward, since by (19) and (20), we have

$$\varphi_S(t/\sigma) = \delta(t) e^{i\theta(t)} = \delta(t)\left(\cos\theta(t) + i \sin\theta(t)\right) \tag{28}$$

where

$$\begin{cases} \delta(t) = \text{absolute values of } \quad (1 - it\gamma_0\sigma_0/2)^{-\left(\frac{2}{\gamma_0}\right)^2} \\ \theta(t) = t\left(\frac{\mu_0 - 2\sigma_0/\gamma_0}{\sigma}\right) + \text{argument of } \quad -\left(\frac{2}{\gamma_0}\right)^2 \log(1 - it\gamma_0\sigma_0/2) \end{cases} \tag{29}$$

Comparing both sides of (26) and (28), and exploring the left hand side of (29), we come to the results of Table 2

D2. Formula (1) via Extended Simpson's Rule

The extended Simpson's rule is adequate to handle the numerical integration of formula (1) and Table 5. Since integration over an infinite interval is practically impossible, it is necessary to integrate over a truncated interval. The limit, R, of the range of integration has to be determined. Also the size, h, of the equally divided subinterval has to be chosen in ultilizing Simpson's rule.

---

33. See [6] p.5-7

34. For a detailed treatment of this, please see Stephen G. Kellison: 'Fundamentals of Numerical Analysis' Richard D. Irwin. Inc. 1975, Chap. 8

If a precision up to the sixth decimal point is required, h=.2 would be satisfactory. The selection for R would be more complicate. A quick and practical way to select the appropriate R is to input t until the value of

$$f(t) \, Sin\left(tx/\sigma + g(t)\right)/(\pi t)$$

is less than, say, 10^{-6}. Choose that t for R. Generally, the appropriate value of R would fall in to the range from 10 to 100, depend on γ_p and λ .

TITLE: PRICING EXCESS-OF-LOSS CASUALTY WORKING COVER REINSURANCE TREATIES

AUTHORS: Mr. Gary Patrik

Mr. Patrik is an Actuarial Director of Prudential Reinsurance Company. He received his FCAS in 1978, and earned an M.S. degree from Northwestern University as well as an M. S. degree from the Graduate School of Business Administration of New York University. Gary is currently, on the CAS Committee on Theory of Risk, the CAS Education and Examination Committee and the AERF Loss Distribution Textbook Selection Committee.

Mr. Russell John

Mr. John is a Senior Actuarial Associate with Prudential Reinsurance Company. He receives his ACAS in May 1980 and holds a PhD degree in Mathematics from Rutgers University.

REVIEWER: Mr. Jerry Miccolis

Mr. Miccolis is Vice President for Tillinghast, Nelson and Warren Inc. He has a B. S. degree in Mathematics from Drexel University in Philadelphia, and received his FCAS in 1979. Jerry is also a member of the American Academy of Actuaries and serves the CAS on the Committee on Theory of Risk as well as the Examination Committee.

I. INTRODUCTION

An excess-of-loss reinsurance treaty provides the primary insurance company (cedant) with reinsurance protection covering a certain layer of loss for a specified category of individual (direct) insurance policies. Hence, for each loss event (occurrence) coming within the terms of the treaty, the reinsurer reimburses the cedant for the dollars of loss in excess of a certain fixed retention up to some maximum amount of liability per occurrence. For example, if the cedant's retention is \$100,000 and the reinsurer's limit of liability is \$400,000, then the reinsurer covers losses in the layer \$100,000 up to \$500,000; in reinsurance terminology, this is the layer \$400,000 excess of \$100,000. The reimbursement generally takes place at the time that the cedant reimburses the injured party. Allocated loss adjustment expenses are usually shared pro rata according to the loss shares, although in a few treaties they may be included in with the loss amounts before the retention and reinsurance limit are applied.

In this paper, casualty coverage will mean either third party liability coverage or worker's compensation coverage, although on certain treaties it may be broader. For example, for automobile insurance, first party coverage may be included within the terms of the excess treaty along with the third party coverage;

in any case, the total loss covered per occurrence is added together before application of the retention and the reinsurer's limit.

A working cover is a treaty on which the reinsurer expects to pay some losses; reinsurance underwriters say that the cover is substantially exposed by the primary insurance policy limits. Typically, layers below $1,000,000 per occurrence for casualty coverage are considered to be working covers. For a more complete discussion of this coverage, see Reinarz (1969), The Insurance Institute of London (1976) or Barile (1978).

An excess-of-loss casualty working cover is typically a large, risky contract. The annual reinsurance premium is usually six figures and quite often is millions of dollars. Although losses are expected, the number of losses to the treaty and their sizes are highly uncertain. Each cedant's insurance portfolio is unique, so there are no simple standard reinsurance rates. Industrywide average increased limits factors might be used as a starting point for pricing; however, competition and uncertainty force the reinsurer to be more sophisticated in his analysis of each proposal. A further complication is that the reinsurer usually has much less information to work with than does his primary insurance colleague. The reinsurer is provided with often vague and incomplete estimates of past and future exposure, of underlying coverage, of aggregate ground-up direct losses, and with some details about the very few historical large losses which are known. The final price will be reached by competitive bidding and by negotiation over particular contract terms. To compete, the reinsurer must work within severe time and manpower constraints to estimate a price which he believes to be adequate and which he can justify to the cedant.

Pricing excess-of-loss casualty working covers with any degree of accuracy is a complex and difficult underwriting and actuarial problem. We believe that the general theoretical pricing problem will remain insolvable: there will always be more questions than there are answers. However, in the spirit of a "Call for Papers", we offer a progress report on our work to date, knowing that we have only the beginnings of a truly satisfying practical solution. We will illustrate the actuarial problem by pricing two relatively simple and representative treaties. The approach is mathematical/actuarial; underwriting considerations are only briefly and incompletely mentioned, although these are very important. Some general solution criteria are presented and some tentative partial solutions are discussed. Although the point of view is that of a reinsurance actuary, we believe that the general approach may be of interest to other actuaries and that some of the particular techniques will be immediately useful to our primary insurance colleagues.

Any complicated procedure such as the one presented in this paper develops over time from the work and ideas of many people.

We wish to acknowledge the help of a few who have contributed to this development: Ralph Cellars, Howard Friedman, Charles Hachemeister, Mark Kleiman, Stephen Orlich, James Stanard and Edward Weissner.

II. TWO TREATY PROPOSALS

Reinsurers often receive proposals for which historical data are virtually non-existent. Such is the case when a newly formed or an about-to-be-formed primary company seeks reinsurance coverage or when an existing company writes a new insurance line or a new territory. There may be some vaguely analogous historical data, general industry information and some underwriting guesses about next year's primary exposure, coverage, rates and gross premium. An example is that of a new doctors' mutual offering professional liability coverage to the members of the medical society in state A.

Example A: A Doctors' Mutual Insurance Company

Proposal

1. reinsured layer: \$750,000 excess of \$250,000 per occurrence; no annual aggregate reinsurance limit; allocated loss adjustment expense shared pro rata according to loss share.

2. underlying coverage: professional liability claims-made coverage for limits of \$1,000,000/\$3,000,000 per claimant/ annual aggregate per doctor using the standard ISO policy form.
3. coverage period: beginning July 1, 1980 and continuous until terminated.
4. reinsurance rate: the offer is 25% of the gross direct earned premium with a 20% ceding commission and brokerage fee (thus, the net rate is 20%).

Information

5. exposure estimate of 500 doctors; no class breakdown.
6. class definitions - identical with ISO classes.
7. list of claims-made rates to be charged by doctor class for \$1M/\$3M limits.
8. summary of calendar/accident year 1974 - 1978 aggregate known losses and earned premiums for state A doctors covered by the BIG Insurance Company.
9. details about the five known losses paid or presently reserved for more than \$100,000 in state A for accident years 1974 - 1978.

10. a booklet describing the organization and financial structure of the doctors' mutual, together with biographies of the principal managers, claims-persons and attorneys and a statement of a get-tough attitude toward defending professional liability claims.

11. other miscellaneous letters and memos stating why this is an especially attractive deal for the reinsurer and the doctors.

It should be apparent that most of this information is only indirectly useful for pricing the reinsurance coverage. The offered rate must be analyzed using analogous industry information. There is great uncertainty regarding the potential loss situation.

At the opposite extreme is the treaty proposal for which there is a great wealth of historical information. This is sometimes the case when a treaty has been in place for many years with only minor changes, such as increasing the primary retention over time to parallel the inflation in individual loss amounts. If a reinsurer has been on the treaty for a few years, his underwriting and claims-persons have gotten to know the primary company people and have audited the treaty accounts. Thus, there is less uncertainty regarding the potential loss situation. A much simplified example of this situation is considered (only one line of business).

Example B: P&C Insurance Company

Proposal

1. reinsured layer: $400,000 excess of $100,000 per occurrence; no annual aggregate reinsurance limit; allocated loss adjustment expense shared pro rata according to loss share.
2. underlying coverage: general liability premises/operations coverage, mainly in state B, written at various limits for bodily injury and property damage liability.
3. coverage period: beginning January 1, 1980 and continuous until terminated.
4. reinsurance rate: the net rate is to be negotiated as a percentage of gross direct earned premium.

Information

5. estimate of 1980 gross direct earned premium.
6. estimate of 1980 premium by policy limit.
7. summary of calendar/accident year 1969 - 1978 aggregate known losses as of 6/30/79 and gross earned premiums for P&C's general liability coverage insurance portfolio.

8. list of rate changes and effective dates for this line of business for 1969 through present and information that no change is contemplated through 1980.

9. detailed listings of all 358 general liability losses occurring since 1969 which were valued greater than $25,000 as of 6/30/75, 6/30/76, . . . or 6/30/79. At each evaluation, the information listed for each loss includes the following:

    a) identification number

    b) accident year (occurrence)

    c) amount of loss paid

    d) amount of loss outstanding

    e) policy limits

The evaluation of these two treaty proposals will illustrate the pricing procedure. Note that for example A we are to evaluate an offered rate, while for example B we are to propose a net rate and negotiate.

Before proceeding with the details, we believe it necessary to discuss some general pricing philosophy.

III. PRICING PHILOSOPHY

An insurance contract may be thought of as a financial stochastic process - a random pattern of pay-ins and payouts over time. The financial repercussions of a casualty excess-of-loss treaty may continue for 20 years or more. Thus, a reinsurer must consider the many aspects of this financial process to be able to estimate prices which are reasonably consistent with broad corporate policy. An actuarial goal is to combine all the contract financial parameters and all the corporate (underwriting) decision-making criteria into one comprehensive premium calculation principle or function - a black box which for each particular treaty produces the final premium or, more realistically, a negotiable premium range. Such a black box will not be purely mathematical, but will require substantial subjective input.

Present actuarial knowledge is short of this utopian goal. However, actuaries and underwriters have identified certain major contract parameters and decision-making criteria which should be considered when evaluating a particular contract. See Pratt (1964), Reinarz (1969), Bühlmann (1970), Gerber (1974) and Freifelder (1976) among others for discussions of premium calculation principles.

We believe that a reinsurer should consider the following items for each treaty either explicitly or implicitly:

1. The potential distribution of the aggregate loss to be ultimately paid by the reinsurer. Although the whole (past and future) coverage period should be considered, most important is the potential distribution of the aggregate loss arising from the next coverage year. The potential distribution of the aggregate loss is based upon the reinsurer's subjective evaluation of the situation and is difficult to specify in detail. Consequently, only certain major characteristics are estimated, such as the expected value, the variance or standard deviation, and certain percentiles, such as the 90th, 95th and 99th.

2. The potential distribution of the cash flow. The overall pattern over time is of interest, but more easily understood is the present value of the cash flow generated by the next coverage year. This random variable is distributed according to various price assumptions and the reinsurer's subjective assessment of the potential distributions of aggregate loss, payout patterns and investment rates-of-return. Since the loss payout varies by line of business, consideration of the potential distribution of this present value for each treaty may provide a more reasonable basis of comparison than does item (1).

3. Various corporate parameters and decision-making criteria. These include the following:

a) the potential distributions of aggregate loss and/or present value of cash flow estimated on the rest of the reinsurer's contract portfolio.

b) the reinsurer's financial surplus, both the current evaluation and the potential distribution of future values due to reserve changes and losses arising from the rest of the contract portfolio.

c) the reinsurer's financial assets and investment opportunities.

d) various corporate goals, e.g., "growth and profits with honor" (David J. Grady, address at the March 7, 1979 Casualty Actuaries of New York meeting).

e) the reinsurer's attitude toward the trade-off of risk versus rate-of-return on each contract and on his whole reinsurance portfolio.

Items (a) - (e) are meant to indicate some of the considerations which might define a utility function for corporate decision-making. For any typical treaty evaluation, it may be possible to localize our attention and only reflect these global considerations indirectly. However, in the long run they may not be ignored.

Other more ambiguous items which a reinsurer might consider include:

4. The surplus necessary to "support" the treaty from the reinsurer's point-of-view. The seller of any insurance or reinsurance contract exposes part of his surplus or net worth to

the risk that the loss will exceed the pure premium. Although it seems reasonable that some amount of surplus might be allocated to support any contract, there is yet no satisfactory theoretical functional definition. Note that this "supporting surplus" per treaty may not sum to the reinsurer's total surplus; he may be interested in surplus allocation on a relative basis: Does treaty A need more "supporting surplus" than treaty B?

5. The potential distribution of rate-of-return on the "supporting surplus" for this treaty relative to the rates-of-return on other treaties in the reinsurer's contract portfolio.

It should be apparent that neither we nor anyone else has a *premium calculation principle which explicitly considers all these* items. They are listed here to illustrate the complexity of the problem of accurately pricing reinsurance treaties. (Indeed, we would argue that it is almost as difficult to price any other large insurance contract or group of contracts.) We believe that thoughtful reinsurance underwriters do evaluate treaty proposals along these or similar lines. To model this process reasonably well is difficult but not impossible, since there are many good theoretical models and estimation techniques available to the modern actuary.

Of all the items, item (1), the potential distribution of aggregate loss to the reinsurer, is the least ambiguous and the most important. Thus, the remainder of this paper concentrates upon the estimation of this distribution for excess-of-loss casualty working covers. We will describe a reasonable mathematical model for this distribution and an estimation procedure for parameterizing the model.

IV. AN AGGREGATE LOSS MODEL

This section describes a mathematical model for the aggregate losses to be paid out on a particular insurance contract. The general insurance loss model will then be specialized for an excess-of-loss reinsurance treaty. The model is based upon the concepts of collective risk theory developed by Bühlmann and others: for example, see Bühlmann (1969) and Beard, Pentikäinen and Pesonen (1977). The model is designed to allow the observer to account for and quantify his uncertainty regarding the "true" distribution of aggregate loss for a particular insurance contract(s). This uncertainty arises from many sources; among them are:

1. Any particular probability model is inexact.
2. Any parameters estimated from sample data are random; that is, subject to sampling errors.
3. The historical loss data may not be at final settlement values, but are themselves random estimates.
4. The proper adjustments for inflation over time are unknown.

5. The underlying insured population for the coverage period to be evaluated is different from the past population.
6. There are often data errors and analytical blunders.

The model will be developed from a subjective Bayesian viewpoint; the particularization of the model is determined from the viewpoint of an observer at a particular time with particular information. An honest competent reinsurer and an honest competent cedant would most likely have different final parameterized models for any given treaty. For a further discussion of subjective or "personal" probability, see Savage (1954) and Raiffa (1968).

The collective risk model describing the distribution of aggregate loss consists of many possible particular probability models, each of which is given a "weight" based upon its subjective likelihood. In this way, the total uncertainty regarding the particular outcome which will be realized is broken down into two pieces: 1) the uncertainty regarding the "best" particular model, sometimes called the parameter risk, and 2) the uncertainty regarding the actual loss value to be realized even when the particular probability model is known, sometimes called process risk. See Freifelder (1976) or Miccolis (1977) for further discussions of these actuarial concepts.

We will use the term "parameter" in a broader sense than is customary. A "parameter" will consist of a complete specification of a particular probability model such as the lognormal, or group of models, together with their usual parameters. Our uncertainty as to which parameter is "best" will be defined by a subjective probability distribution on the set of possible parameters.

It is easier to start with the case where the parameter is known (the particular model is specified). Let the random variable L denote the aggregate loss to be paid out on a given insurance contract for a particular coverage year. We begin by assuming that the total coverage (exposure) can be split into independent homogeneous coverage groups in the following manner. Suppose that L can be written as:

$$(4.1) \qquad L = L_1 + L_2 + \ldots + L_k$$

where L_i = random variable denoting the aggregate loss for group i, i = 1, 2, . . ., k.

Further, suppose that each L_i can be written as:

$$(4.2) \qquad L_i = X_{i1} + X_{i2} + \ldots + X_{iN_i}$$

where N_i = random variable denoting the number of losses (occurrences) for group i.

X_{ij} = random variable denoting the size (loss amount) of the j^{th} loss for group i.

Groups may be defined by any grouping of insureds or coverage which our power of analysis can reasonably and credibly separate. Examples of groups could be:

1. distinct groups of classes of insureds or coverages.

2. similar insureds grouped by distinct policy limit.

3. the overall coverage time period split into sub-periods.

For example A, our groups will be defined by year of coverage and ISO doctor class (the older seven class scheme). For example B, our groups will be defined by combined bodily injury and property damage policy limit.

Let $F(x|\theta) = \text{Prob}[L \leq x|\theta]$ be a particular c.d.f. (cumulative distribution function) for L with known parameter θ. Think of θ as being a comprehensive parameter (vector) containing all the parameters necessary to specify the particular c.d.f.'s for the L_i's, N_i's and X_{ij}'s. Now make the following assumptions:

<u>Assumption 1</u>: Given θ, the L_i's are stochastically independent.

<u>Assumption 2</u>: Given θ, the X_{ij}'s are stochastically independent of the N_i's.

<u>Assumption 3</u>: Given θ, for fixed i, the X_{ij}'s are stochastically independent and identically distributed.

These assumptions split the total coverage into independent homogeneous coverage groups.

The model with known parameter θ has very nice properties. The first property is that $F(x|\theta)$ is the convolution of the c.d.f.'s for individual groups:

$$(4.3) \qquad F(x|\theta) = F_1(x|\theta) * F_2(x|\theta) * \cdots * F_k(x|\theta)$$

where $F_i(x|\theta) = \text{Prob}[L_i \leq x|\theta]$ for $i = 1, 2, \ldots, k$.

From this it follows that the cumulants of L given θ are straightforward sums of the cumulants of the L_i's given θ:

$$(4.4) \qquad K_m(L|\theta) = \sum_i K_m(L_i|\theta)$$

where $K_m(L|\theta)$ is the m^{th} derivative of the logarithm of the moment generating function of L evaluated at 0 (if it exists). Likewise for the $K_m(L_i|\theta)$'s.

See Kendall and Stuart (1966), pp. 157ff, for a discussion of cumulants. In particular, the first three cumulants add:

$$(4.5) \qquad \begin{aligned} K_1(L|\theta) &= E[L|\theta] = \sum_i E[L_i|\theta] \\ K_2(L|\theta) &= Var[L|\theta] = \sum_i Var[L_i|\theta] \\ K_3(L|\theta) &= \mu_3(L|\theta) = \sum_i \mu_3(L_i|\theta) \end{aligned}$$

where $\mu_m(L|\theta) = E[(L - E[L|\theta])^m|\theta]$

Because of assumptions 2 and 3, each $F_i(x|\theta)$ can be written in terms of the c.d.f.'s of N_i and X_i, where X_i is the common loss amount random variable for group i:

$$(4.6) \qquad F_i(x|\theta) = \sum_n \text{Prob}[N_i = n|\theta] \cdot G_i^{*n}(x|\theta)$$

where $G_i(x|\theta) = \text{Prob}[X_i \leq x|\theta]$ for $i = 1, 2, \ldots, k$.

A consequence of (4.6) is that the first three moments of L_i given θ may be written:

$$(4.7) \qquad \begin{aligned} E[L_i|\theta] &= E[N_i|\theta] \cdot E[X_i|\theta] \\ \text{Var}[L_i|\theta] &= E[N_i|\theta] \cdot \text{Var}[X_i|\theta] + \text{Var}[N_i|\theta] \cdot E[X_i|\theta]^2 \\ \mu_3(L_i|\theta) &= E[N_i|\theta] \cdot \mu_3(X_i|\theta) + \mu_3(N_i|\theta) \cdot E[X_i|\theta]^3 \\ &\quad + 3 \cdot \text{Var}[N_i|\theta] \cdot E[X_i|\theta] \cdot \text{Var}[X_i|\theta] \end{aligned}$$

The scheme will be to develop parameterized models for the N_i's and X_i's, calculate their first three moments given θ, and then use (4.7) to calculate the first three moments of the L_i's and use (4.5) to calculate the first three moments of L given θ.

The collective risk model is obtained by deleting the restriction that θ is known. Instead, assume that the set Ω of possible parameters is known and that we can specify a subjective probability distribution $U(\theta)$ on Ω which gives the subjective likelihood of each subset of Ω. Bühlmann (1970) calls $U(\theta)$ a structure function. For simplicity, assume that Ω is finite so that $U(\theta)$ is a discrete probability:

<u>Assumption 4</u>: Ω is the finite set of possible parameters and $U(\theta)$ is the likelihood of the parameter θ.

Ω and $U(\theta)$ specify the observer's uncertainty regarding the "best" parameter.

With Ω and $U(\theta)$ specified, the unconditional c.d.f. $F(x)$ of L is the weighted sum of the conditional c.d.f.'s $F(x|\theta)$:

$$(4.8) \qquad F(x) = \sum_{\theta} F(x|\theta)\cdot U(\theta)$$

Likewise, for each $F_i(x)$, the c.d.f. of L_i.

A consequence of (4.8) is (Bühlmann (1970), p. 66):

$$(4.9) \qquad E[L^m] = \sum_{\theta} E[L^m|\theta]\cdot U(\theta) \qquad \text{for } m = 0, 1, 2, \ldots$$

Likewise, for each L_i.

With θ unknown, assumptions (1) - (3) may no longer hold, for the uncertainty regarding θ may simultaneously affect the model at all levels. For example, the c.d.f.'s of the L_i's are usually subjectively derived from historical data altered by loss development and inflationary trend assumptions. The assumptions made simultaneously about each L_i and L_j are usually not independent, i.e., the particular parameters for the c.d.f. of L_i are correlated with the particular parameters for the c.d.f. of L_j. Symbolically:

$$
\begin{aligned}
E[L_i L_j] &= \sum_\theta E[L_i L_j|\theta]\cdot U(\theta) \\
&= \sum_\theta E[L_i|\theta]\cdot E[L_j|\theta]\cdot U(\theta) \qquad (4.10) \\
&\neq (\sum_\theta E[L_i|\theta]\cdot U(\theta))\cdot(\sum_\theta E[L_j|\theta]\cdot U(\theta))
\end{aligned}
$$

When θ is unknown, equations (4.3) - (4.7) usually no longer hold. In particular, equation (4.5) now holds only for the first moment:

$$
\begin{aligned}
E[L] &= \sum_i E[L_i] \\
K_m(L) &\neq \sum_i K_m(L_i) && \text{for } m \neq 1 \\
E[L^m] &\neq \sum_i E[L_i^m] && \text{for } m \neq 1 \qquad (4.11) \\
\mu_m(L) &\neq \sum_i \mu_m(L_i) && \text{for } m \neq 1
\end{aligned}
$$

Thus, the moments of L must now be evaluated directly from (4.9) by using (4.5) and (4.7); likewise for each L_i. For example, the second moment of L is now written:

$$\begin{aligned} E[L^2] &= \sum_\theta E[L^2|\theta]\cdot U(\theta) \\ &= \sum_\theta \{Var[L|\theta] + E[L|\theta]^2\}\cdot U(\theta) \qquad (4.12) \\ &= \sum_\theta \{(\sum_i Var[L_i|\theta]) + (\sum_i E[L_i|\theta])^2\}\cdot U(\theta) \end{aligned}$$

Continue the expansion using formula (4.7).

Likewise for each L_i.

This general collective risk model may be specialized to the case of an excess-of-loss reinsurance treaty. Suppose that the treaty covers group i losses in the layer from r_i (retention) up to b_i. The general model may be specialized in at least two different ways. The first interpretation views X_i as the excess portion of each loss. We drop the subscript i in the following:

Model 1 Notation:

N = random variable denoting total number of non-zero losses ground-up.

X = random variable denoting that part between r and b of each ground-up loss.

S = random variable denoting the ground-up loss amount.

Given that a loss has occurred, X and S are related by:

$$X = \begin{cases} 0 & \text{if} \quad S \le r \\ S - r & \text{if } r < S < b \\ b - r & \text{if } b \le S \end{cases} \tag{4.13}$$

Thus, the c.d.f.'s of S and X given θ are related by:

$$G_X(x|\theta) = \begin{cases} G_S(r|\theta) & \text{if} \quad x \le 0 \\ G_S(x+r|\theta) & \text{if} \quad 0 < x < b - r \\ 1 & \text{if} \quad b - r \le x \end{cases} \tag{4.14}$$

If N is to denote the number of excess losses, then use the second specialization:

<u>Model 2</u> Notation:

N = random variable denoting the number of excess loss occurrences.

X = random variable denoting the size of an excess loss, given that an excess loss has occurred.

EN = random variable denoting the total number of non-zero ground-up losses, called "base number".

S = random variable denoting the ground-up loss amount.

With known parameter θ, the c.d.f.'s of N and BN are related by:

(4.15)
$$\text{Prob}[N=n|\theta] = \sum_{m\geq n}\{\text{Prob}[BN=m|\theta]\cdot\binom{m}{n}\times (1 - G_S(r|\theta))^n \cdot G_S(r|\theta)^{m-n}\}$$

where $G_S(r|\theta) = \text{Prob}[S \leq r|\theta]$

In particular, it is easy to show that:

(4.16)
$$E[N|\theta] = E[BN|\theta]\cdot(1 - G_S(r|\theta))$$

Likewise, the c.d.f.'s of X and S for Model 2 are related by:

(4.17)
$$G_X(x|\theta) = \begin{cases} 0 & \text{if } x \leq 0 \\ G_S(x+r|\theta)\cdot(1 - G_S(r|\theta))^{-1} & \text{if } 0 < x < b - r \\ 1 & \text{if } b - r \leq x \end{cases}$$

Model 1 is easier to work with since the definition of N remains the same when different retentions are considered. But, it is easy to trade back and forth between the two models and, most importantly, they both yield identical answers for the distribution of L. We prefer to use Model 1, so hereafter N will be the number of non-zero ground-up losses.

The next three sections show how this general model may be used to evaluate the loss potentials of particular treaties. To do so, we must:

1. specify the homogeneous groups.
2. specify the set of possible parameters Ω and the subjective likelihood $U(\theta)$, of each θ in Ω.
3. calculate (using a computer package) the moments and approximate various percentiles of L from the moments of the N_i's and X_i's given the θ's.

V. PARAMETER ESTIMATION: EXAMPLE A

The most difficult part of this aggregate loss evaluation procedure is estimating the parameters to be used in the models. The estimation for A Doctors' Mutual Insurance Company, example A, will illustrate the case where there are no credible historical loss data directly related to the exposure. In this case, general industry information must be used together with substantial judgement. In general, in this situation we presently estimate three parameters based upon low, medium and high loss frequency and severity assumptions (We purposely use the word "medium" to avoid the statistical theoretic connotations of words such as "mean" and "median".) For example A, the estimates will be based upon Insurance Services Office ratemaking data and further modified by judgement based upon the NAIC Medical Malpractice Closed Claim Surveys (1977) and (1978).

The groups for example A are selected to be the seven doctors classes in the old ISO class plan because we believe there are sufficient data to separate these classes for loss frequency and severity. The complete parameter matrix is displayed in Table 5A. It looks formidable but is really quite simple; much of it is repetitive and based upon standardized judgement. Each class is represented by three rows: the low θ is the first row for each class, the medium θ is the second row for each class and the high θ is the third row for each class. In Section VII these parameters will be input to a Prudential Reinsurance Company computer package named RISKMODEL which will calculate the moments of the aggregate loss L for the layer $750,000 excess of $250,000 for the coverage year 1980/81 using the formulas from Section IV. The package also approximates selected percentiles of the distribution of L.

The form of the parameterized c.d.f.'s we shall use for the distribution of the number of loss occurrences N_i for class i is the negative binomial defined in Appendix D. Thus, we must specify two parameters for each c.d.f.; we will specify $E[N_i|\theta]$ and the ratio $Var[N_i|\theta] \div E[N_i|\theta]$ for each class i for each θ. The expected number of ground-up loss occurrences $E[N_i|\theta]$ is based upon the exposure and loss frequency estimates in Table 5A, columns (2) and (3). The estimates of exposure by class are based upon ISO exposure data

TABLE 5A

GROUP	EXPOSURE	FREQUENCY	VENI-CENT	DST CODE	PAR 1	PAR 2	TRUNC. PT T	EXCESS PROP XP	WGT	INDEX
CLASS1 1	215.0000	.0050	1.0000	2	23640.0000	1.4040	1000	.0000	.2500	1
CLASS1 2	215.0000	.0062	1.5000	2	18450.0000	1.2980	1000	.8560	.5000	0
CLASS1 3	215.0000	.0075	2.0000	2	18155.0000	1.1910	1000	.8380	.2500	0
CLASS2 1	77.0000	.0072	1.0000	2	23640.0000	1.4040	1000	.0000	.2500	1
CLASS2 2	77.0000	.0090	1.5000	2	18450.0000	1.2980	1000	.8560	.5000	0
CLASS2 3	77.0000	.0108	2.0000	2	18155.0000	1.1910	1000	.8380	.2500	0
CLASS3 1	66.0000	.0085	1.0000	2	23923.0000	1.4650	1000	.8560	.2500	1
CLASS3 2	66.0000	.0106	1.5000	2	20106.0000	1.2700	1000	.8860	.5000	0
CLASS3 3	66.0000	.0127	2.0000	2	20597.0000	1.1890	1000	.8950	.2500	0
CLASS4 1	10.0000	.0104	1.0000	2	22253.0000	1.4650	1000	.8560	.2500	1
CLASS4 2	10.0000	.0130	1.5000	2	21724.0000	1.2780	1000	.8860	.5000	0
CLASS4 3	10.0000	.0156	2.0000	2	21762.0000	1.1890	1000	.8950	.2500	0
CLASS5 1	46.0000	.0130	1.0000	2	27911.0000	1.4650	1000	.8560	.2500	1
CLASS5 2	46.0000	.0163	1.5000	2	25450.0000	1.2700	1000	.8860	.5000	0
CLASS5 3	46.0000	.0195	2.0000	2	24031.0000	1.1890	1000	.8950	.2500	0
CLASS6 1	35.0000	.0169	1.0000	2	23923.0000	1.4650	1000	.8560	.2500	1
CLASS6 2	35.0000	.0212	1.5000	2	20106.0000	1.2780	1000	.8860	.5000	0
CLASS6 3	35.0000	.0254	2.0000	2	20597.0000	1.1890	1000	.8950	.2500	0
CLASS7 1	51.0000	.0156	1.0000	2	30569.0000	1.4650	1000	.8560	.2500	1
CLASS7 2	51.0000	.0195	1.5000	2	25692.0000	1.2700	1000	.8860	.5000	0
CLASS7 3	51.0000	.0234	2.0000	2	26320.0000	1.1890	1000	.8950	.2500	0

and the assumption that there will be 500 doctors. Possible variance of the actual exposures from these estimates will be simply accounted for when selecting the low and high frequency estimates. The medium frequency (ground-up) estimates are derived in Appendix A, p. A1. They are based upon projections of overall countrywide doctor loss frequency at the mid-point (January 1, 1981) of coverage year fiscal 1980/81, modified by various offsets: 1) class, 2) state, 3) year in claims-made program (in this case, first year) and 4) contagion (multiple doctors per incident). It is necessary to use a contagion factor to adjust the basic ISO data, which are number of occurrences per doctor, since the treaty will cover loss per occurrence for all covered doctors added together. All the offsets are selected on the basis of ISO data and NAIC (1977, 1978) information. The low and high loss frequencies are selected to be ± 20% of the medium loss frequencies; this is pure judgement to reflect the uncertainty regarding the actual exposure and the "true" expected frequency per class. The ratio $Var[N_i|\theta] \div E[N_i|\theta]$ values 1.0 (low), 1.5 (medium) and 2.0 (high), Table 5A, column (4), are selected on the basis of research by the ISO Increased Limits Subcommittee.

The parameters for the loss amount c.d.f.'s are in Table 5A, columns (5) - (9). The number 2 in column (5) specifies to the computer package RISKMODEL that the form of the c.d.f. is the 4-parameter modified Pareto distribution defined in Appendix D; the other choices are 1 = lognormal and 3 = Weibull. Columns (6) - (9) are its four parameters for each class and each θ. We and the ISO Increased Limits Subcommittee have found this general Pareto c.d.f. to be very useful for describing loss amount distributions. The particular parameters derived on Appendix A, pp. A2 and A3 are based upon ISO countrywide loss amount data and modified by various offsets (class, state and contagion) selected on the basis of other ISO data and NAIC (1977, 1979) information. Note that the offsets apply to the β parameter (PAR1) only. We do not presently offset according to year in claims-made program, although we might if we ever see any claims-made loss data sufficient for this purpose. The low, medium and high parameters are selected from c.d.f.'s fitted to five policy years of ISO data via the maximum likelihood techniques described in Patrik (1980) and are indeed the low, medium (all five years combined) and high c.d.f.'s.

Column (10) of Table 5A displays the subjective weights assigned to the three parameters. In this case, they are purely judgemental, with the medium parameter assigned a likelihood of 50% and the low and high parameters assigned 25% likelihoods.

VI. PARAMETER ESTIMATION: EXAMPLE B

The parameter estimation for example B, the exeess proposal for P&C Insurance Company's general liability coverage, will illustrate the case where there are credible historical loss data directly related to the exposure. In this case, we will use as much of the data as we can to select the homogeneous coverage groups, to estimate the forms of the loss amount c.d.f.'s and to estimate some θ's and $U(\theta)$'s (the loss count c.d.f.'s are assumed to be adequately modeled by negative binomial distributions). Recall from Section II that the proposal is for $750,000 excess of $250,000 and that the P&C Insurance Company has provided a detailed history of large losses (greater than $25,000), gross earned premiums, an overall rate history and more.

The steps of the procedure we will follow are:

1. Select the homogeneous coverage groups.
2. Decide which historical exposure years are most indicative of (can be easily adjusted to) next year's exposure.
3. Estimate loss amount inflationary trend factors.
4. Select a primary retention to directly evaluate loss count and amount distributions for the next coverage year and restrict attention to those

large losses whose trended values are greater than this retention. This retention is not necessarily the proposed retention, but is instead the one which we believe will yield the most credible estimates of the potential loss.

5. Decide how to adjust the large loss data to an ultimate settlement basis.
6. Estimate ground-up loss amount c.d.f.'s for the next coverage year, both forms and parameters, from the large loss data and general information.
7. Estimate the number of excess IBNR losses (excess of the deflated values of the selected retention (4)).
8. Estimate excess loss frequencies for the next coverage year.
9. Estimate base (ground-up) loss count c.d.f.'s for the next coverage year based upon (6), (8) and the estimated exposure.
10. Select the parameter weights $U(\theta)$.

The procedure for example B will follow this outline very cleanly. In practice, however, any of the steps may be reversed and any of the decisions may be changed later during the procedure if the analysis so indicates.

We decided not to display the complete P&C Insurance Company data in an appendix for three reasons:

1. We would like to focus on the general procedure, not all the details. Most of the detailed steps could be done in many different ways.
2. The data are voluminous.
3. The data, used with the primary company's permission, should remain confidential.

Many summary exhibits are displayed in Appendix B.

Step 1

The groups are defined by the major policy limits based upon the policy limits listed on the large loss records and P&C Insurance Company's estimate of their policy limits distribution for 1980. However, the general liability coverage will be analyzed as a whole; thus, the parameters of the estimated ground-up loss amount c.d.f.'s and the loss frequencies will be the same for each group - only the policy limits and the underlying exposure will be different. The complete parameter matrix which will later be input to the RISKMODEL computer package is displayed in Table 6A. In this case, there are four policy limit groups: \$200,000, \$250,000, \$350,000 and \$500,000 or more; there are four parameters θ: the first is the combination of the first row for each group, and so on.

TABLE 6A

GROUP		EXPOSURE	FREQUENCY	V[N]÷E[N]	DST CODE	PAR 1	PAR 2	TRUNC. PT T	EXCESS PROP XP	WGT	INDEX
GL/200	1	1175.0000	.0108	1.5000	2	124016.0000	3.6795	0	1.0000	.1000	1
GL/200	2	1175.0000	.0135	2.0000	2	89251.0000	3.1290	0	1.0000	.4000	0
GL/200	3	1175.0000	.0096	1.5000	2	130747.0000	3.8769	0	1.0000	.1500	0
GL/200	4	1175.0000	.0104	2.0000	2	130693.0000	3.7558	0	1.0000	.3500	0
GL/250	1	1175.0000	.0108	1.5000	2	124016.0000	3.6795	0	1.0000	.1000	1
GL/250	2	1175.0000	.0135	2.0000	2	89251.0000	3.1290	0	1.0000	.4000	0
GL/250	3	1175.0000	.0096	1.5000	2	130747.0000	3.8769	0	1.0000	.1500	0
GL/250	4	1175.0000	.0104	2.0000	2	130693.0000	3.7558	0	1.0000	.3500	0
GL/350	1	2350.0000	.0108	1.5000	2	124016.0000	3.6795	0	1.0000	.1000	1
GL/350	2	2350.0000	.0135	2.0000	2	89251.0000	3.1290	0	1.0000	.4000	0
GL/350	3	2350.0000	.0096	1.5000	2	130747.0000	3.8769	0	1.0000	.1500	0
GL/350	4	2350.0000	.0104	2.0000	2	130693.0000	3.7558	0	1.0000	.3500	0
GL/500+	1	18800.0000	.0108	1.5000	2	124016.0000	3.6795	0	1.0000	.1000	1
GL/500+	2	18800.0000	.0135	2.0000	2	89251.0000	3.1290	0	1.0000	.4000	0
GL/500+	3	18800.0000	.0096	1.5000	2	130747.0000	3.8769	0	1.0000	.1500	0
GL/500+	4	18800.0000	.0104	2.0000	2	130693.0000	3.7558	0	1.0000	.3500	0

Step 2

We restrict our attention to the large loss data from accident years 1973 through 1978 since we believe that these data are more easily adjustable to 1980 level in a reasonable manner. Also, there does not appear to be any significant development of loss counts or amounts beyond the 78 month evaluations of the data presented in P&C's June 30, 1975, . . ., June 30, 1979 loss evaluations. With this decision, we still have quite enough data, over 200 large losses, to analyze.

Step 3

Many different loss amount inflationary trend models may be developed using many different economic and actuarial assumptions. We shall use two very simple models:

1. Exponential trend model: ISO general liability bodily injury average loss amounts of various kinds from the past several years may be fit by exponential curves in the usual manner. In this case, our model produces an annual trend estimate of 16.8%.

2. Econometric trend model: Slightly more sophisticated trend estimates are derived via a primitive but reasonable econometric model using the Bureau of Labor Statistics' Consumer Price Index and its Medical Care Services component as independent variables and some ISO loss amount

index as the dependent variable. The trend factors to adjust each accident year's data to 1980 level are displayed in Appendix B, p. B2, column (1).

Loss parameters will be derived separately from the two sets of data adjusted by these two trend models. In general, use as many reasonable trend models as possible and assign subjective weights to them.

Step 4

Our objective is to estimate 1980 ground-up loss amount and loss count c.d.f. models which produce accurate estimates of the losses in the layer $400,000 excess of $100,000. However, to estimate these models, it is not necessary to restrict our attention to only those historical losses whose 1980 level values are greater than $100,000. With the exponential and econometric trend models, a 1980 retention of $75,000 deflates to 1973 values of $25,291 and $25,299, respectively (see column (2) of Appendix B, pp. B1 and B2). Since these deflated values are larger than $25,000, the 1973 - 78 large loss data contain all known losses whose 1980 values are larger than $75,000. Furthermore, more credible excess frequency and loss amount estimates may be obtained from evaluating a lower retention of $75,000. That is, there are 171 (exponential) and 158 (econometric) known losses whose 1980 values are greater than $75,000 (see Appendix B, pp. B1 and B2), while only 109 (exponential) and 104 (econometric) have 1980 values greater than $100,000. Therefore, we restrict our attention to those large losses whose 1980 level values are greater than $75,000. The 1980 level average values and number of occurrences at each evaluation date are shown in Appendix B, pp. B1 and B2.

Step 5

For each historical coverage year, we want an estimate of the distribution of ultimate settlement values (1980 level) of losses greater than $75,000. The age-to-age development factors displayed in Appendix B, pp. B1 and B2, for the 1980 level average values indicate that the large loss distribution for the recent years will change as more losses pierce the retention and as the losses are settled. Thus, these data must be adjusted. In this case we observe that the loss amount distribution appears to develop little beyond the 42 month evaluation. Also, the two years for which we can expect the data to substantially develop, 1977 and 1978, have only 14 and 3 large losses respectively. Thus, in this case we choose to use multiplicative average size development factors applied to the large loss values. These factors are displayed in Appendix B, pp. B1 and B2. (For a more sophisticated approach, which simultaneously accounts for the development of loss counts and amounts, see Hachemeister (1976)).

Step 6

The 1980 loss amount c.d.f.'s are derived from four data sets by using the maximum likelihood estimation techniques and testing procedures described in Patrik (1980). The data sets are:

1. The large losses together with their policy limits adjusted to 1980 level via the exponential trend model and developed to ultimate settlement.
2. Same as (1) except that the losses and policy limits are censored at (limited to) $500,000.
3. The large losses together with their policy limits adjusted to 1980 level via the econometric trend model and developed to ultimate settlement.
4. Same as (3) except that the losses and policy limits are censored at $500,000.

Censorship at $500,000 is used in (2) and (4) for two reasons:

1. The proposed reinsurance layer stops at $500,000. Thus, we may focus upon the loss amount distribution below $500,000.
2. In general, we have found that censored (by policy limits) loss amount c.d.f.'s estimated via the method of maximum likelihood fit better when there are some losses at the censorship points: the parameter estimates appear to have smaller sample error.

However, the data in this case have no losses at their policy limits.

The parameters for c.d.f.'s (1) - (4) are displayed in Table 6A, columns (5) - (9). Both the Kolmogorov-Smirnov Test and an "actuarial ad-hoc expected value test" (see Patrik (1980)) show the Pareto model fitting much better than either the lognormal or the Weibull models. Thus, each selected c.d.f. is Pareto (column (5) entry is 2). The column (8) and (9) entries are selected for convenience to be 0 and 1, respectively, because we are not concerned with the lower end of the loss amount distribution. See Appendix D and note that if XP = 1, then the four parameter model reduces to a two parameter model with the parameters PAR1 and PAR2 in Table 6A, columns (6) and (7). C.d.f.'s (2) and (4) fit well, while the fit of (1) and (3) is only fair. This information will be used later when selecting the subjective likelihoods (weights) of the parameters.

Step 7

The number of IBNR (incurred but not reported) 1980 level losses excess of $75,000 for each year 1973, . . ., 1978 are estimated using a method developed by James Stanard and described in Patrik (1978). The first step is to estimate a c.d.f. model for the distribution of report lags. In this case, the report lag is defined as the time in months between the date of occurrence of a loss and the date its 1980 level incurred value first exceeds \$75,000. Weissner (1978) showed how to estimate this c.d.f. using the method of maximum likelihood when the data include month of occurrence and month of report for every loss. However in this case, such detail is not available: the data have only year of occurrence (accident) and year of report. Thus, we select a report lag c.d.f. model by comparing the actual number of occurrence age-to-age factors in Appendix B, pp. B1 and B2, to tables of annual age-to-age factors generated by various theoretical report lag distributions, such as the exponential, lognormal or Weibull. In this case, a Weibull distribution with parameters $\beta = 34.0$ and $\delta = 2.75$ (see Appendix D) appears to describe both sets of actual age-to-age factors best; so we will use it to calculate IBNR. The annual age-to-age factors generated by this Weibull are the row underlined in the table in Appendix B, p. B3. The IBNR calculations are displayed in Appendix B, pp. B4 and B5.

Step 8

Appendix B, pp. B6 and B7, displays the estimated IBNR per year (column (4)) and the implied 1980 level frequency excess-of-\$75,000 per year (column (6)) with respect to gross direct earned premium at present (1980) rate level (column (2)). Columns (7) and (8) display our estimates of the 1980 level base frequency per year. We use the term "base frequency" to distinguish these numbers from the true ground-up loss frequency. The base frequencies are slightly fictitious numbers derived solely as convenient input for the RISKMODEL computer package (table 6A, column (3)). They are interpolated downward from the excess frequencies by use of the previously selected loss amount c.d.f. models. For example, the base frequency of .0108 for 1973 in column (7) of Appendix B, p. B6, is derived from the excess frequency of .0019 in column (6) via:

$$\begin{aligned}
&(\text{excess frequency}) \div \text{Prob}[X>\$75{,}000 \mid \text{c.d.f.}(1)] \\
(6.1)\qquad &= (.0019) \div \left(\frac{\beta}{\beta + 75{,}000}\right)^{\delta} \\
&= .0108
\end{aligned}$$

where $\beta = 124{,}016$ and $\delta = 3.6795$.

The base frequencies with respect to all four loss amount c.d.f.'s are displayed in Appendix B, pp. B6 and B7, along with four selected values which are input in Table 6A, column (3).

Step 9

The negative binomial c.d.f. is selected as the general form for the distribution of N_i, the number of 1980 base losses for policy limit group i. The expected value for each particular c.d.f. is the base frequency times the estimate of the 1980 gross direct earned premium in Table 6A, column (2). The ratios $\text{Var}[N_i|\theta] \div E[N_i|\theta]$ in column (4) are again selected on the basis of research by the ISO Increased Limits Subcommittee.

Step 10

The parameter weights $U(\theta)$ in Table 6A, column (10), are selected on the following basis:

1. Each trend model is given weight .50.
2. The weight selected for loss amount c.d.f. (2) together with its implied base frequency is .40 (out of .50 possible) since it fit best; the remaining .10 goes to c.d.f. (1). Likewise, loss amount c.d.f. (4) together with its implied base frequency is given a weight of .35 because of its good fit, with the remaining .15 going to c.d.f. (3).

As a final remark on the parameter estimation for example B, it should be apparent that if we believe that the P&C large loss data is not fully credible, then we can append more parameters based upon general industry information as in example A. The parameter weights would be adjusted accordingly, perhaps via some credibility procedure.

VII. MOMENTS AND PERCENTILES OF THE DISTRIBUTION OF AGGREGATE LOSS

This section describes a computer package named RISKMODEL which takes information such as in Tables 5A and 6A and transforms it into moments and percentiles of the distribution of aggregate loss for any selected mixture of loss layers. Tables 7A, 5A and 7B-7D document a RISKMODEL run for example A; the run for example B is contained in Appendix C and Tables 6A and 7E. In both cases the printout displays both the package interrogatories and the user's input. Almost complete runs are displayed so that the reader can see how easily the complicated model formulas translate into a working computer package; the only parts eliminated are the step-by-step data input process and some ending details regarding further displays and memory storage.

Table 7A displays the beginning of the RISKMODEL run for example A. The user enters the group names "class 1, class 2, . . ., class 7", specifies that there will be three parameters and indicates that he wants the limits matrix LIM in the package to be assigned the elements of a previously created matrix LIMA. Since the proposed coverage is $750,000 excess of $250,000, the loss layers we want to consider are 0 - $250,000 and $250,000 - $1,000,000; we observe the output for the lower layer to provide an extra check on the reasonableness of the output for the excess layer. For each group (class), there are two rows with lower and upper limit columns and a third column, INDEX, which indicates when there is a change in group.

The user next specifies that he wants the parameter matrix PAR in the package to be assigned the elements of a previously

TABLE 7A

```
      RICKMODEL
DO NOT PANIC IF YOU MAKE AN ERROR WHILE INPUTTING.
OPPORTUNITY TO CHANGE LATER.

ENTER MAJOR GROUP NAMES AS FOLLOWS:   /GRP1/GRP2..........
NOTE: MUST BE IN QUOTES.   FOR MORE THAN 1 LINE OF INPUT, USE ,O
O:
      '/CLASS1/CLASS2/CLASS3/CLASS4/CLASS5/CLASS6/CLASS7'

ENTER THE NUMBER OF PARAMETERS,  E.G.   5
O:
      3

DO YOU WISH TO  (1) INPUT VECTOR OF LIMITS,  OR
(2) USE MATRIX OF LIMITS PREVIOUSLY CREATED. 1 OR 2.
O:
      2
ENTER THE NAME OF THE MATRIX OF LIMITS PREVIOUSLY CREATED
NOTE: NAME SHOULD HAVE PREFIX LIM
LIMA

DO YOU WISH TO SEE THE LIM MATRIX.  Y OR N
Y
                         LIMITS
             LOWER               UPPER               INDEX
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0
                 0              250000                 1
            250000             1000000                 0

DO YOU WISH TO MAKE ANY CHANGES IN THE LIM MATRIX.  Y OR N
N

DO YOU WISH TO
(1) INPUT VECTOR OF PARAMETERS FOR THE FIRST SUBGROUP  OR
(2) USE MATRIX OF PARAMETERS PREVIOUSLY CREATED.  1 OR 2
O:
      2
ENTER THE NAME OF THE MATRIX OF PARAMETERS PREVIOUSLY CREATED
NOTE: NAME SHOULD HAVE PREFIX PAR
PARA

DO YOU WISH TO SEE THE PAR MATRIX.  Y OR N
Y
```

created matrix PARA. The parameter matrix was displayed in Table 5A.

Table 7B continues the run after the display of the parameter matrix PAR. Next displayed is a matrix of intermediate calculations for layer 1: 0 - \$250,000. The notation here is:

(7.1) A = layer lower bound (here $A = 0$)

B = layer upper bound (here $B = 250{,}000$)

S = ground-up loss amount random variable

$P[S>A] = 1-G_i(A|\theta)$ for each group i for each θ

$P[S>B] = 1-G_i(B|\theta)$ for each group i for each θ

$E[S{*}m] = \int_A^B x^m \, dG_i(x|\theta)$ for each group i for each θ

where $m = 1, 2, 3$

$$G_i(x|\theta) = \text{Prob}[S_i \leq x|\theta]$$

These values will be used to calculate the moments of the aggregate loss L given θ by using formula (4.7). They are displayed so that the user can check that the run is going alright.

Table 7C continues the run with a display of a matrix of intermediate calculations for layer 2: \$250,000 - \$1,000,000. These are similar to those for layers 1 except that here A = 250,000 and B = 1,000,000. Next input are the selected ε's

TABLE 7B

DO YOU WISH TO MAKE ANY CHANGES IN THE PAR MATRIX. Y OR N
N

GROUPS AND PARAMETER INPUT COMPLETED
TO PROCESS INTERMEDIATE CALCULATIONS, HIT EXECUTE

DO YOU WISH TO PRINT THE INTERMEDIATE CALCULATIONS,
P[S>A],P[S>B],E[S],E[S*2],E[S*3]. Y OR N.
Y

INTERMEDIATE CALCULATIONS USED THROUGHOUT MOMENT CALCULATIONS

LAYER 1

GROUPS	P[S>A]	P[S>B]	E[S]	E[S*2]	E[S*3]
CLASS1 1	1.000	.021	2.216E04	1.813E09	2.415E14
CLASS1 2	1.000	.026	2.258E04	1.878E09	2.546E14
CLASS1 3	1.000	.034	2.363E04	2.065E09	2.868E14
CLASS2 1	1.000	.021	2.216E04	1.813E09	2.415E14
CLASS2 2	1.000	.026	2.258E04	1.878E09	2.546E14
CLASS2 3	1.000	.034	2.363E04	2.065E09	2.868E14
CLASS3 1	1.000	.024	2.392E04	1.985E09	2.660E14
CLASS3 2	1.000	.032	2.500E04	2.154E09	2.960E14
CLASS3 3	1.000	.042	2.717E04	2.459E09	3.456E14
CLASS4 1	1.000	.026	2.478E04	2.093E09	2.824E14
CLASS4 2	1.000	.034	2.585E04	2.261E09	3.124E14
CLASS4 3	1.000	.044	2.802E04	2.571E09	3.632E14
CLASS5 1	1.000	.030	2.642E04	2.304E09	3.146E14
CLASS5 2	1.000	.038	2.745E04	2.468E09	3.444E14
CLASS5 3	1.000	.050	2.962E04	2.787E09	3.973E14
CLASS6 1	1.000	.024	2.392E04	1.985E09	2.660E14
CLASS6 2	1.000	.032	2.500E04	2.154E09	2.960E14
CLASS6 3	1.000	.042	2.717E04	2.459E09	3.456E14
CLASS7 1	1.000	.033	2.795E04	2.507E09	3.463E14
CLASS7 2	1.000	.043	2.893E04	2.667E09	3.756E14
CLASS7 3	1.000	.055	3.110E04	2.994E09	4.303E14

TABLE 7C

LAYER 2

GROUPS	P[S>A]	P[S>B]	E[S]	E[S*2]	E[S*3]
CLASS1 1	.021	.003	8.015E03	4.093E09	2.431E15
CLASS1 2	.026	.005	9.754E03	5.103E09	3.097E15
CLASS1 3	.034	.007	1.227E04	6.522E09	4.014E15
CLASS2 1	.021	.003	8.015E03	4.093E07	2.431E15
CLASS2 2	.026	.005	9.754E03	5.103E09	3.097E15
CLASS2 3	.034	.007	1.227E04	6.522E09	4.014E15
CLASS3 1	.024	.003	9.025E03	4.622E09	2.753E15
CLASS3 2	.032	.006	1.176E04	6.177E09	3.763E15
CLASS3 3	.042	.009	1.513E04	8.059E09	4.967E15
CLASS4 1	.026	.004	9.701E03	4.973E09	2.964E15
CLASS4 2	.034	.006	1.253E04	6.589E09	4.016E15
CLASS4 3	.044	.009	1.605E04	8.556E09	5.276E15
CLASS5 1	.030	.004	1.108E04	5.690E09	3.396E15
CLASS5 2	.030	.007	1.409E04	7.419E09	4.527E15
CLASS5 3	.050	.010	1.790E04	9.551E09	5.897E15
CLASS6 1	.024	.003	9.025E03	4.622E09	2.753E15
CLASS6 2	.032	.006	1.176E04	6.177E09	3.763E15
CLASS6 3	.042	.009	1.513E04	8.059E09	4.967E15
CLASS7 1	.033	.005	1.248E04	6.423E09	3.841E15
CLASS7 2	.043	.008	1.566E04	8.258E09	5.045E15
CLASS7 3	.055	.011	1.973E04	1.055E10	6.520E15

TO PROCESS MORE INTERMEDIATE CALCULATIONS, HIT EXECUTE

ENTER EPSILON(S) FOR WHICH PROB(LOSS>MAX. PROB. LOSS) = EPSILON. (0<ε≤.5)
⎕:
      .1 .05 .01
NOW FOR THE FINAL PRINTOUT

ENTER COMPANY NAME
EXAMPLE A: A DOCTORS MUTUAL INSURANCE COMPANY

ENTER YOUR NAME (EG. J. SMITH)
HOWARD H. FRIEDMAN

ENTER TODAY'S DATE (EG. JAN. 1, 1979)
APRIL 1, 1980

ENTER IN PARENTHESIS AND QUOTES A SEVEN CHARACTER NAME FOR THE UNITS
(E.G. '(DOCTORS)' OR '__(BEDS)_')
⎕:   OF EXPOSURE CENTERED IN 9 SPACES
      '(DOCTORS)'

ADJUST PAPER TO TOP OF NEW PAGE & HIT EXECUTE

TABLE 7D

EXAMPLE A: A DOCTORS MUTUAL INSURANCE COMPANY

LAYER 1

GROUPS	LIMITS LOWER ($ 000)	LIMITS UPPER ($ 000)	EXPOSURE (DOCTORS)	EXPECTED NUMBER OF LOSSES	EXPECTED LOSS ($)	STANDARD DEVIATION ($)	COEFF. OF SKEWNESS	MAXIMUM PROBABLE LOSS ONE IN 10.0 YEARS ($)	20.0 YEARS ($)	100.0 YEARS ($)
CLASS1	0	250	215.000	1.34	39,072	75,227	2.993	160,419	227,692	380,569
CLASS2	0	250	77.000	.69	20,626	53,964	4.128	113,657	172,766	310,046
CLASS3	0	250	66.000	.70	23,634	59,345	3.917	124,600	187,304	332,704
CLASS4	0	250	10.000	.13	4,500	26,204	8.026	45,800	91,600	236,386
CLASS5	0	250	46.000	.75	20,357	66,755	3.606	139,709	206,652	360,754
CLASS6	0	250	35.000	.74	25,015	61,076	3.810	128,230	191,699	338,324
CLASS7	0	250	51.000	1.00	40,319	80,936	3.055	170,547	243,018	410,534
TOTALS			500.000	5.35	182,404	169,614	1.513	427,306	534,476	765,907

LAYER 2

GROUPS	LIMITS LOWER ($ 000)	LIMITS UPPER ($ 000)	EXPOSURE (DOCTORS)	EXPECTED NUMBER OF LOSSES	EXPECTED LOSS ($)	STANDARD DEVIATION ($)	COEFF. OF SKEWNESS	MAXIMUM PROBABLE LOSS ONE IN 10.0 YEARS ($)	20.0 YEARS ($)	100.0 YEARS ($)
CLASS1	250	1000	215.000	.04	11,003	78,1[illegible]	8.493	110,030	221,660	600,947
CLASS2	250	1000	77.000	.02	5,733	54,155	11.790	57,334	114,668	573,330
CLASS3	250	1000	66.000	.02	7,061	62,591	10.637	70,611	141,222	643,414
CLASS4	250	1000	10.000	.00	1,403	27,904	23.894	15,032	[illegible]	140,[illegible]
CLASS5	250	1000	46.000	.03	9,103	71,249	9.414	91,030	102,059	668,333
CLASS6	250	1000	35.000	.02	7,475	64,405	10.359	74,752	149,503	648,163
CLASS7	250	1000	51.000	.04	13,508	86,970	7.750	135,081	270,163	711,734
TOTALS			500.000	.18	55,367	176,305	3.862	354,204	539,083	966,439

PREPARED BY: HOWARD H. FRIEDMAN
DATE: APRIL 1, 1980

(.10, .05, .01) for the aggregate loss distribution percentiles. In the package, the $1 - \varepsilon$ percentile, L_ε, the point which L has subjective probability ε of exceeding, is called "the maximum probable loss for one in 'ε^{-1}' years". This wording was chosen to be more meaningful to the underwriters who see the main output.

The main output is displayed in Table 7D. Various information about the distribution of aggregate loss for each layer is shown. The display should be self-explanatory to actuaries. Note for example, the amount of "risk" being assumed by the reinsurer as evidenced by the coefficient of skewness: 1.513 for the primary layer versus 3.862 for the excess layer. Or, notice the coefficients of variation: .930 (169,614 ÷ 182,404) for the primary layer versus 3.184 (176,305 ÷ 55,367) for the excess layer. Approximations of the aggregate loss percentiles are in the last three columns.

There are many methods for approximating the percentiles of a distribution. The method used by RISKMODEL is the NP-approximation described by Beard, Pentikäinen and Pesonen (1969 - 2nd ed., 1977). This approximation is given by:

$$(7.2) \qquad L_\varepsilon \doteq E[L] + (\mathrm{Var}[L])^{\frac{1}{2}} \cdot \{z_\varepsilon + \frac{\gamma_1}{6} \cdot (z_\varepsilon^2 - 1)\}$$

where L_ε is minimal such that $\text{Prob}[L>L_\varepsilon] \le \varepsilon$

$z_\varepsilon = \Phi^{-1}(1-\varepsilon)$ for Φ the standard normal (0,1) c.d.f.

$\gamma_1 = \mu_3(L) \div (\text{Var}[L])^{3/2}$, the coefficient of skewness.

A problem with the NP-approximation is that if γ_1 is very large (say $\gamma_1 > 8$), then for certain values of ε, the approximation is much too large. However, there is a natural bound on L_ε which RISKMODEL uses to bound the NP-approximation. This bound is:

$$L_\varepsilon \le \varepsilon^{-1} \cdot E[L] \tag{7.3}$$

The necessity of this Chebyshev-like bound is seen immediately from:

$$\begin{aligned} E[L] &= \int_0^\infty x \cdot dF(x) && \text{since } F(x) = 0 \text{ for } x < 0 \\ &\ge \int_{L_\varepsilon}^\infty x \cdot dF(x) && \text{since } L_\varepsilon \ge 0 \\ &\ge \int_{L_\varepsilon}^\infty L_\varepsilon \cdot dF(x) \\ &= \varepsilon \cdot L_\varepsilon \end{aligned} \tag{7.4}$$

The extreme values of γ_1, which trigger this bound on the NP-approximation seem to occur only when the expected number of loss occurrences is very small. For example, the bound occurs in the example A main output, Table 7D, for the excess layer for each individual class when $\varepsilon = .10$, .05 and sometimes .01; in each case, the expected number of excess losses is less than .05. It does not happen for the overall excess layer where the expected number of losses is .18.

Thus, in certain extreme situations, the NP-approximation may not be very accurate. In fact, there has been quite a discussion in the recent literature regarding the accuracy of the NP-approximation versus its various alternatives. The reasonable alternatives presently include: 1) approximation via simulation, 2) an NP3-approximation which uses the fourth moment of L in addition to the first three and 3) approximation via the 3-parameter gamma distribution. See the argument carried on in Kauppi and Ojantakanen (1969), Seal (1977), Pentikäinen (1977) and Seal (1979) and also the discussion in Cummins and Freifelder (1978).

The reasons to use the NP-approximation are:

1. it is easier to compute than any of its reasonable alternatives.
2. in most situations, it is just as good.
3. it is slightly conservative; that is, L_ε is less than the NP-approximation.

In particular, it is as good as the alternatives for the usual excess-of-loss casualty working cover situation. Beard, Pentikäinen and Pesonen (1977), p. 5, said it well: "Thus it is important not to develop mathematical tools of disproportionate accuracy (and complication) without regard to the context in the problem being solved".

The example B run, Appendix C, has four policy limit groups and four parameters (see p. C1). The reason for grouping by policy limit should be obvious. Again, the limits and parameter matrices have been previously input. Since the proposed coverage is \$400,000 excess of \$100,000, the loss layers of interest are 0 - \$100,000 and \$100,000 - min {\$500,000, policy limit}. The parameters, Table 6A, were discussed in detail in Section VI. The intermediate calculations and the ϵ selection (pp. C2 and C3) are analogous to example A.

The main output is displayed in Table 7E. Again note the "risk" being assumed by the reinsurer as evidenced by the coefficient of skewness: .216 for the primary layer versus .437 for the excess layer. Or again notice the coefficients of variation: .129 (1,247,991 ÷ 9,678,618) for the primary layer versus .287 (641,998 ÷ 2,238,766) for the excess layer. Note that there is much less uncertainty in example B than there was for example A. Since we are using "base frequencies" as explained in Section VI, the expected number of losses in layer 1 are probably understated; the expected loss in layer 1 may also be understated. The estimates for layer 2 have no known systematic bias.

TABLE 7E

EXAMPLE B: P&C INSURANCE COMPANY-GENERAL LIABILITY

LAYER 1

GROUPS	LIMITS LOWER ($ 000)	LIMITS UPPER ($ 000)	EXPOSURE (000)	EXPECTED NUMBER OF LOSSES	EXPECTED LOSS ($)	STANDARD DEVIATION ($)	COEFF. OF SKEWNESS	MAXIMUM PROBABLE LOSS ONE IN 10.0 YEARS ($)	MAXIMUM PROBABLE LOSS ONE IN 20.0 YEARS ($)	MAXIMUM PROBABLE LOSS ONE IN 100.0 YEARS ($)
GL/200	0	100	1175.000	13.58	483,931	219,139	.662	780,360	895,753	1,100,600
GL/250	0	100	1175.000	13.59	483,931	219,139	.662	780,360	885,753	1,100,600
GL/350	0	100	2350.000	27.17	967,862	314,939	.481	1,207,750	1,529,073	1,812,045
GL/500+	0	100	18800.000	217.33	7,742,894	1,070,248	.227	9,140,652	9,572,683	10,411,509
TOTALS			23500.000	271.66	9,678,618	1,247,991	.216	11,307,066	11,808,457	12,780,606

- 450 -

LAYER 2

GROUPS	LIMITS LOWER ($ 000)	LIMITS UPPER ($ 000)	EXPOSURE (000)	EXPECTED NUMBER OF LOSSES	EXPECTED LOSS ($)	STANDARD DEVIATION ($)	COEFF. OF SKEWNESS	MAXIMUM PROBABLE LOSS ONE IN 10.0 YEARS ($)	MAXIMUM PROBABLE LOSS ONE IN 20.0 YEARS ($)	MAXIMUM PROBABLE LOSS ONE IN 100.0 YEARS ($)
GL/200	100	200	1175.000	1.46	77,023	79,995	1.224	190,046	236,486	335,192
GL/250	100	250	1175.000	1.46	91,014	100,994	1.352	235,895	296,023	427,286
GL/350	100	350	2350.000	2.92	213,774	180,223	1.123	466,456	567,054	782,016
GL/500+	100	500	18800.000	23.37	1,856,156	600,305	.486	2,656,854	2,926,009	3,467,655
TOTALS			23500.000	29.21	2,238,766	641,998	.437	3,091,686	3,374,779	3,938,912

PREPARED BY: RALPH M. CELLARS
DATE: OCTOBER 31, 1979

VIII. CONCLUSION

We have described a procedure for estimating the distribution of the aggregate loss for the next coverage year of an excess-of-loss casualty working cover reinsurance treaty. Recall that for both treaty proposals, for each individual loss the reinsurer shares the allocated loss adjustment expense (ALAE) pro rata according to his share of the loss (the reinsurer's unallocated loss adjustment expense is included in his general overhead expense). The ALAE share increases the reinsurer's aggregate loss by 3% to 6% depending upon the line of business and the excess layer. For both examples, we will increase all aggregate loss figures by 5%.

According to the list in Section III, there are four more general items to consider before deciding about the adequacy of the rate offered on example A or before proposing a rate for example B. Without offering complete, elegant solutions, let us briefly consider those items (2) - (4).

Item (2) is the potential distribution of cash flow. Both proposals are fairly typical excess-of-loss casualty working covers which we may assume will have standard monthly or quarterly premium payment patterns and typical long tail casualty loss pay-out patterns. That simple general cash flow models can be constructed should surprise no one who has read the CAS exam materials. In the long run, such general models should be constructed so that any two treaty proposals can be compared to each other. However, even without such models explicitly set up, we can say something about these two treaty proposals. For instance, based upon typical medical malpractice claims-made loss payment patterns, the one year aggregate loss expected values or higher percentiles for example A could be discounted from 10% to 15% on a present value basis with respect to rates-of-return on investments of 5% or greater. Based upon typical casualty loss payment patterns, the discount for example B would be 10% to 20%. The present values of the premium payments for both examples would be discounted around 5%. How this is viewed by the reinsurer depends upon items (3) - (5).

Item (3) is the collection of the reinsurer's various corporate financial parameters and decision-making criteria. Assuming that the reinsurer is at least moderate sized and is in good financial condition, then neither proposal in isolation leads to overwhelmingly complex decision problems; there is nothing unusual or very exciting here. It is highly unlikely that either treaty by itself could hurt such a reinsurer very much. However, the loss results of a whole portfolio of typical medical malpractice treaties, for example, would be correlated and could hurt a lot if priced badly.

Item (4) is the surplus necessary to "support" a treaty from the reinsurer's point-of-view and item (5) is the potential distribution of the rate-of-return on this "supporting surplus". These are very ambiguous but we believe useful concepts. Strictly for illustration, let us define an ad-hoc measure of supporting surplus for our two treaty examples. In each case, we will consider the supporting surplus to be the difference of the 90th percentile of the distribution of aggregate loss and ALAE minus the pure premium (that part of the premium available to pay losses).

The A Doctors' Mutual Insurance Company proposal, example A, is expected to be profitable to the reinsurer based upon the 1980/81 expected aggregate loss of $55,367 in the layer $750,000 excess of $250,000 (Table 7D) and an expected net reinsurance premium of $115,248 (Appendix A, p. A3). But the 90th percentile of the reinsurer's subjective distribution of aggregate loss is $354,284 (Table 7D), over three times the net premium. This is very risky, and our ad-hoc supporting surplus is (1.05 x $354,284) - (.97 x $115,248) = $260,208 (take 3% out of the net premium for overhead expenses). The expected rate-of-return on this supporting surplus is 21% (((.97 x $115,248) - (1.05 x $55,367)) ÷ $260,208). The reinsurer's decision to accept or reject the proposal would be based upon his attitude toward risk and upon the extra premium he wants for assuming such risk.

Example B could be profitable to the reinsurer if he can negotiate a reasonable net rate with the P&C Insurance Company. Exactly what the final rate will be depends upon the two companies' attitudes toward risk, their separate evaluations of the loss potential, the rates that are available for such coverage in the reinsurance marketplace and finally the amount of premium that P&C is collecting from his insureds for the layer $400,000 excess of $100,000. A quick check of the ISO increased limits factors for state B for this coverage, i.e., the premises/operations bodily injury table B (ISO Subline Code 314), indicates that about 15% of P&C's gross general liability premium is collected for this layer. Since the expected excess aggregate loss is $2,238,766 (Table 7E) and the expected gross direct earned premium is $23,500,000 (Table 7E total exposure), there is room to negotiate.

Purely for illustration, suppose that a flat net rate of 12% is negotiated for example B. Then the reinsurer's premium is .12 x $23,500,000 = $2,820,000 and his pure premium is .97 x $2,820,000 = $2,735,400. The 90th percentile of the reinsurer's subjective distribution of aggregate loss is $3,091,686, so our ad hoc supporting surplus is (1.05 x $3,091,866) - $2,735,400 = $511,059. The expected rate-of-return on this supporting surplus is 75% (($2,735,400 - 1.05 x $2,238,766) ÷ $511,059).

If the insurer and the reinsurer disagree strongly on the loss potential, the rate could be negotiated to include a profit commission arrangement by which they would share good years and bad years fairly. Reinsurance contract wording is often very inventive; treaties are custom-made for the particular situation;

the terms are adjusted to suit both parties. This is an example of a fundamental principle of reinsurance: reinsurance works best when it is a long term beneficial partnership between the parties.

We hope you noticed that the models, estimation techniques and decision procedures presented in this paper are not really specific to excess-of-loss reinsurance. They may be useful for pricing any large casualty contracts; with suitable modifications, they are useful for property insurance also. You may have noticed that we have presented no cookbook formulas for pricing reinsurance; the area is too rich in diversity and too interesting for such simplistic nonsense. We consider the work described here as only the beginning of a truly satisfying pricing procedure.

We close by noting that the Bibliography contains some papers on excess reinsurance pricing in addition to those previously mentioned. You will find most of these to be informative and interesting.

BIBLIOGRAPHY

Ammeter, Hans (1955). "The Calculation of Premium-Rates for Excess of Loss and Stop Loss Reinsurance Treaties", Non-Proportional Reinsurance, ed. S. Vajda, Arithbel S.A.

Barile, Andrew J. (1978). Reinsurance . . . A Practical Guide, Interstate Service Corporation.

Beard, R.E., Pentikäinen, T. and Pesonen, E. (1977). Risk Theory, The Stochastic Basis of Insurance, Second Edition. Chapman and Hall.

Benktander, Gunnar (1977). "The Use of Actuarial Techniques in Non-Life Reinsurance: The Underwriting of Risks", Reinsurance Offices Association.

Bühlmann, Hans (1970). Mathematical Methods in Risk Theory, Springer-Verlag.

Cummins, J. David and Freifelder, Leonard R. (1978) "A Comparative Analysis of Alternative Maximum Probable Yearly Aggregate Loss Estimators", The Journal of Risk and Insurance, Vol. 45, No. 1, pp. 27-52.

Ferguson, Ronald (1978). "Actuarial Note on Loss Rating" and reviews by David J. Grady and Gary Patrik, Proceedings of the Casualty Actuarial Society, Vol. 65.

Freifelder, Leonard R. (1976). A Decision Theoretic Approach to Insurance Ratemaking, S.S. Heubner Foundation Monograph, No. 4, Richard D. Irwin, Inc. See also the review by Gary Patrik in The Actuarial Review, Vol. 4, No. 3, and author's reply in The Actuarial Review, Vol. 4, No. 4.

Gerber, Hans U. (1974). "On Additive Premium Calculation Principles", The ASTIN Bulletin, Vol. 7, pp. 215-222.

Hachemeister, Charles A. (1976). "Breaking Down the Loss Reserving Process", Casualty Actuarial Society Loss Reserves Symposium, Chicago, Illinois.

Harding, V. (1968). "The Calculation of Premiums for Excess of Loss Reinsurance of Motor Business", Trans. 18th International Congress of Actuaries, Vol. 2, pp. 445-457.

Hastings, N.A.J. and Peacock, J. B. (1975). Statistical Distributions: A Handbook for Students and Practitioners, J. Wiley & Sons.

Insurance Institute of London, The (1976). Excess of Loss Methods of Reinsurance: Report by Advanced Study Group No. 201.

Johnson, Norman L. and Kotz, Samuel (1969). Distributions in Statistics: Discrete Distributions, Houghton-Mifflin Company.

Johnson, Norman L. and Kotz, Samuel (1970). Distributions in Statistics: Continuous Univariate Distributions - 1. Houghton-Mifflin Company.

Jung, Jan (1963). "On the Use of Extreme Values to Estimate the Premium for an Excess of Loss Reinsurance", The ASTIN Bulletin, Vol 3, Part 2, pp. 178 - 184.

Kauppi, Lauri and Ojantakanen, Pertti (1969). "Approximations to the Generalized Poisson Function", ASTIN Bulletin, Vol. 5, Part 2, pp. 213 - 226.

Kendall, Maurice G. and Stuart, Alan (1969). The Advanced Theory of Statistics, Vol. 1, Hafner Publishing Company.

Miccolis, Robert S. (1977). "On The Theory of Increased Limits and Excess of Loss Pricing", and discussion by Sheldon Rosenberg, PCAS, Vol. 64, pp. 27 - 73.

National Association of Insurance Commissioners (1977). NAIC Malpractice Claims, Vol. 1, No. 4.

National Association of Insurance Commissioners (1978). NAIC Malpractice Claims, Vol. 2, No. 1.

Patrik, Gary (1978). "An Actuarial Procedure for Estimating a Reinsurance Company's IBNR", Proceedings of the Insurance Accounting and Statistical Association, pp. 531 - 534.

Patrik, Gary (1980). "Estimating Casualty Insurance Loss Amount Distributions", submitted to the PCAS.

Pentikäinen, T. (1977). "On the Approximation of the Total Amount of Claims", ASTIN Bulletin, Vol. 9, pp. 281 - 289.

Pratt, John W. (1964). "Risk Aversion in the Small and in the Large", Econometrica, Vol. 32, No. 1 - 2, pp. 122 - 136.

Raiffa, Howard (1968). Decision Analysis: Introductory Lectures on Choices Under Uncertainty, Addison-Wesley.

Reinarz, Robert C. (1969). Property and Liability Reinsurance, Mission Publishing Company.

Savage, Leonard J. (1954). The Foundations of Statistics, John Wiley & Sons, Inc.

Seal, Hilary L. (1977). "Approximations to Risk Theory's F(x, t) by Means of the Gamma Distribution", ASTIN, Vol. 9, pp. 213 - 218.

Seal, Hilary L. (1979). "Letter to the Editor", ASTIN, Vol. 10, p. 130.

Straub, E. (1971). "Estimation of the Number of Excess Claims by Means of the Credibility Theory", the ASTIN Bulletin, Vol. 5, Part 3, pp. 388 - 392.

Verbeek, H. G. (1972). "An Approach to the Analysis of Claims Experience in Motor Liability Excess of Loss Reinsurance", The ASTIN Bulletin, Vol. 6, Part 3, pp. 195 - 202.

Weissner, Edward W. (1978). "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood". Proceedings of the Casualty Actuarial Society, Vol. 65.

APPENDIX A

EXAMPLE A: A DOCTORS' MUTUAL INSURANCE COMPANY

Parameter Selection

(1) Doctor Class	(2) Frequency Offset	(3) Medium Frequency	(4) Severity Offset	(5) Low β	(6) Medium β	(7) High β
1	.90	.0062	1.00	23,640	18,450	18,155
2	1.30	.0090	1.00	23,640	18,450	13,155
3	.65	.0106	.90	23,923	20,106	20,597
4	.80	.0130	.95	25,253	21,224	21,742
5	1.00	.0163	1.05	27,911	23,458	24,031
6	1.30	.0212	.90	23,923	20,106	20,597
7	1.20	.0195	1.15	30,569	25,692	26,320

(1) ISO old class plan.

(2) Selected on the basis of ISO data; the class 1, 2 countrywide mean frequency is selected to be .0385 and the class 3 - 7 countrywide mean frequency is selected to be .0904 for 1/1/81.

(3) The state A frequency offset is selected to be .90; the first year claims-made offset is selected to be .25; the contagion (multiple doctors per incident) is selected to be .80. Together with col. (2), these offset the countrywide mean frequencies in note (2). For example A, the low and high frequencies are selected to be ± 20% of the medium frequencies.

(4) Selected on the basis of ISO data.

(5)- (7) The state A severity offset is selected to be .70; the contagion offset is selected to be 1.25. Together with col. (4), these offset the countrywide β parameters on p.A2.

APPENDIX A

EXAMPLE A: A DOCTORS' MUTUAL INSURANCE COMPANY

General Loss Amount Distribution Model

		Countrywide Loss Amount Parameters: 1/1/81			
		β	δ	t	XP
Physicians	- low	27,017	1.484	1000	.808
(1, 2)	- medium	21,086	1.293	1000	.856
	- high	20,749	1.191	1000	.838
Surgeons	- low	30,378	1.465	1000	.856
(3 - 7)	- medium	25,531	1.278	1000	.886
	- high	26,155	1.189	1000	.895

The parameters are selected based upon ISO medical malpractice data via maximum likelihood estimation - See Patrik (1980). The general loss amount c.d.f. is the 4-parameter Pareto described in Appendix D.

APPENDIX A

EXAMPLE A: A DOCTORS' MUTUAL INSURANCE COMPANY

Estimated Premium: 7/1/80 - 6/30/81

(1) Doctor Class	(2) # in Class	(3) 1980 1M/3M Rate	(4) 1980 1M/3M Premium
1	215	$ 400	$ 86,000
2	77	720	55,440
3	65	1,200	78,000
4	11	1,600	17,600
5	46	2,000	92,000
6	35	2,400	84,000
7	51	3,200	163,200
	500		$576,240

(1) These are older ISO doctor class plan.

(2) Based upon ISO doctor distribution and the estimate of 500 doctors.

(3) First year claims-made rates to be used by A Doctors' Mutual Insurance Company.

(4) The reinsurance net premium is .20 x $576,240 = $115,248.

EXAMPLE B: P&C INSURANCE COMPANY

Average Incurred (Ground-Up) and Occurrence Loss Development
Excess of $75,000 at 1980 Level as of 6/30/79

Exponential Trend Model

Trend Factor	Deflated Retention	Accident Year		18	30	42	54	66	78
2.966	$25,291	1973	Avg. $	NA	141,778	170,039	162,867	159,706	159,117
			#		19	30	45	44	46
2.539	29,540	1974	Avg. $	117,249	134,211	165,289	173,331	183,696	
			#	4	17	39	44	43	
2.174	34,502	1975	Avg. $	92,268	103,421	113,232	127,553		
			#	2	21	28	41		
1.861	40,299	1976	Avg. $	112,482	109,284	109,583			
			#	5	10	24			
1.593	47,069	1977	Avg. $	0	100,650				
			#	0	14				
1.364	54,976	1978	Avg. $	103,172					
			#	3					

(Age in Months: 18, 30, 42, 54, 66, 78)

Average Incurred Age-to-Ultimate Factors

	18 - Ult.	30 - Ult.	42 - Ult.	54 - Ult.	66 - Ult.
Actual*	1.73	1.22	1.04	1.02	1.00
Selected	1.20	1.12	1.05	1.02	1.00

Occurrence (count) Age-to-Age Factors

	18 - 30	30 - 42	42 - 54	54 - 66	66 - 78
Actual**	5.64	1.81	1.34	.98	1.05
Selected	5.22	1.94	1.36	.98	1.05

* based on weighted average incurred
** based on average number of occurrences

APPENDIX B

B1

EXAMPLE B: P&C INSURANCE COMPANY

Average Incurred (Ground-Up) and Occurrence Loss Development
Excess of $75,000 at 1980 Level as of 6/30/79

Econometric Trend Model

Trend Factor	Deflated Retention	Accident Year		Age in Months 18	30	42	54	66	78
2.964	$25,299	1973	Avg. $	NA	141,960	170,257	163,076	159,911	159,321
			#		19	30	45	44	46
2.420	30,992	1974	Avg. $	111,925	140,014	160,026	172,256	183,063	
			#	4	14	38	41	40	
2.012	37,283	1975	Avg. $	100,600	118,962	124,117	131,588		
			#	1	11	18	32		
1.721	43,584	1976	Avg. $	121,726	121,905	102,667			
			#	5	9	23			
1.506	49,817	1977	Avg. $	0	95,201				
			#	0	14				
1.350	55,546	1978	Avg. $	102,150					
			#	3					

Average Incurred Age-to-Ultimate Factors

	18 - Ult.	30 - Ult.	42 - Ult.	54 - Ult.	66 - Ult.
Actual*	1.16	1.12	1.03	1.02	1.00
Selected	1.20	1.12	1.05	1.02	1.00

Occurrence (count) Age-to-Age Factors

	18 - 30	30 - 42	42 - 54	54 - 66	66 - 78
Actual**	4.80	2.06	1.37	.98	1.05
Selected	5.22	1.94	1.36	.98	1.05

* based on weighted average incurred
** based on average number of occurrences

EXAMPLE B: P&C INSURANCE COMPANY

WEIBULL DISTRIBUTION
ACCIDENT YEAR AGE-TO-AGE FACTORS*

MEAN MON.	PARAMETERS SCALE	SHAPE	18 TO 30	30 TO 42	42 TO 54	54 TO 66	66 TO 78	78 TO 90	90 TO 102	102 TO 114	114 TO ULT
27	30.000	2.500	4.050	1.798	1.218	1.042	1.004	1.000	1.000	1.000	1.000
28	31.000	2.500	4.117	1.863	1.245	1.053	1.006	1.000	1.000	1.000	1.000
28	32.000	2.500	4.179	1.886	1.272	1.064	1.009	1.001	1.000	1.000	1.000
29	33.000	2.500	4.235	1.927	1.299	1.077	1.013	1.001	1.000	1.000	1.000
30	34.000	2.500	4.288	1.966	1.326	1.091	1.017	1.002	1.000	1.000	1.000
31	35.000	2.500	4.336	2.003	1.352	1.105	1.022	1.003	1.000	1.000	1.000
32	36.000	2.500	4.381	2.030	1.378	1.120	1.027	1.004	1.000	1.000	1.000
33	37.000	2.500	4.422	2.071	1.403	1.135	1.034	1.005	1.001	1.000	1.000
34	38.000	2.500	4.460	2.102	1.428	1.151	1.041	1.007	1.001	1.000	1.000
35	39.000	2.500	4.495	2.131	1.452	1.167	1.047	1.010	1.001	1.000	1.000
27	30.000	2.750	4.679	1.897	1.219	1.031	1.002	1.000	1.000	1.000	1.000
28	31.000	2.750	4.760	1.954	1.246	1.041	1.003	1.000	1.000	1.000	1.000
28	32.000	2.750	4.839	2.007	1.278	1.052	1.005	1.000	1.000	1.000	1.000
29	33.000	2.750	4.912	2.061	1.310	1.065	1.007	1.000	1.000	1.000	1.000
**30	34.000	2.750	4.978	2.110	1.342	1.079	1.010	1.001	1.000	1.000	1.000
31	35.000	2.750	5.038	2.157	1.374	1.095	1.014	1.001	1.000	1.000	1.000
32	36.000	2.750	5.094	2.201	1.406	1.110	1.019	1.002	1.000	1.000	1.000
33	37.000	2.750	5.145	2.242	1.437	1.128	1.024	1.003	1.000	1.000	1.000
34	38.000	2.750	5.193	2.281	1.468	1.145	1.031	1.004	1.000	1.000	1.000
35	39.000	2.750	5.236	2.310	1.497	1.164	1.038	1.005	1.000	1.000	1.000

* Expected value of annual age-to-age factors that would be generated if the report lags of losses occurring in each month are distributed according to the Weibull distribution with specified parameters.

** Report lag c.d.f. selected with respect to both trend models.

APPENDIX B

APPENDIX B

EXAMPLE B: P&C INSURANCE COMPANY

Number of IBNR Occurrences Excess of $75,000 at 1980 Level as of 6/30/79

Total number of IBNR occurrences excess of $75,000 for accident years 1973 - 78 as of 6/30/79 are estimated using the method described in Patrik (1978).

$$\text{Total IBNR} = \frac{(\text{Known})\cdot\omega}{1-\omega}$$

= 87.2 and 80.4 with respect to the exponential and econometric trend models, respectively.

where

Known = total number of known occurrences excess of $75,000 for accident years 1973 - 78 as of 6/30/79.

= 171 and 158 with respect to the exponential and econometric trend models, respectively.

$$\omega = \frac{\sum_m EP_m\cdot[1 - W(x_m)]}{EP}$$

= .3375 for months m such that $1/73 \leq m \leq 12/78$

EP_m = monthly exposure base, in this case GL gross direct earned premium at present rates, for $1/73 \leq m \leq 12/78$.

$EP = \sum_m EP_m$ for $1/73 \leq m \leq 12/78$

APPENDIX B

$W(\cdot)$ = selected report lag c.d.f. (see p.B3).

x_m = maximum observable report lag; that is, for accident month m the difference between 6/30/79 and the mid-point of m.

Letting IBNR(x;6/30/79) denote the number of IBNR occurrences for accident year x as of 6/30/79, the total IBNR is allocated to accident year x using the formula:

$$\text{IBNR}(x;6/30/79) = R \cdot \sum_m EP_m \cdot [1 - W(x_m)]$$

where

$$R = \frac{\text{Known + Total IBNR}}{EP},$$

$1/x \leq m \leq 12/x$, and $x = 73,\ldots,78$.

The assumptions underlying this IBNR method are:

1. homogeneous coverage groups
2. the ratio of ultimate number of occurrences to earned exposure is constant and independent of time
3. the report lag distribution does not vary with occurrence date.

EXAMPLE B: P&C INSURANCE COMPANY

Excess and Base Frequencies and Excess IBNR
by Accident Year at 1980 Level

Exponential Trend Model

Accident Year	Present Level Gross Direct Earned Premium (000)	Occurrences Excess of $75,000 Known (6/30/79)	IBNR** (6/30/79)	Ultimate	Frequency Excess Of $75,000	Base Frequency* c.d.f.(1)	c.d.f.(2)
1973	$24,524	46	0	46.0	.0019	.0108	.0128
1974	21,860	43	.5	43.5	.0020	.0114	.0135
1975	19,435	41	3.2	44.2	.0023	.0131	.0155
1976	19,685	24	12.5	36.5	.0019	.0108	.0128
1977	21,137	14	28.6	42.6	.0020	.0114	.0135
1978	22,701	3	42.4	45.4	.0020	.0114	.0135
Selected	-	-	-	-	-	.0108	.0135

* Base frequency = excess frequency divided by the probability of an occurrence exceeding $75,000 for loss amount c.d.f.(1) and c.d.f.(2).

** Based on the IBNR method described in Appendix B, pp. B4 and B5.

EXAMPLE B: P&C INSURANCE COMPANY

Excess and Base Frequencies and Excess IBNR
by Accident Year at 1980 Level

Econometric Trend Model

Accident Year	Present Level Gross Direct Earned Premium (000)	Occurrences Excess of $75,000 Known (6/30/79)	IBNR** (6/30/79)	Ultimate	Frequency Excess Of $75,000	Base Frequency* c.d.f.(3)	c.d.f.(4)
1973	$24,524	46	0	46.0	.0019	.0101	.0104
1974	21,860	40	.4	40.4	.0018	.0096	.0099
1975	19,435	32	2.9	34.9	.0018	.0096	.0099
1976	19,685	23	11.5	34.5	.0018	.0096	.0099
1977	21,137	14	26.4	40.4	.0019	.0101	.0104
1978	22,701	3	39.2	42.2	.0019	.0101	.0104
Selected	-	-	-	-	-	.0096	.0104

* Base frequency = excess frequency divided by the probability of an occurrence exceeding $75,000 for loss amount c.d.f.(3) and c.d.f.(4).

** Based on the IBNR method described in Appendix B, pp. B4 and B5.

APPENDIX C

EXAMPLE B: P&C INSURANCE COMPANY RISKMODEL RUN

```
      RISKMODEL
DO NOT PANIC IF YOU MAKE AN ERROR WHILE INPUTTING,
OPPORTUNITY TO CHANGE LATER.

ENTER MAJOR GROUP NAMES AS FOLLOWS.   /GRP1/GRP2..........
NOTE: MUST BE IN QUOTES.   FOR MORE THAN 1 LINE OF INPUT, USE .0
0:
      '&GL/200&GL/250&GL/350&GL/500+'

ENTER THE NUMBER OF PARAMETERS,  E.G.   5
0:
      4

DO YOU WISH TO  (1) INPUT VECTOR OF LIMITS,  OR
(2) USE MATRIX OF LIMITS PREVIOUSLY CREATED. 1 OR 2.
0:
      2
ENTER THE NAME OF THE MATRIX OF LIMITS PREVIOUSLY CREATED
NOTE: NAME SHOULD HAVE PREFIX LIM
LIMP&C

DO YOU WISH TO SEE THE LIM MATRIX.  Y OR N
Y
                        LIMITS
              LOWER                UPPER               INDEX
                  0                100000                 1
             100000                200000                 0
                  0                100000                 1
             100000                250000                 0
                  0                100000                 1
             100000                350000                 0
                  0                100000                 1
             100000                500000                 0

DO YOU WISH TO MAKE ANY CHANGES IN THE LIM MATRIX.  Y OR N
N


DO YOU WISH TO
(1) INPUT VECTOR OF PARAMETERS FOR THE FIRST SUBGROUP  OR
(2) USE MATRIX OF PARAMETERS PREVIOUSLY CREATED.  1 OR 2
0:
      2
ENTER THE NAME OF THE MATRIX OF PARAMETERS PREVIOUSLY CREATED
NOTE: NAME SHOULD HAVE PREFIX PAR
PARP&C

DO YOU WISH TO SEE THE PAR MATRIX.  Y OR N
Y
```

(The PAR matrix is displayed in Table 6A)

APPENDIX C

EXAMPLE B: P&C INSURANCE COMPANY RISKMODEL RUN

DO YOU WISH TO MAKE ANY CHANGES IN THE PAR MATRIX. Y OR N
N

GROUPS AND PARAMETER INPUT COMPLETED
TO PROCESS INTERMEDIATE CALCULATIONS, HIT EXECUTE

DO YOU WISH TO PRINT THE INTERMEDIATE CALCULATIONS,
P[S>A],P[S>B],E[S],E[S*2],E[S*3]. Y OR N.
Y

INTERMEDIATE CALCULATIONS USED THROUGHOUT MOMENT CALCULATIONS

LAYER 1

GROUPS		P[S>A]	P[S>B]	E[S]	E[S*2]	E[S*3]
GL/200	1	1.000	.114	2.544E04	1.270E09	8.047E13
GL/200	2	1.000	.095	2.394E04	1.147E09	7.112E13
GL/200	3	1.000	.122	2.591E04	1.312E09	8.384E13
GL/200	4	1.000	.118	2.568E04	1.292E09	8.225E13
GL/250	1	1.000	.114	2.544E04	1.270E09	8.047E13
GL/250	2	1.000	.095	2.394E04	1.147E09	7.112E13
GL/250	3	1.000	.122	2.591E04	1.312E09	8.384E13
GL/250	4	1.000	.118	2.568E04	1.292E09	8.225E13
GL/350	1	1.000	.114	2.544E04	1.270E09	8.047E13
GL/350	2	1.000	.095	2.394E04	1.147E09	7.112E13
GL/350	3	1.000	.122	2.591E04	1.312E09	8.384E13
GL/350	4	1.000	.118	2.568E04	1.292E09	8.225E13
GL/500+	1	1.000	.114	2.544E04	1.270E09	8.047E13
GL/500+	2	1.000	.095	2.394E04	1.147E09	7.112E13
GL/500+	3	1.000	.122	2.591E04	1.312E09	8.384E13
GL/500+	4	1.000	.118	2.568E04	1.292E09	8.225E13

APPENDIX C

EXAMPLE B: **P&C** *INSURANCE COMPANY RISKMODEL RUN*

LAYER 2

GROUPS		P[S>A]	P[S>B]	E[S]	E[S*2]	E[S*3]
GL/200	1	.114	.029	1.147E04	1.623E09	2.387E14
GL/200	2	.095	.025	9.503E03	1.342E09	1.972E14
GL/200	3	.122	.031	1.233E04	1.747E09	2.572E14
GL/200	4	.118	.031	1.195E04	1.692E09	2.490E14
GL/250	1	.114	.017	1.414E04	2.217E09	3.718E14
GL/250	2	.095	.015	1.171E04	1.834E09	3.074E14
GL/250	3	.122	.018	1.522E04	2.391E09	4.015E14
GL/250	4	.118	.018	1.474E04	2.315E09	3.886E14
GL/350	1	.114	.007	1.705E04	3.073E09	6.254E14
GL/350	2	.095	.007	1.418E04	2.562E09	5.233E14
GL/350	3	.122	.008	1.837E04	3.316E09	6.757E14
GL/350	4	.118	.008	1.780E04	3.214E09	6.550E14
GL/500+	1	.114	.003	1.893E04	3.848E09	9.493E14
GL/500+	2	.095	.003	1.507E04	3.259E09	8.154E14
GL/500+	3	.122	.003	2.037E04	4.144E09	1.021E15
GL/500+	4	.118	.003	1.977E04	4.026E09	9.944E14

TO PROCESS MORE INTERMEDIATE CALCULATIONS, HIT EXECUTE

ENTER EPSILON(S) FOR WHICH PROB(LOSS>MAX. PROB. LOSS) = EPSILON. (0<ε≤.5)
D:
.1 .05 .01
NOW FOR THE FINAL PRINTOUT

ENTER COMPANY NAME
EXAMPLE B: P&C INSURANCE COMPANY-GENERAL LIABILITY

ENTER YOUR NAME (EG. J. SMITH)
RALPH M. CELLARS

ENTER TODAY'S DATE (EG. JAN. 1, 1979)
OCTOBER 31, 1979

ENTER IN PARENTHESIS AND QUOTES A SEVEN CHARACTER NAME FOR THE UNITS
(E.G. '(DOCTORS)' OR '__(BEDS)_') OF EXPOSURE CENTERED IN 9 SPACES .
D:
' (000) '

ADJUST PAPER TO TOP OF NEW PAGE & HIT EXECUTE

(The main output is displayed in Table 7E)

APPENDIX D

Probability Distribution Definitions

Negative Binomial

density: $f(x|p,\alpha) = \binom{\alpha+x-1}{\alpha} p^{\alpha}(1-p)^{\alpha}$ for $x = 0,1,2,\ldots.$

where $p,\alpha > 0$.

This is our basic model of the loss occurrence (count) process. Note, if $\mathrm{Var}[N] \div E[N] = 1$, then RISKMODEL assumes that the occurrence process is Poisson with $\lambda = E[N]$.

Four Parameter Loss Amount Distributions

$$\text{c.d.f:}\quad G_S(x|\alpha,\beta,t,XP) = \begin{cases} \dfrac{XQ}{H(t|\alpha,\beta)} \cdot H(x|\alpha,\beta) & \text{for } 0<x\leq t \\ XQ + XP\cdot\{H(x|\alpha,\beta)-H(t|\alpha,\beta)\} & \text{for } x>t \end{cases}$$

where $t \geq 0$, $0 < XP \leq 1$

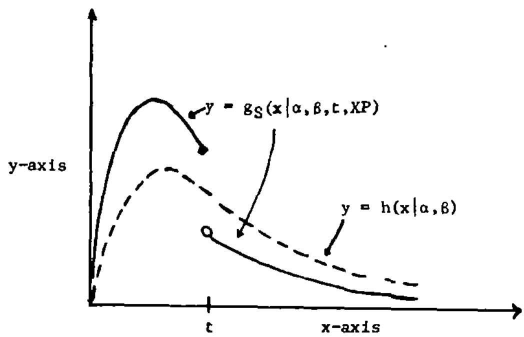
$$XQ = 1 - XP\cdot\{1 - H(t|\alpha,\beta)\}$$

$H(x|\alpha,\beta)$ is some c.d.f. for $x > 0$ with parameters (α,β). RISKMODEL's present library of choices for $H(\cdot|\alpha,\beta)$ are (1) = lognormal, (2) = Pareto and (3) = Weibull. Definitions of each of these distributions are given below.

$$\text{density:}\quad g_S(x|\alpha,\beta,t,XP) = \begin{cases} \dfrac{XQ}{H(t|\alpha,\beta)}\, h(x|\alpha,\beta) & \text{for } 0<x\leq t \\ XP\cdot h(x|\alpha,\beta) & \text{for } x>t \end{cases}$$

APPENDIX D

A graph of the density $g(\cdot|\alpha,\beta,t,XP)$ in general looks like:



(1) Lognormal

c.d.f: $H(x|\mu,\sigma^2) = \Phi\left(\frac{\log x - \mu}{\sigma}\right)$ for $0 < x < \infty$

where $\Phi(\cdot)$ is the standard normal (0,1) c.d.f. and $\sigma > 0$, $-\infty < \mu < \infty$.

density: $h(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\{-\frac{(\log x - \mu)^2}{2\sigma^2}\}$

(2) Pareto

c.d.f: $H(x|\beta,\delta) = 1 - \left(\frac{\beta}{x+\beta}\right)^\delta$ for $x \geq 0$

where $\beta, \delta > 0$.

density: $h(x|\beta,\delta) = \delta\beta^\delta(x+\beta)^{-\delta-1}$

APPENDIX D

(3) Weibull

c.d.f: $H(x|\beta,\delta) = 1 - \exp\{-\left(\frac{x}{\beta}\right)^{\delta}\}$ for $x \geq 0$

where $\beta,\delta > 0$

density: $h(x|\beta,\delta) = \delta\beta^{-\delta}x^{\delta-1}\exp\{-\left(\frac{x}{\beta}\right)^{\delta}\}$

For more details on probability distributions, see Hastings and Peacock (1975) or Johnson and Kotz (1969,1970).

PRICING EXCESS-OF-LOSS CASUALTY WORKING COVER

REINSURANCE TREATIES

by Gary Patrik and Russell John

Discussion by Jerry A. Miccolis

GENERAL COMMENTS

This is an interesting paper. It presents a progress report on the analytical approach one large reinsurer is developing toward the pricing of excess casualty coverage. The approach is an analytical one, in that pricing decisions are made on the basis of information generated by a theoretical pure premium distribution fitted to sample data.

The authors illustrate their techniques via two examples: one a new doctors' mutual (for which there is precious little historical data) and the other an excess-of-loss treaty between a reinsurer and a large primary insurer (for which there is a wealth of detailed pricing information).

A note on format. The authors present their work in phases: description of the coverages, the pricing approach, the model, parameter estimation, results, and conclusions. Within most phases, the two examples are presented separately with much of the technical detail left to appendices. The order of presentation is a matter of personal preference, but I had a much easier time following the flow of the paper by reorganizing it so that I could trace the complete development of first one example through all phases, and then the other.

PRICING PHILOSOPHY

In Section III, the authors mention five items to consider in pricing a reinsurance treaty: 1) the distribution of aggregate loss of the treaty, 2) the distribution of the cash flow of the treaty, 3) a number of corporate criteria (including other treaties in the reinsurer's portfolio, surplus, assets, investment opportunities, corporate goals, and corporate views on risk vs. rate-of-return), 4) "needed surplus" to support the treaty, and 5) distribution of the rate of return on "needed surplus" for each treaty in the reinsurer's portfolio. The paper concentrates on item (1), the distribution of aggregate loss to the reinsurer under the treaty, citing it as "the least ambiguous and most important" item of the five.

I balk slightly at the term "most important". I would select any of the other four items as being more important than item (1), and I suppose that when the perfect pricing model is someday developed, all five items will be thoroughly treated. However, given the present state of our art/science, I grant that knowledge of item (1) is a prerequisite to intelligent formulation of a model treating the latter four items, and in this sense, then, it may be the most important, and deserves our current attention.

AGGREGATE LOSS MODEL

The conceptual meat of the paper is contained in Section IV. The portfolio is assumed to consist of several groups of "independent" risks. (I'll explain the use of the quotation marks shortly.) *Each group has its own distribution of number of claims, and its own* distribution of size of loss for each claim. Further, the specification of these distributions is contained in a "parameter vector", θ, for each group. This is a convenient formulation, since anything that might cause the losses in different groups to move together (e.g., inflation) can be parameterized and thrown into the parameter vector. This allows one to state that the conditional distributions (given θ) for all groups are mutually independent (and hence the quotes above). The authors *show how the necessary conditional* distributions are derived and their moments computed, and then describe how to weight these moments together to arrive at the moments of the unconditional distribution of aggregate loss for the entire portfolio.

SOME TECHNICAL POINTS

There are some errata in the version of the paper that I received. They are itemized in the Appendix following.

Below are some random thoughts of a somewhat technical nature arranged in no particular order:

Section IV

1. It should be noted that unless some grouping of the portfolio is found such that the authors' three assumptions preceding equation (4.3) are met (or at least approached), then there is no advantage to grouping.

2. A simple description of the convolution concept before equation (4.3) might be useful to the lay reader.

3. I'm not sure the presentation would suffer if the notion of "cumulants" was never introduced. Equations (4.5) could be derived without them.

4. Equations (4.7) might deserve derivation in an appendix.

Section V

5. The next-to-last paragraph, last sentence, mentions low, medium and high loss-amount c.d.f.'s. On what basis are these c.d.f.'s characterized low, medium and high? (unlimited mean? coefficient of variation?)

Section VII

6. It is interesting to note that the structure function (i.e., the subjective distribution function of the parameter vector θ) in Example A does not permit much mixing of the frequency and

severity distributions. That is, the "low" claim-count c.d.f. always occurs in conjunction with the "low" loss-amount c.d.f. and similarly for the medium and high c.d.f.'s.

7. To a casual reader of the risk theory papers in the authors' bibliography, it would appear controversial that Patrik and John claim good results for the NP-approximation when the coefficient of skewness is fairly large (i.e., $2 < \gamma_1 < 8$). I think further elaboration by the authors on their position and its apparent conflict with the views of some of risk theory's pioneers would be extremely enlightening.

8. I wonder if a simulation approach would not produce more cost-effective results. In particular, I wonder if it would eliminate recourse to the "Chebyshev-like bound" of equation (7.3).

9. Despite the above comments, I certainly agree with the authors that too much concern over an approximation technique may miss the point. There is so much opportunity for error (the specification of trend and loss development, the choice of the general form of the c.d.f.'s, the use of broad industry data in Example A) that perhaps nothing more than ballpark estimates should be strived for.

Appendix A

10. Page A2: One would expect a smooth progression of parameters for the loss-amount c.d.f. as one moves from low to medium to high. This is not the case for the XP parameter for physicians nor

for the β parameter for surgeons. A rationale for these apparent reversals might be instructive.

Section VIII

11. In the discussion of item (2), mention is made of discounting the future cash flow. I think treatment of this topic is incomplete without consideration of the potentially offsetting phenomenon of inflation on outstanding losses.

12. The discussion of items (4) and (5) contain some ad-hoc measures of supporting surplus and the expected return on such surplus. These measures are elegant in their simplicity and usefulness.

13. The next-to-last paragraph claims that the paper has application beyond excess-of-loss reinsurance. I'd like to issue a warning against using the model (in particular the four-parameter loss amount c.d.f. of Appendix D) for pricing coverage at limits near the truncation point, t. The four-parameter c.d.f. was derived in the context of increased limits pricing where the truncation point was well below basic limits. In this context, the shape of the c.d.f. to the left of t is immaterial, and the form chosen in Appendix D for $G_s(x|\alpha, \beta, t, XP)$ for $x \leq t$, is as arbitrary as any other choice. Indeed, all that matters is that $G_s(x|\alpha, \beta, t, XP)$ reach XQ by the time x reaches t from the left, and the route $G_s(\cdot)$ takes to get to XQ is quite irrelevant. As it happened, the ISO Increased Limits Subcommittee only decided to use a truncation point

in the first place because no theoretical c.d.f. could be found to fit empirical data from first dollar. Further, since in increased limits pricing the concern is with the "tail" of the distribution, it was only necessary to find a distribution which fit the empirical data to the right of a chosen truncation point. This background should be kept in mind when applying the authors' model to anything other than increased limits or excess-of-loss pricing.

SUMMARY

This is an important paper, as are all intelligent attempts at modelling an insurance process. It takes some high-powered techniques and applies them in a practical way. It claims application for risk theory techniques beyond the boundaries imposed by the originators of those techniques. In the "spirit of a Call for Papers" it should generate much response among actuaries and hopefully some suggestions for future enhancement.

I commend the authors on their progress thus far.

Appendix
Sheet 1

ERRATA

1. Section IV, first paragraph: the first reference should be "Bühlmann (1970)".

2. Section IV, group definition: the passage reads, "... For example A, our groups will be defined by year of coverage and ISO doctor class ...". Since the authors are only dealing with the first year of coverage in example A, the groups are defined solely by ISO doctor class.

3. Section IV, definition of cumulant following equation (4.4): "... the moment generating function of L evaluated at 0 ..." should read "... the moment generating function of L given θ evaluated at 0 ...".

4. Section IV, following equation (4.10): "When θ is unknown, equations (4.3) - (4.7) usually no longer hold. In particular, equation (4.5) now holds only for the first moment ..." should read "When θ is unknown, the <u>unconditional</u> <u>counterparts</u> to equations (4.3) - (4.7) usually do not hold. In particular, equation (4.5) would hold <u>unconditionally</u> only for the first moment ...".

5. Section IV, equation (4.14): As it stands, the equation imparts a positive probability to the event of a negative claim. The equation should read:

Appendix
Sheet 2

$$G_x\ (x|\theta) = \begin{cases} 0 & \text{if } x < 0 \\ G_s(x + r|\theta) & \text{if } 0 \leq x < b - r \\ 1 & \text{if } b - r \leq x \end{cases}$$

6. Section V, next-to-last paragraph: The reference, "NAIC (1977, 1979)" should read "NAIC (1977, 1978)".

7. Appendix A, page A1: Footnotes (2) and (3) are confusing and do not seem to match their respective columns.

8. Appendix D, page D1: The specification of the Negative Binomial density function is wrong. It should be:

$$f(x|p,\alpha) = \binom{\alpha + x - 1}{x} p^{\alpha}\ (1 - p)^{x} \quad \text{for } x = 0,\ 1,\ 2\ \ldots$$

where $0 < p < 1$ and $\alpha = 1, 2, 3 \ldots$

The density may be generalized to the case of non-integer α as follows:

$$f(x|p,\alpha) = \frac{\Gamma(x + \alpha)}{x!\ \Gamma(\alpha)}\ p^{\alpha}\ (1 - p)^{x} \qquad \text{for } x = 0,\ 1,\ 2\ \ldots$$

where $0 < p < 1$ and $\alpha > 0$

Appendix
Sheet 3

9. Bibliography: The following reference is missing:

Weissner, Edward (1978) "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood", PCAS, Volume 65.

TITLE: EXPERIENCE RATES AS ESTIMATORS: A SIMULATION OF THEIR BIAS AND VARIANCE

AUTHOR: Mr. James N. Stanard

Mr. Stanard is Director, Treaty Marketing with Prudential Reinsurance Company. He received his B. A. from Lehigh University and his FCAS in 1977. Jim is a member of the CAS Examination Committee.

REVIEWER: Mr. John Robertson

Mr. Robertson is Assistant Actuary with Fireman's Fund Insurance Company. John received his B. A. degree from Howard and his M. A. from the University of California (Berkeley). He received his ACAS in 1979 and is a member of the American Academy of Actuaries and the Mathematics Association of America.

Using an individual insured's own past loss experience to arrive at its rate is a procedure that is used in many different areas of insurance. In addition to the formal individual risk rating plans, ad hoc procedures of this type are used in large risk departments of primary companies, excess and surplus lines companies, treaty and facultative reinsurers, and by various types of insurance consultants.

The purpose of this paper is to discuss the concepts of bias and variance of experience rating procedures[1], and illustrate these concepts by using a computer simulation model to examine the properties of some simple experience rating techniques. We will also discuss the effect that the misestimation of an insured's true loss potential has on the "risk" that the insurer faces. The rating techniques used are not represented as being the best available -- however, the paper presents some useful results concerning the superiority of certain types of techniques.

EXPERIENCE RATES AS ESTIMATORS

View the loss process as follows: a given insured's losses during an accident year "a" are random variables drawn from some probability distribution determined by a vector of parameters θ_a. Let θ represent a vector containing all the parameters from the first accident year of the experience period thru the year to be rated (denoted y). So

$$\theta = (\theta_1, \ldots, \theta_y)$$

1. For the purposes of this paper, define "experience rate" as a rate quoted to a given insured where the expected losses portion of the rate is wholly or predominantly determined by the insured's own loss experience over the past several years. Note that the term insured here could refer to anything from an individual auto to an entire insurance company (under a treaty reinsurance agreement).

Let X be a vector representing the insured's known loss experience during the experience period. X is a random sample drawn from the distributions determined by θ.

Let the ultimate losses that a particular insured will have for the policy period to be rated be a random variable "L". The purpose of the experience rate is to give the "best" estimate of E(L) [2]. E(L) is some function of the θ_y, whereas the experience X was drawn from distributions determined by $\theta_1,\ldots,\theta_{y-1}$. In order for X to be useful in estimating E(L), there must be some relationship between $\theta_1,\ldots,\theta_{y-1}$ and θ_y.

The simplest assumption would be that $\theta_1 = \ldots = \theta_y$, that is that an insured's loss potential is constant over the experience period. A more refined model would be that the severity and frequency componants of the θ_i's would be influenced by inflationary trends and by changes in a measurable exposure base[3], and that, after proper adjustments for these, the parameters would be stable over time.

The experience rating procedure is an estimator[4] of E(L); it is some function "R" of the insured's past known loss and exposure information X [5]. A perfect experience rating system would be a function R such that R(X) = E(L). However, X is also a random variable, so ful-

2. This paper will only consider estimates of E(L). In real life cases, we might want estimates of other attributes of the distribution of L, such as Var(L) or 95% percentile of L.
3. Such as number of cars in a commercial fleet, or subject premium in a reinsurance treaty.
4. An estimator is a function of a random sample and is therefore a random variable; an estimate is the result of the estimator function applied to a particular realization of the random variable, and is therefore itself a particular number.
5. Consider X to be a vector containing all pertinant rating information.

filling this condition is not possible, except by chance. We can, however, hope that $R(X)$ is an unbiased estimator of $E(L)$, that is, that $E(R(X)) = E(L)$.

We would also like $R(X)$ to be close of $E(L)$, on the average. One common way of expressing this is to minimize $E(R(X) - E(L))^2)$, the mean square error (MSE), which for an unbiased estimator is equivalent to minimizing $Var(R(X))$. For many simple statistical models, the form of estimator R that satisfies these criteria can be explicitly calculated. This is referred to as a UMVU[6] estimator.

For large samples, the Maximum Likelihood Estimator (MLE) usually satisfies these properties (asymptotically). However, there are reasons why we cannot always use the MLE, the main one being that in order to calculate it we must explicitly know the forms of the probability distributions that generate X. Of course, we can specify a model of the process that we believe is "reasonable" (as is done later in this paper), but there still are several problems. First, the MLE can be very difficult to calculate; second, although it is known to have good properties for large samples, it may be a bad estimator for smaller samples (it is usually biased); third, while it may be a good estimator if the model we assume is in fact the true one, it may be a bad estimator for a different model -- that is, it may not be robust.

6. Uniform Minimum Variance Unbiased.

The approach taken in this paper is to take several ad hoc (but hopefully reasonable) estimation techniques and examine their properties by a computer simulation model. Briefly, for an individual insured, the computer generates several accident years of known loss experience (X_i for the ith trial) from distributions with fixed parameters. It then applies several rating techniques to this set of known losses, arriving at several different estimates of E(L). The estimates and the actual ultimate losses are stored. This whole process (generating experience, then calculating estimates) is repeated several hundred times -- using the same underlying distributions and parameters. It can then be determined how well the estimates $R(X_i)$ fared as "guesses" of E(L), and which estimator function R does the best[7].

COMPUTER MODEL

An individual insured's past experience was "rerun" several hundred times in order to see how the results of a single rating method would be distributed.

Each iteration produced a set of loss experience for six accident years -- a five year experience period to rate from and the experience for the year to be rated (denoted y = 6). Not only was the ultimate experience generated for each of these years, but also the portion of it that would be known at any point in time.

7. E(L) can in principle be calculated explicitly from θ. However, for the loss generating model that was used, the calculation is quite complex, so the actual loss outcomes L_i were used to estimate E(L). The standard errors on these estimates were small compared with the standard errors of the estimates of E(R).

A single accident year for a single iteration was generated as follows[8]:

A random number of losses, N, was drawn from a Normal[9] distribution with mean = 40, variance = 60.

For each of the N claims, the following random variables were drawn:

M_i = Date of loss within year (Uniform with minimum = 0, maximum = 1)

Q_i = Report Lag (waiting time between accident date and report date) (Exponential with mean = 1.5 years)

All experience was viewed as being analyzed as of year-end, so a claim would first become known in $\lceil(M_i + Q_i - 1)$ years after the accident year[10].

P_i = Payment Lag (waiting time between report date and payment date) (Exponential with mean = 1 year)

So $\lceil(M_i + Q_i + P_i - 1)$ is the number of years after the accident year that the claim is paid[11]. Let "a" denote the accident year, a = 1,...,y. Then $\lceil(M_i + Q_i + P_i + a - 1)$ is the year of payment of the claim, where year 1 is the first year of the experience period.

8. The computer model allows the choice of several different distributions with arbitrary parameters. The distributions and parameters specified here were the principle, but not only ones, that were used.
9. The normal distribution was chosen as an approximation for the negative binomial, which is more difficult to simulate. Also, N was restricted to be between 1 and 65.
10. The APL symbol "$\lceil$", referred to as "ceiling", means "the smallest integer greater than", Note that if $M_i + Q_i < 1$ the claim is reported during the accident year, "zero" years after the accident year.
11. Note that the maximum value allowed was 10 years.

An inflation index $I(M_i + Q_i + P_i + a - 1)$ of 8% per year (others were tested as will be explained in the results), from year 1 until the year of payment was assumed to affect the expected value of the payment distribution.

The random payment amount of C_i was drawn from a Lognormal distribution with $\mu = 8 + \ln I(M_i + Q_i + P_i + a - 1)$, and $\sigma^2 = 2.5$. This means that the mean and standard deviation trended at 8% per year.

So far, the number of claims, and (for each of these claims) the report date, the payment date and final payment amount have been determined. The last thing to do is set the reserve on each open claim. Each reserve was set as an unbiased guess of what the claim would settle for, if it closed in the year for which the reserve was being set.

For each claim that was reported but unpaid for at least a year, a random Reserve Error, V_i, was drawn from a Lognormal distribution with mean = 1, and variance = 2. This was multiplied by the final payment amount and the result was trended backwards from the payment year to the year for which the reserve is being set. Two things are important to note:

1. The reserve error is only chosen once for each claim, regardless of how many years it remains open, so the reserve, once set, will mearly be updated each year for inflation, and 2. this system leads to under reserving -- by the amount of future inflation[12].

12. A method of setting reserves at V times the ultimate payment, which does not lead to under reserving, was tested, but it made no significant difference in the results.

The known loss amount at time "t" on i^{th} loss from accident year "a" =

$$K_i(a,t) = \begin{cases} 0 & \text{if } M_i + Q_i > t \\ C_i \, V_i \, \dfrac{I(a-1+t)}{I(a-1+M_i+R_i+P_i)} & \text{if } M_i + Q_i \leq t < M_i + Q_i + P_i \\ C_i & \text{if } M_i + Q_i + P_i \leq t \end{cases}$$

So the actual ultimate losses $L = \sum_{i=1}^{N} C_i$.

The full experience matrix known at the beginning of year y for an insured would be

$$\begin{pmatrix} \sum_{i=1}^{N_1} K_i(1,1) & \cdots & \sum_{i=1}^{N_1} K_i(1,y-a) \\ \vdots & \cdot^{\cdot^{\cdot}} \quad \cdot^{\cdot^{\cdot}} & 0 \\ & & \vdots \\ \sum_{i=1}^{N_{y-1}} K_i(y-1,1) & 0 \cdots & 0 \end{pmatrix}$$

This represents the familiar "loss development triangle". We will denote such an experience matrix by "$" and the triangle of claim counts by "#"[13].

Once the experience matrices $ and # have been calculated for one iteration, they are used as input for several different rating techniques (estimators of E(L)). These will be described in the "RATING METHODS" section.

13. The results to date are based on rating methods that use $ and/ or # as their input statistics. Of more interest are techniques that use triangles of some function of each known loss (such as losses truncated at basic limits).

CREDIBILITY

Often in experience rating we wish to use some outside experience that we believe is "related" to the insured in question. For example, we may use an insured's own basic limits experience, but rely on outside information for loss development factors, trend factors and expected excess losses.

The model underlying the use of this outside data is that the particular insured being rated was randomly selected from the group of all potential insureds of the same type. Therefore, the θ that we are trying to estimate is a realization of a random variable. θ's probability distribution is referred to as a structure function $U(\theta)$ [14]. If we have statistics available for many other insureds we can estimate certain properties of the group of all potential insureds (referred to as the collective). This then gives us valid information to use in estimating E(L) for a particular insured. Credibility theory addresses the question of how to combine data from the collective with data from the individual insured to arrive at the estimator of E(L) with the best properties[15]. The MSE's for certain credibility systems have been explicitly calculated[16] however, trend has seldom been[17] and loss development has not been addressed.

14. H. Bühlmann, Mathematical Methods in Risk Theory, Springer-Verlag, 1970
15. More precisely credibility theory restricts itself to linear combinations of collective data and individual risk data.
16. Fl. DeVylder, Introduction to the Actuarial Theory of Credibility.
17. C. Hachemeister, "Credibility for Regression Models with Application to Trend" in Credibility Theory and Applications ed. P. Kahn, Academic Press, 1975.

Several rating techniques that use outside information, in particular trend factors, were tested in the simulation. These tests are not strictly valid within the framework of credibility theory because the trend factors have not been estimated from a collective -- they have simply been postulated[18]. However, in all cases several different trend factors have been tested, including ones known to be wrong, in order to test the sensitivity of the rating method to incorrect assumptions about trend.

RATING METHODS

The following methods that used only the information contained in $ were tested:[19]

Method #1: Loss dollars are projected to ultimate by age-to-age factors. A least squares line[20] (restricted to a slope ≥ 0) is fitted to the five ultimate results to project the sixth year.

Method #2a, b, c, d, e:

Loss dollars are projected to ultimate using age-to-age factors. The ultimate result for each accident year is trended to the current year by multiplying it by an inflation factor raised to the appropriate power. The 5 trended results are then averaged to predict the current year. Cases, a, b, c, d and e refer to trend factors of 0%, 5%, 8%, 12%, and 15%, respectively.

18. To the extent that trend factors serve to project the effects of future inflation rather than adjust experience for the effects of inflation during the experience period, one would probably not want to estimate inflation from the data anyway, but rather use an exogenuous factor based on economic considerations.
19. A numerical example of each rating technique is contained in Appendix B.
20. Unrestricted linear and exponential fits were tested and gave similar results.

Method #3a, b, c, d, e:

"Adjustment to Total Known Losses method"

Estimated expected ultimate losses $= \left(\sum_{j=1}^{5} K_j \cdot (1+i)^{6-j}\right) \div \left(\sum_{j=1}^{5} \frac{1}{f_j}\right)$

Where K_j is known losses at current year for accident year j

f_j is the age-to-ultimate factor for accident year j

i is a trend factor which was set at 0%, 5%, 8%, 12% and 15% for 3a thru 3e, respectively.

The derivation of this formula is given in Appendix A.

The following rating methods using both $ and # were tested:

Method #4a, b, c, d, e:

Claim counts are projected to ultimate using age-to-age factors. The estimate of the 6th year is the average of these five results. This is multiplied by actual average known claim size, trended to the current year as in Method 2.

Method #5a, b, c, d, e:

Same as Method 4 except ultimate claim counts are projected by the Adjustment to Total Known Losses method.

Method #6: Ultimate claim counts by year are projected by the Adjustment to Total Known Losses method. For each accident year these are multiplied by actual average claim size. The results are trended by linear least squares (restricted to a slope ≥ 0) to project the sixth year.

RESULTS

The computer simulation model was written in APL and run on an IBM 5110 mini-computer. Creating six year's experience (five years to rate from, and one year as the policy's experience) for an average of forty losses per year, then applying twenty different rating techniques to the known losses took about 5 minutes, so 500 iterations took about 42 hours to run.

The simulation was run under four different sets of parameters. The first set were the ones given in the previous section. The second set were the same except that the severity trend was 8% the first four years (during the experience period) and 12% thereafter. The third set was the same as the first except the expected value and standard deviation of number of claims (N) increased by 5% each accident year starting with an expect number of 25 the first year. This could be used to reflect either an increase in exposure units not reflected in the rating method, or an unsuspected frequency trend. For the fourth run, the distributions were set as uniform to test the robustness of the previous results to wild departures in the form of the distributions. Exhibit 1 gives a summary of the parameters in each of the above cases. It also shows true[21] E(L) for each case -- this is the value we wish the rating techniques to be close to most of the time.

Exhibit 2 shows the simulation results of the distribution of R (the experience rated estimate of E(L)) for the first set of parameters.

21. Actually this value is also an estimated one, see note 7.

EXHIBIT 1

The four sets of parameters against which the rating methods were tested.

	#1	#2	#3	#4
E(L) for year 6	\$731,000	\$837,000	\$575,000	\$1,664,000
Standard Error of Estimate of E(L)	\$8,000 based on 2025 iterations	\$13,000 based on 1000 iterations	\$15,000 based on 500 iterations	\$29,000 based on 1000 iterations
Number of Losses N	Normal $\mu = 40$ $\sigma^2 = 60$		Normal $\mu = 25 \times (1.05)^j$ $\sigma^2 = 40 \times (1.05)^{2j}$ $j = 0, \ldots, 5$	Uniform max = 30 min = 1
Date of Loss within year M_i	Uniform max = 1 min = 0			
Report Lag Q_i	Exponential mean = 1.5			Uniform max = 4 min = 0
Payment Lag P_i	Exponential mean = 1			Uniform max = 4 min = 0
Payment Amount C_i	Lognormal $\mu = 8 + \ln I\,(a + M_i + Q_i + P_i - 1)$, $\sigma^2 = 2.5$ (mean = 10,405 x I, standard deviation = 34,793 x I)			Uniform max = 100,000 x $I\,(a + M_i + Q_i + P_i - 1)$ min = 1
	$I(t) = 1.08^{t-1}$	$I(t) = \begin{cases} 1.08^{t-1} & t \leq 5 \\ 1.08^4 \times 1.12^{t-5} & t > 5 \end{cases}$	$I(t) = 1.08^{t-1}$	$I(t) = \prod_{j=1}^{t} (1+r_j)$ where r_j generated randomly uniform (.2, 0)
Reserve Error V_i	Lognormal = -.549 = 1.099 (mean = 1, variance = 2)			Uniform max = 2 min = 0

EXHIBIT 2

ACTUAL INFLATION 8% PER YEAR
TRUE E(L) = $731,000

Rating Method	Information Used	Chosen Trend Factor	Distribution of R: Bias	Distribution of R: Standard Deviation
1	$	fit	$240,000	$870,000
2c′	$	8%	50,000	370,000
3c′	$	8%	- 30,000	250,000
4c′	$, #	8%	- 70,000	220,000
5c′	$, #	8%	- 90,000	180,000
6	$, #	fit	- 50,000	540,000

The bias (E(R-L))column shows whether each method will produce too much or not enough premium on the average. The standard deviation of R measures of how wide a range of results the various methods will give. Because there is only one right answer ($731,000), the smallest possible spread of estimates is the most desirable.

The six rating methods can be sorted into two groups depending on how they handle trend. Methods 1 and 6 fit a least squares line thru the estimated ultimate results for the past five years to project the sixth year. They, therefore, try to estimate the underlying trend based solely on the insured's experience. Both of these perform poorly in terms of standard deviation, and method 1 is highly biased.

Methods 2 thru 5 use a postulated trend factor that adjusts each accident year to current level. Of course, the bias for the version using an 8% trend factor (Methods 2c thru 5c) should be low because that is the true inflation rate underlying the model. The bias need not be zero because the rating technique may not take inflation into account exactly the way the loss generating model does (the rating techniques all trend past accident years to the current year whereas in the loss model inflation acts on all open claims across calendar years).

A way of reflecting trend in Methods 2 thru 5 that appears to be superior[22] to trending each accident year separately and averaging the results (as is done in methods 2 thru 5,b thru e) is to adjust

22. The conditions under which each of the two methods are superior are discussed in Appendix C. The simulation did not provide conclusive results either way.

the untrended result for three years trend; in other words, trend the average result rather than average the trended results. The methods labled 2c′ thru 5c′ are ones for which the untrended results (2a thru 5a) were adjusted by (1.08^3). [23] The bias and standard deviations shown in Exhibit 2 for methods 2c′ thru 5c′ were not arrived at by *simulation, but rather a straight adjustment of the simulated* results for methods 2a thru 5a (the untrended versions).

Methods that use both \$ and # (methods 4 thru 6) have a smaller variance than those that use \$ alone (methods 1 thru 3). [24] However, all the ones using \$ and # tested here suffer from a serious defect. That is that they have no way of detecting reserve deficiencies from the data. In this model, (the expected value of) reserves are deficient to the extent of future inflation, so this leads to a downward bias in the techniques. Methods that analyze loss development from \$ can attempt to detect such under reserving (at least to the extent that the earliest experience year is truly fully mature).

One obvious conclusion is that the more things we try to estimate from the data (e.g., trend, reserve deficiency) the higher the variance of the estimator will be. This suggests that for a given set of data we should be realistic about what effects we can estimate from it. This is, of course, the "full credibility" question: "How much data do you need to give your estimator satisfactory variance?" In the case of the risk sizes used in this simulation it seems that one should not

23. Actually the unbiased adjustment is $5/\sum_{i=1}^{5} \frac{1}{1.08^i}$, which is very close to 1.08^3.
24. This is plausable result, which should be true in all but very unusual cases. However, it should be noted that the loss model further tilts results in this direction because it uses constant frequency parameters.

try to estimate trend (methods 1 and 6) but one can use a method that is sensitive to reserve deficiency (method 3).

Method 3c′ gives the best overall result with a variance slightly higher than methods 4c′ and 5c′, but the smallest (absolute value of) bias of any method. It is interesting that the Adjustment to Total Known Losses method (methods 3 and 5), which takes the total known losses for all years and divides that sum by an overall adjustment factor for loss development, has a smaller variance than simply projecting accident years to ultimate and averaging the results (methods 2 and 4). This is analoguous to the earlier comment about more efficient trend adjustment. Appendix C shows that under some conditions this is a Best Linear Unbiased Method.

The calculation of the bias and standard deviation for any of the methods 2 thru 5 where a trend factor different from 8% was (incorrectly) selected is straight forward:

$$\text{bias for trend } r = (E(L) + \text{bias for } 8\%)\left(\frac{1 + r}{1.08}\right)^3$$

$$\text{std. dev. for trend } r = (\text{std dev for } 8\%)\left(\frac{1 + r}{1.08}\right)^3$$

A 50% error in selecting r (i.e., 12% or 4% instead of 8%) will introduce a bias of about + 12% to an otherwise unbiased technique.

Exhibit 3 shows how well each method performed under parameter sets 2, 3, and 4. Remember set 2 has an accelerating severity trend, set 3 has a frequency trend and set 4 uses all uniform distributions.

EXHIBIT 3

Rating Method	Parameters #2 True E(L) = $837,000 Bias R	Std. Dev. R	Parameters #3 True E(L) = $575,000 Bias	Std. Dev. R	Parameters #4 True E(L) = $1,664,000 Bias	Std. Dev. R
1	$170,000	$1,140,000	$100,000	$780,000	$800,000	$5,120,000
2c′	- 30,000	460,000	- 60,000	290,000	130,000	1,870,000
3c′	- 120,000	260,000	-120,000	210,000	-280,000	380,000
4c′	- 170,000	250,000	-140,000	170,000	-330,000	470,000
5c′	- 190,000	230,000	-130,000	160,000	-370,000	330,000
6	- 140,000	540,000	-100,000	360,000	-460,000	570,000

Method 2c′ does the best in terms of bias, however, has a high standard deviation. This exhibit shows that the ranking of methods from low to high standard deviation and from low to high bias seems to be fairly insensitive to changes in parameters. However, performance in terms of absolute value of bias depends on how the trend underlying the model compares with the trend chosen in the rating method.

VALIDITY OF THE RESULTS

Two issues should be considered when assessing the validity of the results.

1. Are 500 iterations a sufficient number to give stable estimates of the mean and variance of the distribution of R? The standard errors of the bias can be estimated as $\left(\frac{\mathrm{Var}(R-L)}{500}\right)^{\frac{1}{2}}$

   A 95% confidence interval around the estimates of the bias shown in exhibits 2 and 3 should be roughly two standard errors on either side of the estimated value.

   Taking Var(R-L) to equal $\widehat{\mathrm{Var}}(R) + \widehat{\mathrm{Var}}(L)$ where these are the variances estimated by the simulation, give standard errors of the bias estimates ranging from \$15,000 to \$30,000 (the rating methods with larger Var(R) having larger standard errors). This means that a rating method that is actually unbiased could show a bias of roughly + \$50,000 based on 500 simulations.

   The stability of the estimates of Var(R) are not known.

Note that because several (but not all) rating methods were tested during the same computer run (the same set of 500 simulated experience periods) there is a positive covariance between the estimates of E(R-L) (and also Var(R)) for rating methods 1 thru 4,and 5 and 6, but the estimates between these two groups of methods are independent.

2. Are the results specific to the form of the loss generating model that was used; how different would the ranking of efficiency of the rating methods been under a somewhat different model?[25] Many possibilities suggest themselves: inflation may affect different sizes of losses differently, reserves may be set in a different fashion with strengthenings occurring during a calendar year across all accident years, frequency and severity may not be independent. At least the model has shown that an extreme change in parameters (set 4) does not affect the conclusions greatly.

At the time of the writing of this paper, the computer model was not sufficiently sophisticated to test rating techniques of real interest, such as ones that adjust losses for changes in exposure during the experience period, ones that truncate losses at various levels, credibility weighing techniques and excess of loss experience rating techniques. Hopefully simulation results on some of these types of techniques will be available for presentation at the Spring meeting.

25. One error in the current model is that the severity distribution should allow for claims closed without payment.

RISK

Viewing premium as a random variable raises some new issues in the calculation of profit loading.

The random variable of ultimate interest to an insurer is its profit[26] on a given insured or group of insureds.

Let U be the random variable underwriting profit on the individual insured.

Let π be a fixed profit loading[27]

Let R be the experience rated estimator of E(L)

So U = Experience rated premium - L = $(\pi + R) - L$

The variance of profit on a single insured is

$$\begin{aligned} Var(U) &= Var(\pi + R - L) \\ &= Var(R) + Var(L) - 2\,Cov(R,L) \end{aligned}$$

R is based on known losses for prior years whereas L is losses for the period to be rated. We have been assuming that loss occurrences are independent, so Cov(R,L) = 0.

26. For simplicity's sake we are ignoring investment income considerations here.
27. Of course, this term should depend on the "riskiness" of insured.

If U were not random, the insurer would face no risk or variability of results. The insurer's risk[28] arises from the variability of U, which in turn arises from the variability of both R and L.

The insurer is frequently in a situation of being one of several companies quoting prices from which the insured will pick the lowest. This means that E(U) no longer equals $E(R) - E(L) + \pi$ (or bias plus loading) but rather

$$E(R \mid R+\pi < k) - E(L) + \pi$$

where k is the minimum of the other quoted prices for the insured.

Consider an unbiased rating technique R . Assume that the "proper" expected profit margin (based on risk considerations) has been determined to be π'. That is, we wish

$$E(U) = \pi'$$

$$E(U) = E(R \mid R + \pi < k) - E(L) + \pi$$

$$= E(R) - E(L) - (E(R) - E(R \mid R + \pi < k)) + \pi$$

So $\pi = \pi' + (E(R) - E(R \mid R + \pi < k))$

This says that the profit margin added to an unbiased estimate of expected losses should contain two pieces, 1. a risk loading (π') and 2. a factor to load for the antiselection you expect to suffer

28. The proper measure of "risk" for an insurer (or in fact for any financial transaction) is a much debated topic. Two the leading candidates are Var(U), which seems to be favored by actuaries, and Cov(U,M) where M is the return the entire market of assets, which arises from the CAPM. The CAPM unfortunately implies that insurance underwriting is almost riskless, because $Cov(U,M) = Cov(R,M) + Cov(L,M)$ (with our independence assumptions) and both of the terms on the right should be near zero.

in a competative bidding situation (if your quote is accepted, it is more likely that you underestimated expected losses).[29] Notice that an estimator R with a smaller variance will be desirable because it will decrease both components of the loading.

29. Two implications of this:
    1. in a renewal situation with no outside quotes, an insurer should be able to quote a lower price than otherwise because he will not need this loading
    2. the more companies quoting, the higher this loading should be

APPENDIX A

Derivation of "Adjustment to Total Known Losses method"

For Accident Year j

k_j = Known losses (thru current year)

u_j = Actual ultimate losses

$IBNR_j$ = IBNR

f_j = Age-to-ultimate factor

i = Trend factor

u = True expected losses for year 6

Assume 1) $u = \frac{1}{5} \sum_{j=1}^{5} u_j (1+i)^{6-j}$

2) $u_j = k_j + IBNR_j$

So $5 \cdot u = \sum_{j=1}^{5} K_j (1+i)^{6-j} + \sum_{j=1}^{5} IBNR_j (1+i)^{6-j}$ (*)

Assume 3) $IBNR_j = (1-\frac{1}{f_j}) \frac{u}{(1+i)^{6-j}}$

So $\sum_{j=1}^{5} IBNR_j (1+i)^{6-j} = u(5-\sum_{j=1}^{5} \frac{1}{f_j})$

substituting this into (*) gives

$$u = \left(\sum_{j-1}^{5} k_j (1+i)^{6-j}\right) \div \left(\sum_{j=1}^{5} \frac{1}{f_j}\right)$$

Numerical Examples of Rating Methods

Let

$$\$ = \begin{pmatrix} \$242{,}744 & \$202{,}907 & \$216{,}946 & \$223{,}772 & \$243{,}633 \\ 135{,}700 & 536{,}598 & 608{,}794 & 636{,}252 & 0 \\ 70{,}734 & 535{,}107 & 733{,}341 & 0 & 0 \\ 42{,}031 & 222{,}841 & 0 & 0 & 0 \\ 185{,}689 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(1) Accident Year	(2) Age to Age Factor[30]	(3) Age to Ultimate Factor	(4) Known Losses	(5) Ultimate Losses	(6) Reciprocal of Age to Ultimate Factor
1	1	1	$243,633	$243,633	1
2	1.0888	1.0888	636,252	692,751	.9184
3	1.0415	1.1340	733,341	831,609	.8818
4	1.2232	1.3871	222,841	309,103	.7209
5	3.0485	4.2285	185,689	785,186	.2365
Total			$2,021,756	$2,862,282	3.7576

Method #1 Column 5 projected to accident year 6 by linear least squares = $782,294

Method #2c′ ((Sum of column 5) ÷ 5) x $5/\sum_{i=1}^{5} \frac{1}{1.08^i}$ = $716,877

Method #2c ($243,633 x 1.08^5

+ 692,751 x 1.08^4

⋮

+ 785,186 x 1.08) ÷ 5 = $711,317

Method #3a ((Sum of column 4) ÷ (Sum of column 6)) = $538,044

30. e.g., $1.0415 = \dfrac{636{,}252 + 223{,}712}{608{,}794 + 216{,}946}$

Let

$$\# = \begin{pmatrix} 10 & 21 & 28 & 28 & 29 \\ 23 & 45 & 49 & 52 & 0 \\ 14 & 44 & 54 & 0 & 0 \\ 11 & 29 & 0 & 0 & 0 \\ 11 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(7) Accident Year	(8) Age-to-Age Factor	(9) Age to Ultimate Factor	(10) Known Counts	(11) Ultimate Counts	(12) Reciprocal of Age to Ultimate Factor
1	1	1	29	29	1
2	1.0357	1.0357	52	53.86	.9655
3	1.0390	1.0761	54	58.11	.9293
4	1.1909	1.2815	29	37.16	.7803
5	2.3966	3.0712	11	33.78	.3256
Total			175	211.91	4.0007

Method #4a ((Sum of Column 11) ÷ 5) x (($243,633 ÷ 29
+ 636,252 ÷ 52
+ 733,341 ÷ 54
+ 222,841 ÷ 29
+ 185,689 ÷ 11) ÷ 5) = $498,263

Method #5a ((Sum of Column 4) ÷ (Sum of Column 12)) = $505,351

(13) Accident Year	(14) $(1 - (12)) \times \frac{211.91}{5}$	(15) (10) + (14)	(16) (15) x (5) ÷ (10)
1	0	29	$243,633
2	1.46	53.46	654,116
3	3.00	57.00	774,082
4	9.31	38.31	294,381
5	28.58	39.58	668,143

Method #6 Column 16 projected to accident year 6 by linear least squares = $673,657.

APPENDIX C

Comparison of "adjusting, then averaging" vs "averaging, then adjusting"

Let X_i be a random variable representing observed losses for accident year i

Assume that these losses arise from distributions with expected values that are constant over time, except for an adjustment factor. This adjustment factor can represent either a loss development factor or a trend factor or both.

So $X_i = \frac{\mu}{a_i} + \epsilon_i \quad i=1,\ldots,n$

where μ = underlying expected losses

a_i = non-random adjustment factor (≥ 1)

ϵ_i = random error $E(\epsilon_i) = 0$, $Var(\epsilon_i) = \sigma_i^2$

We wish to estimate μ

Let $\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i a_i$

This represents trending (and/or developing) known losses for each year and averaging the results

Let $\hat{\mu}_2 = \left(\sum_{i=1}^{n} X_i\right) \div \left(\sum_{i=1}^{n} \frac{1}{a_i}\right)$

This represents the "adjustment to total known losses method"

It is easy to see that both $\hat{\mu}_1$ and $\hat{\mu}_2$ are unbiased, ie

$E(\hat{\mu}_1) = E(\hat{\mu}_2) = \mu$

Calculate the Best Linear Unbiased Estimate (B.L.U.E.)[31] of μ

That is, find weights c_i such that $\hat{\mu}\left(=\sum_{i=1}^{n} c_i X_i\right)$ is unbiased and has minimum variance.

So minimize $\mathrm{Var}\left(\sum_{i=1}^{n} c_i X_i\right)$ subject to $E\left[\sum_{i=1}^{n} c_i X_i\right] = \mu$

$$\mathrm{Var}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 \sigma_i^2$$

$$E\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} \frac{c_i}{q_i}\mu = \mu \Rightarrow \sum_{i=1}^{n} \frac{c_i}{q_i} = 1$$

Let

$$L = \sum_{i=1}^{n} c_i^2 \sigma_i^2 + \lambda\left(1 - \sum_{i=1}^{n} \frac{c_i}{q_i}\right)$$

$$\frac{\partial L}{\partial c_i} = 2c_i\sigma_i^2 - \frac{\lambda}{q_i} = 0 \qquad i = 1, \cdots, n$$

So $c_i = \dfrac{\lambda}{2q_i\sigma_i^2}$, $\displaystyle\sum_{i=1}^{n} \frac{c_i}{q_i} = \sum_{i=1}^{n} \frac{\lambda}{2q_i^2\sigma_i^2}$

So $\lambda = \dfrac{2}{\displaystyle\sum_{i=1}^{n} \frac{1}{q_i^2\sigma_i^2}}$

So $c_i = \dfrac{1}{q_i\sigma_i^2} \dfrac{1}{\displaystyle\sum_{j=1}^{n} \frac{1}{q_j^2\sigma_j^2}}$

Now consider various possibilities for σ_i^2

1. Let $X_i q_i = \mu + e_i$ where $\mathrm{Var}(e_i) = \sigma^2 \ \forall i$

   This means that $\epsilon_i = \dfrac{e_i}{q_i}$ so $\sigma_i^2 = \dfrac{1}{q_i^2}\sigma^2$ so $c_i = q_i$

   Therefore $\hat{\mu}_1$, is the BLUE

2. Let $\dfrac{\mathrm{Var}(X_i)}{E[X_i]} = k \quad \forall i$

   So

$$\frac{\sigma_i^2}{\mu/q_i} = \frac{\sigma_i^2 q_i}{\mu} \Rightarrow \sigma_i^2 q_i = k\mu$$

31. The approach of calculating the B.L.U.E. was suggested by Aaron Tenenbein, Associate Professor, Statisics and Actuarial Science.

This means that $c_i = 1 \Big/ \sum_{i=1}^{n} \frac{1}{q_i}$

Therefore $\hat{\mu}_2$ is the BLUE

As was discussed in the results section, $\hat{\mu}_2$ performed better than $\hat{\mu}_1$, in the simulation.

ADDITIONAL REFERENCES

Bicker, Peter and Doksum, Kjell, Mathematical Statistics: Basic Ideas and Selected Topics, Holden-Day, 1977

Fishman, George, Concepts and Methods in Discrete Event Digital Simulation, John Wiley & Sons, 1973

EXPERIENCE RATES AS ESTIMATORS: A SIMULATION OF THEIR BIAS AND VARIANCE
By James N. Stanard

REVIEWED BY John P. Robertson

Mr. Stanard's paper opens a new area of actuarial reasearch, namely the use of simulation to investigate the reliability of commonly used pricing (and related) models. He is not using simulation to forecast insurance results directly, but rather to determine how well a given technique for such forecasting can be expected to perform. I believe this is not a paper to read to find final answers, but rather to find groundbreaking results from a technique which should become more widely used.

In this review, I will comment on the interpretation of the results from the standpoint of bias and variance, clarify (I hope) the algebra underlying the derivation of the "Adjustment to Total Known Losses Method" and conclude with some comments on other possible areas of application of this technique.

BIAS OF RESULTS

On an initial reading of the paper, I was surprised by the biases developed by the various experience rating procedures. These range from +30% to -10% of the expected losses under the first set of parameters and are statistically significant since they are quite a bit larger than the standard error of the estimate of the expected losses. The techniques used are similar to commonly used ones which are not normally assumed to be biased. It is possible, however, to see the causes for the bias in the procedures used.

Mr. Stanard notes the underreserving of claims by the amount of future inflation. This combined with the use of loss development factors that only go to the fifth year, produce a downward bias in Methods 1 through 3. Methods 4 through 6 are given downward biases both by the use of claim development factors that only go through year 5, and by use of the actual average known claim size. Given the reserving method used, the average actual known claim size for a given accident year will tend to increase as the accident year develops since new reported claims come in at a higher average amount and outstanding reserves are increased for inflation. The latter has two effects on Method 6; it serves to reduce the average claim size (as in Methods 4 and 5) and since it reduces the latest year the most and the earliest year the least, it biases the average trend downwards. Finally, the restriction of fitted slopes to be positive in Methods 1 and 6 contributes an upward bias to these methods.

The above effects are all of the significant sources of bias in the cases where the trend factor used is equal to the true inflation rate underlying the model (8% per year). (There is another minor source which I will discuss in the "Adjustments to Total Known Losses" section.) I do not agree with Mr. Stanard's comment that in this case "the bias need not be zero because the rating technique may not take inflation into account exactly the way the loss generating model does". If the ultimate losses for each accident year were projected without bias (which, of course, they're not), then any of the Methods 2 through 5 should give unbiased results. Methods 1 and 6 would also give unbiased results if exponential fits were used (the straight line gives a very slight downward bias) and if the slope of the fitted line were allowed to become negative. Clearly, if the trend factor used is not equal to the underlying rate of inflation, then a bias will be introduced into any of the above methods.

I remain surprised by the positive bias of $50,000 reported for Method 2c' under the first set of parameters. The only major source of bias in Method 2 is the downward one due to lack of full development. The possible likely error in the bias, as noted by Mr. Stanard in his section on "Validity of the Results" is about $50,000. The direct interpretation of the simulation is, therefore, that Method 2 is very unlikely to have a negative bias. This seems to me to be in conflict with the "a priori" expectation of Method 2's bias. The fact that the second set of parameters gives a negative bias does not help explain this conflict since under the second set of parameters all of the biases (including method 2's) have moved down about $90,000 from the biases under the first set of parameters.

The fact that biases exist in the methods under consideration is interesting, but I hope my discussion has shown that their existence is not too surprising, since reasons can be found for expecting bias. Of more interest is the relative magnitudes of the biases, since these are harder to predict in advance. Also Methods 1 and 6 have both positive and negative sources of significant bias, and predicting which will win out would not be easy by a priori methods.

VARIANCE OF RESULTS

I believe that the simulated standard deviations are of far greater interest than the biases. Bias in pricing methods is something that actuaries are used to dealing with and there are obvoius techniques for eliminating the bias in Methods 2 through 5. (Under

Methods 1 and 6, it would be difficult to part with the restriction of non-negative slope; hence, eliminating bias for these methods is not as easy.)

The standard deviations are not only harder to predict than bias, but they cannot be "repaired" in the sense that a suspected non-zero bias can be repaired. I do not disagree than Method 3c' gives the best overall result of the methods tested, but I would nominate Method 5c' as having the greatest promise since it shows the least variance. In most applications of Method 5, the lack of full development of claim count would be apparent and some adjustment could be made to approximate full development. Similarly, the development to ultimate of the average claim size in each accident year could probably be addressed. Hence, eleminating the bias in Method 5 could likely be achieved. As the underlying parameters are changed, I think Method 5's advantages become clear.

Moving from the first set of parameters to the second (that is, missing a change in the trend) influences the bias of all the methods similarly. If all were unbiased under the first set of parameters, all would be biased by about $-90,000 under the second set. While Method 5 does not do any better than the other methods here, it doesn't do any worse either. The various methods do not react as uniformly to the introduction of an unsuspected claim count trend. The bias of Method 5 changes less from the first to the third set of parameters than any of the other methods. No matter what set of parameters is used, Method 5 shows the least variance.

Obviously, neither Mr. Stanard's paper nor my discussion will prove that some one experience rating technique is the ultimate such technique. I do hope that I have shown that Mr. Stanard's results already contain much information of value in choosing a technique and that the variance information he gives is likely to be as useful or more useful than the bias information.

ADJUSTMENT TO TOTAL KNOWN LOSSES METHOD (ATTKLM)

The ATTKLM is an interesting method for applying development factors to reported results. My purpose here is to show the algebra underlying the general application of this method, and to tie together all of the places it is actually used in Stanard's paper.

The Appendix to this discussion shows that if $A_iB_i=C$ for all i in some set, then $(\Sigma A_i) \div (\Sigma 1/B_i)=C$ and that is essentially the ATTKLM. The application shown in Standard's Appendix A (used in Methods 3, 5, and 6) is the first application in my appendix. Additionally, Methods 2c' through 5c' use the same theory but with the B_i as trend factors. Stanard justifies this latter use in his Appendix C, wherein he notes that he is also justifying the use as loss or claim development factors. While his Appendix A and my appendix show that the ATTKLM will not introduce any bias into an experience rating method, his Appendix C makes a strong argument for expecting less variance in results when this method is used.

The paper effectively compares the use of the ATTKLM to the more normal "adjust, then average" since this is the essential difference between Methods 2 and 3 and also between Methods 4 and 5. In either case, the ATTKLM shows less variance than the "adjust, then average" method. It would be interesting to compare Methods 2c through 5c to Methods 2c' through 5c' to further test this comparison. Stanard's footnote 22 notes "the simulation did not provide conclusive results either way." Since the average loss development factors are likely to be larger than the average trend factors, it is possible that the test using loss development factors is more likely to show a difference.

As a final comment on the use of the ATTKLM, I must point out that as used in methods 3c' and 5c', it introduces a slight bias. To illustrate with Method 3c', Stanard has taken

$$\left(\tfrac{1}{5}\Sigma k_j\right) \div \left[\left(5 / \Sigma (1/f_j)\right) \times \left(5 / \Sigma\left(1/(1.08)^{6-j}\right)\right)\right]$$

where it is more correct to take

$$\left(\tfrac{1}{5}\Sigma k_j\right) \div \left[5 / \Sigma (1/f_j)\left(1/(1.08)^{6-j}\right)\right]$$

One cannot keep spinning off factors B_i since after the first time the relationship $A_iB_i=C$ fails to hold. The effect of this difference is to introduce a very slight negative bias in Methods 3c' and 5c', which coincidentally is approximately offset by the use of $(1.08)^3$ in place of $5 \div \Sigma 1/(1.08)^{6-j}$.

CONCLUSION

In most applications of Stanard's technique, one is not going to be able to specify the distributions underlying the experience (if one could, then one could estimate mean losses far more accurately than any normal experience rating method allows). Thus, the most significant conclusions to be drawn from the simulations are not in the area of what model best fits some given data, but rather are in the area of choosing a technique in the absence of any information other than the reported results.

Areas worthy of further investigation include refinements to the underlying assumptions to make the model more realistic (including application of credibility weighting techniques), use for larger models (what variation should be anticipated in Homeowners or Auto Liability indications when the standard ratemaking techniques are used?), and the use for testing possible variance in rating or ratemaking methods due to particular components of the methods. Obviously the larger the model, either in terms of number of assumptions needed or the number of claims and other items which may need to be simulated, the greater the possibility of the cost of running the simulations of becoming prohibitive. Even though the simulations presented are based on relatively simple experience rating techniques, it is clear that a great deal of work was required to achieve the results.

In summary, Mr. Stanard has provided a very interesting and useful paper, both from the standpoint of the results of the simulations given and also because of the introduction of the "average, then adjust" method of applying trend and development factors.

APPENDIX

The following lemma generalizes the result of Stanard's Appendix A. Applications include the two used in his paper and two new ones.

Lemma: If $A_iB_i=C$ for i=1 to n

Then $(\Sigma_i A_i) \div \Sigma_i (1/B_i) = C$ (1)

Also $(\Sigma_i W_i A_i) \div \Sigma_i (W_i/B_i) = C$ for any W_i (2)

Proof: $A_i = C/B_i$

$\Sigma A_i = C \times \Sigma (1/B_i)$

$(\Sigma A_i) \div \Sigma (1/B_i) = C$

This establishes (1). (2) follows by substituting W_iA_i and B_i/W_i for A_i and B_i in (1).

Applications: For these applications, think of k_j, u_j, and u as expected values rather than as actual reported values.

1) $A_j=k_j\ (1+i)^{6-j}$, $B_j=f_j$, $C=u=A_j \cdot B_j = k_j \cdot f_j (1+i)^{6-j}$

Then $u = (\Sigma k_j (1+i)^{6-j}) \div (\Sigma 1/f_j)$

This is the application given in Stanard's Appendix A. The fact that $u=k_j f_j (1+i)^{6-j}$ follows from his Assumption 3) as follows:

$IBNR_j = u_j - k_j = u_j - u_j/f_j = (1-1/f_j)(u_j)$ (Definitions)

$= (1-1/f_j) \dfrac{u}{(1+i)^{6-j}}$ (By his assumption x(3))

Thus, $u_j = \dfrac{u}{(1+i)^{6-j}} \therefore u = u_j (1+1)^{6-j} = k_j f_j (1+i)^{6-j}$

2) $A_j=u_j, B_j=(1+i)^{6-j}, C=u=A_jB_j$

Then $u = (\Sigma u_j) \div \Sigma 1/(1+i)^{6-j}$

$= (\Sigma u_j) \times (5 \div \Sigma 1/(1+i)^{6-j})$

This is the method used to adjust untrended results per footnote 23 of Stanard's paper.

3) $A_j=k_j$, $B_j=f_j(1+i)^{6-j}$, $C=u=A_jB_j$

Then $u=(\Sigma k_j) \div \Sigma(1/f_j(1+i)^{6-j})$

These first three examples show that f_j and $(1+i)^{6-j}$ play symmetric roles in the projection of results.

4) "Ratemaking"

Let A_i be the reported loss ratio (developed to ultimate or not) and let B_i be the ratio of loss trend to premium trend (if any) from the i^{th} year to the period the rates apply to (times an ultimate development factor if not included in A_i). Then $(\Sigma W_iA_i) \div (\Sigma W_i/B_i)$, where the W_i are weightings of the various years such as 10-15-20-25-30, gives an estimate of the ultimate loss ratio analogous to the commonly used $\Sigma W_iA_iB_i$. Mr. Stanard's results indicate that the former may show less variation about the true mean loss ratio than the latter.

TITLE: ACTUARIAL ISSUES TO BE ADDRESSED IN PRICING INSURANCE COVERAGES

AUTHOR: Mr. E. James Stergiou

Mr. Stergiou is Vice President of Woodward and Fondiller, Inc., consulting actuaries. He is a graduate of City College of New York with a B.S. degree in Mathematics and received his FCAS in 1979. In addition, Jim is a member of the American Academy of Actuaries.

REVIEWER: Mr. Sheldon Rosenberg

Mr. Rosenberg is an Associate Actuary and Manager with Insurance Services Office. He received his FCAS in 1976 and has served on the CAS Examination Committee.

CO-REVIEWER: Mr. Aaron Halpert

Mr. Halpert is an Actuarial Supervisor for Insurance Services Office. He has an M.S. degree in Mathematics and is a student of the CAS.

A Statement of The Problem

A recent article in the Journal of Commerce cited an address given at the convention of The National Association of Casualty & Surety Agents by its President, Mr. John S. Childress, Vice President of Marsh & McLennan. His remarks emphasized the value of educating the public regarding casualty ratemaking procedures and company needs.

Speaking before NACSA, Mr. Childress noted that bills are being introduced at both federal and state levels "by legislators of Senator Metzenbaum's persuasion" that would substantially restructure the insurance business. He went on to say that, "Not only is the way insurance operates being questioned, but its credibility is on the line as well. We need to better explain ourselves, as there are those who question the integrity of our business simply because they do not understand it. That is clearly our fault. We must educate the public, the legislatures, and other government officials, who are all demanding accurate, credible and understandable answers to their questions. They have a need to know why we use certain rating classifications and not others." He indicated that we must make a genuine effort to understand the need of our insureds and help them comprehend the breadth as well as the limitations of their insurance policies. At the same time, Mr. Childress states "we must thoroughly examine our industry and find a reasonable balance between the profit motive of private enterprise and our responsibility to society at large. Society has changed faster than we expected, and we will have to adjust ourselves to the needs of the society in which we live. However, before we do that, the public must also be made to understand the price they will have to pay for these changes. Once consumers have a clearer picture of what is involved, they may be able to judge better what changes are either necessary or desirable, especially when they desire the same degree of protection and service or better, than what they already receive."

We believe Mr. Childress accurately gauges what one of the main challenges to the insurance industry will be in the 1980's. It is our responsibility as actuaries to explain to outside sources (state agencies, consumer groups, etc.) the factors that are (<u>or should be</u>) considered in rate reviews and filings and quantify them as much as possible. This paper begins to respond to this challenge by first describing the variables which should be expressly considered in rate filings, discussing the reasons for their inclusion and quantification and finally suggesting guidelines for the structuring of rate filings in a manner that will assist all sides (state departments, consumer groups and the company) to properly evaluate an entity's rate level requirements.

Introduction

During the course of our actuarial careers, many of us face the prospect of having our rate filings challenged by a State Insurance Department, a consumer bureau or our company management. The depth and level of the questions posed depend upon the quality and quantity of the reviewer's actuarial knowledge and/or personnel. This is true even when formal rate filings are not produced and rate changes are obtained via informal rate reviews not submitted to outside sources.

Some of these questions deal with the standard issues that have traditionally been considered when analyzing a filing. These are:

(1) The derivation of Loss Development Factors.

(2) The derivation of Trend Factors.

(3) Expense Provisions, including those for Loss Adjustment Expenses and the reflection of Investment Income.

(4) Reconciliation of filing results with those of the Annual Statement.

Until recently, the treatment of these variables was rather perfunctory in nature, and is summarized in Part 1 of this paper. However, there are several hidden variables which impact on insurer results and therefore affect each of these four items. Unfortunately, these hidden variables have, in many instances, not been analyzed carefully during the preparation of the rate revision and this has caused state agencies and consumer advocates particularly to contest rate filings and criticize ratemaking procedures. In the absence of this outside criticism, omission of the consideration of these variables can often lead to an inadequate, excessive or unfairly discriminatory rate structure.

The problem cannot be attributed to a lack of understanding since most, if not all, ratemakers are familiar with these missing factors. Instead, the blame lies with their lack of quantification and reflection in the rate structure through the filing's statistical support.

The first part of this paper consists of a discussion of the present methodologies underlying a rate filing. In the second part the author illustrates, by means of hypothetical examples, biases resulting from the application of these current treatments which affect the calculation of the factors enumerated in items (1)-(3) above. Part three suggests various tests that should be administered during the preparation of a rate filing or review in order to anticipate any questions that may arise.

included in that section is a proposed list of interrogatories to be answered concerning the filer's reserving, claim settlement rate, changes in its business mix and other relevant factors.

Part four consists of a sample newsletter written in laymen's language containing a brief description of ratemaking procedures along with an explanation for rate adjustments. Distribution to policyholders of this summary might help educate the general public about insurance projections and remove some of the potential causes of consumer dissatisfaction and distrust of the industry.

Part I: The Traditional Approach

A. Loss Development Factors

The loss development model usually followed in the past and, still employed to a great extent presently, calls for an age-to-age valuation of incurred losses over an experience period. This typical loss development exhibit is similar to that shown in Part A of Exhibit I. The selected "link" ratios, displayed on the bottom of Part B of that table, are frequently the result of computing the mean of those determined arithmetically above for the particular maturity studied. From these results, completion ratios are calculated. Application of these factors to the losses reported to date would yield a projection of ultimate incurred losses for each accident or policy year, as displayed in Part C of Exhibit I.

Conspicuously absent in this treatment are explicit measures of the variables which impact greatly on the loss growth curves defined by the link ratios. These variables reflect the following changes during the experience period (or expected to occur subsequently) used in the filing or subsequent thereto:

(i) filer's reserving policy and claim settlement procedures

(ii) policyholder profile such as the book of business such as by classification, territory, liability limit

(iii) propensity of late reported claims and all variables attendant thereto based on (i) and (ii) above

(iv) cause of loss for packaged policies

It would be desirable, therefore, to include within the filing or review analysis, statistical support quantifying each of the above variables in order to properly gauge the company's need for a rate revision.

All of these are combined and hidden by the simplistic traditional loss development method described above, resulting in possible biases in the results. In view of these possible distortions, therefore, it is not surprising that, in this age of consumerism, the past has finally caught up. Insureds and their political representatives have become more adept at discerning these inherent biases and have of late been requesting measurements of the effects of these variables on loss growth patterns and therefore on loss development factors.

B. Trend Factors

Traditionally, calendar year trend data of the type displayed in Exhibit II have been used to derive factors reflecting insurance inflation. This measure is comprised of two aspects: frequency and severity of claims and increases in fixed and variable expense costs.

As is seen, Exhibit II utilizes calendar year paid (i.e., closed) claim data usually fitted to an exponential curve to project measures of average frequency and severity.

However, any such treatment and calculation of trend is faulty because it ignores variables which impact on a filer's implicit trend data, viz:

(i) No determination is made as to whether the average maturity level of the closed claims in Exhibit II has changed over time. Hence, it is possible that at any point in time, there might be a large number of either older or younger claims being settled which might in turn yield larger or smaller than usual claim sizes. This, naturally, would throw off the results significantly. Claims closed during an accident year's first maturity may exhibit different claim cost trends than those closed later and there might have been a shift in the average age of closure during the experience period. It would therefore seem desirable to include a trend exhibit showing changes in frequency and severity by maturity within accident year, for example.

(ii) No correlation is made between changes in claim cost and frequency and changes in the company's policyholder profile during the experience period. A shift in the mix of insureds may affect loss growth patterns in the areas of propensity to sue and rapidity of reporting. The influence of these changes by classification and territory is obvious. It is, for example, known that urban insured's claim costs are higher than those of rural policyholders and that younger drivers tend to sue for less than those of middle age (i.e., or higher wage earning) group.

(iii) With respect to packaged policies, most filings do not include a cause of loss data breakdown. Such an analysis would provide trends of exactly where the loss dollars are coming from. Then, separate trend factors could be computed, using the filer's actual state experience or industrywide data in that particular state for that specific cause of loss (fire, liability, theft, etc.). Such cause of loss data would also be helpful in predicting trends in loss development patterns. Just as different loss development patterns are applicable for each cause of loss, so are different trend factors, both with respect to frequency and severity.

Once again, the problem with the traditional statistical support, as shown in Exhibit II, centers around the fact that the impact of these important variables, (i) - (iii) above, is not quantified. As in the case of loss development factors, severe biases can result when there has been a change in a particular aspect contributing to the final result.

C. Expense Provisions

The method traditionally used to quantify overhead expenses is illustrated in Exhibit III. Frequently, a provision for each expense item is developed using ratios of costs to written and/or earned premiums for the most recent several years.

However, there are two significant shortcomings in this method:

(i) Expenses are taken as a percentage of past calendar year collected earned premiums. The effect of prior year rate changes on these ratios and a list of budgetary estimates for the coming year are not considered.

(ii) The traditional loss ratio method of ratemaking, by which expense provisions are reflected in the rate structure in deriving a gross rate or level change, assumed that <u>all</u> expenses vary <u>directly</u> with losses. In fact, only some of the total insurer's expenses vary with premiums while others are relatively fixed. Separate treatments of these different types of overhead costs are required in order to arrive at a fair rate level.

Several states have begun recognizing the need and propriety of separating fixed from variable expenses and territorial flattening of fixed costs in their ratemaking models. This has proven beneficial to insureds who, in the past, might have faced as large an increase in the expense portion of the rate, which often totaled 35% of the total rate, as they did in the loss portion. This was particularly true in the professional liability area during the middle 1970's.

D. Reconciliation of Filer's Results With Those Shown In The Annual Statement

Increased consumer consciousness has given rise to policyholder indignation regarding companies' financial results. Insureds, cognizant of satisfactory earnings records enjoyed by insurance companies, fail to comprehend the necessity for large rate increases. It is vital that we in the insurance industry be equipped to explain this seeming contradiction.

In actuality, Annual Statement results are not and should not be used as statistical support for rate filings. Statement losses reflect countrywide data with all coverages, voluntary and assigned risk, Bodily Injury and Property Damage, net as to reinsurance, and so on, combined. If premiums and losses shown in the Statement were required to be as detailed (or anywhere near so) as those required in a rate filing or review, an eighteen wheel Mack truck would be needed to deliver a company's Statement to the insurance department.

Be that as it may, there is no doubt that some important parts of rate filings do not appear to be treated anywhere near as rigorously as they are on the filer's Statement. The main reference herein concerns reserve levels. A company's Annual Statement contains reserves which, in most cases, have been calculated with the utmost care to reflect the items of Subsection A above. Input from actuarial and other departments is used to properly evaluate a company's loss and loss adjustment reserve requirements, particularly in the area of case reserving adequacy, rate of settlement, trends and changes in the book of business. On the other hand, the derivation of loss development factors shown in rate filings (Exhibit 1) is comparatively cursory in nature. If one performed such an analysis using the data in Part 2 of Schedule P of a company's Statement, the resulting ultimate losses and/or statement reserves would, in most cases, be far different than those predicted in the company's balance sheet. As we stated earlier, this often results from not analyzing separately the above variables.

The above Subsections A - D of Part One have summarized the major issues the writer feels should be addressed when preparing a rate analysis or filing.

Part 2 will examine, for each of the three major adjustments (loss development, trend and expenses) discussed above, some of the biases which can occur without a complete and thorough review of the variables which impact on these factors.

Part 2: Discussion of Hidden Biases in The Traditional Approach

A. Loss Development Factors

(i) Changes in a filer's reserving policy

Consider the situation where a company has altered its reserving policy and is now reserving more adequately at the case level than it was in the past. At this particular point in time, therefore, the company's incurred losses are at a more mature level than at comparable dates in the prior accident or policy years.

However, application of historically derived loss development factors, particularly using the method described in Exhibit I, to the present, more adequate valuation of incurred losses would bias the filer's estmate of ultimate losses dramatically upward.

We describe below an example of how changes in the reserving policy of the filer can affect results. Part A of Exhibit IV displays the outstanding loss portion of the Exhibit I incurred development. Using the claim counts in Part B of Exhibit IV, average outstanding loss costs are computed in Part C. As evidenced upon examination of the last diagonal, a significant increase in this average outstanding loss cost appears to have occurred during the latest, 1978, calendar year. This, in the absence of large claims (for which these average reserves should always be examined for bias), may mean that the insurer has begun to reserve more adequately during 1978. Before any subsequent adjustment is made, this should be confirmed by questioning company personnel in the claim or underwriting area.

The Exhibit I completion ratio of 2.137 for example, which is applied to the 1978 incurred losses to date of $400,000 results from past reserve deficiencies. Applying it to the 1978, more adequate, first maturity incurred losses results in the possibly overstated ultimate loss projection of $854,800 produced in Column (3) of Exhibit I.

It is important to realize that, in any test of reserve adequacy similar to that shown in Exhibit IV, allowance must be made to reflect normal inflationary pressures that manifest themselves in rising claim costs. In Exhibit IV we notice that 1974-78 calendar year changes in the first maturity (i.e., 12 months) hovered in the 25% to 35% area. If we assume that external sources indicate that company claim costs are increasing 10% per year, we may conclude that the insurer has adopted a policy, express or implied, increasing its reserve adequacy.

This apparent strengthening at the case reserve level will greatly affect the future incurred loss growth pattern and should be recognized when selecting completion ratios. As we indicated earlier, any failure to reflect these changes will bias the projection of ultimate incurred losses dramatically upward. An adjustment should therefore be made to the link ratios produced in Exhibit I to reflect this change in the reserving policy of the carrier.

It is, therefore, imperative to include, in the rate filing or review, an exhibit measuring explicit or implied changes in the company's reserving policy.

One way to correct these biases is suggested in the description and numerical example shown in Exhibit V-A, wherein the most recent average outstanding loss cost for each past accident or policy year is used and prior year's average outstanding losses (for the same maturity) are adjusted backward by an estimated inflation factor. These "smoothed" average costs are then multiplied by the corresponding outstanding claim counts for the maturity/accident year cell to obtain total "adjusted" outstanding losses. When paid losses are added in the corresponding cells, the resulting artificial incurred development pattern can be used to calculate development factors and finally, ultimate incurred losses.

As can be seen from Exhibit V-B, this would result in incurred losses smaller by about $262,000 (i.e., $3,194,746 per Exhibit V-B compared with $3,456,639 per Exhibit I) or 8% as compared with the unadjusted results using the traditional approach in Exhibit I.

Of course, the above example may be an over simplification of the approach needed to be taken. However, the fact remains that, currently, little effort is expended by filers in the determination and quantification of changing reserving practices. This failure leaves the industry susceptible to criticism from state agencies and consumer groups.

(ii) Changes in the filer's rate of settling claims

A situation similar to that described above can result from a modification in the rate of claim settlement.

As before, application of historically derived paid and/or incurred development factors to present, more mature or adequate valuation of paid or incurred losses would likewise bias the results.

To illustrate, Part A of Exhibit VI has been prepared to test paid loss development data. Normally, in a fast closing line like Homeowners or Physical Damage, such paid loss input can be used in estimating a carrier's ultimate losses. Parts B and C of Exhibit VI display the projection of ultimate losses of $3,154,653 resulting from these empirical results using the traditional development approach. No test has yet been made to ascertain whether the company's claim settlement pattern has changed. Let us assume we perform such a study in Exhibits VII-A and VII-B.

Exhibit VII-A sets forth a simple age-to-age pattern displaying ratios of the number of paid claims during a period to the total number of incurred claims by maturity within accident year. Although use of report year data in the manner described in Part 3 of this paper is preferable, many insurers do not have such data readily available. However, the Exhibit VII-A calculation is usually available and can be used for our purpose. As can be seen in Exhibit VII-A, therefore, these "disposal" ratios indicate an increase in the rapidity rate of claim settlements.

If we reflect this speed-up in a modified paid development approach in the manner described beginning in Sheet 1 of Exhibit VII-A, we would apply a factor smaller than that employed in Part C of Exhibit VI for each year. The change in the ultimate loss projection between Exhibits VII-B and VI exceed $600,000 or about 19% due to this adjustment. Basically, the method involves an adjustment using the same ratio of paid to total losses going back in time by maturity within accident or policy year.

Changes in the reserving and claim settlement policy of the filer, either explicit or implied, constitute the two most prolific influences in a historical loss growth pattern. Estimation of these influences would greatly aid in providing a more accurate loss and, therefore, rate picture.

Unfortunately, most companies make no effort, at the present time, to explain and document these variables in their filings.

The above studies to determine changes in reserving policy or settlement rates do not have to be confined to state data. Regional or countrywide data could be used to determine a filer's reporting pattern and its reserving and claim settlement policies. The broad conclusions reached from these studies can be applied to the statewide data used in the filing, when relevant. The main point here is that these items should be considered since they do eventually impact on rate levels. Inclusion of some of these studies can improve a filer's credibility with its insureds and state agencies.

(iii) Changes in the policyholder profile

Loss growth patterns are significantly affected by the territorial and classification mix of the insurer. A shift in the policyholder profile will obviously bear on the results. Illustratively, a more urban insured profile will give rise to a more litigious claims picture in the future. This will in turn cause a more protracted loss growth pattern for the more recent accident years and hence, greater weight should be placed thereon (or alternatively projected loss development factors should be used) if this type of insured profile will continue in the future.

Similarly, if the filer has changed its policyholder profile by classification and is, for example, insuring more young drivers, the loss growth pattern should be different in the future than was the case historically.

These policyholder mixes would affect implicitly both claim settlement and reserving policies and to the extent possible their influence within these areas should be studied.

Statistical tests of significance can be made to determine if a correlation might exist between territorial or class splits and changes in the insurer loss growth pattern. The input for such tests may be in the form presented in Exhibit VIII and either a judgmentally or statistically based adjustment may be made in the filer's loss development factor.

In rate filings there is little, if any, evidence of this type of analysis at the present time. There is usually no indication as to what the current policyholder profile is, vis-a-vis territory and classification and how it has changed during the experience period. Furthermore, no attempt is made to correlate and quantify changes in the policyholder mix with factors affecting loss development patterns, such as propensity to sue and early reporting of claims.

(iv) Propensity for late reported claims

When performing a comprehensive reserve study, the actuary usually separates loss data by report year to determine and quantify changes in the development patterns of losses for claims reported early as compared with those reported late. The same type of analysis should be done for rate filings. Usually, late reported claims will exhibit characteristics different with respect to both development and trend from those which are reported during the accident year. As we indicated in the preceding subsection, such a study can and should be considered with changes in the exposure profile of a company. As a minimum, loss development data already submitted in a filing or prepared for a rate review should be broken down between claims reported within the accident or policy year and late reported to allow for a more detailed study. If these data lack sufficient credibility, regional or countrywide statistics could be used to document these effects.

(v) Packaged policies

With respect to packaged policies, we believe that cause of loss data should be made available so that the reviewer can make a determination as to exactly which direction the preponderance of losses are taking. If the liability component is increasing, for example, a more protracted loss growth pattern can be expected in the future and vice-versa. An exhibit such as that shown herein, i.e., Exhibit IX, could be prepared when reviewing a rate structure.

All of the above items are considered carefully when quantifying a particular company's reserve level for Annual Statements or other purposes. It seems logical to expect that similar care be given to quantifying these reserves on a by-state and subline basis for rate filing purposes.

B. Trend Factors

The current valuation of trend factors, in most cases, leaves much to be desired. Ordinarily, a table such as that set forth in Exhibit II is displayed and the problems attendant with the procedure followed using this table were discussed in Part 1 of this paper.

A demonstration of the biases caused by combining claims of different age groups is seen from Exhibit II-A. This table displays average paid claims cost data by maturity within accident year. It is demonstrated that there are different trends existing at various maturity levels. If we ascertain that the most recent average cost corresponds to an average settlement date of 2.0 years as compared with say an historical average of 3.0 years prevailing during the past, the implied trend factor would be different by, say, 10%. This is extremely prevalent where there has been a shift toward a policyholder profile of insureds who tend either to have their case closed earlier or later than in the company's past profile.

It must be recognized that there are multiple components comprising the trend factor. Each of the issues addressed above should be quantified, as much as possible, with respect to both loss frequency and severity. Furthermore, when exposure is measured by the amount of insurance purchased, premium trend factors must be used as an offset to the loss costs in order to recognize inflation resulting in increasing insurance-to-value.

When government indices are used, the filing should statistically correlate insurance company results (severity and frequency) with those using Consumer Price or related indices. Such a correlation can take the form described in Exhibit II-B by maturity within accident year. Once we establish this correlation, these external indications can be used and the argument that they are not relevant to insurance industry data can be diffused. Of course, these severity trends should always be used in conjunction with frequency trends to obtain a total pure premium trend picture.

C. Expense Provisions

The insuring public and state agencies will, for the most part, no longer accept the old notion that all expenses can be assumed to vary with losses and premiums as implied by the formula dictated by the old loss ratio method, viz:

$$\text{Indicated Rate Change} = \frac{\text{Rate Level Loss Ratio}}{\text{Expected Loss Ratio}} - 1.000$$

Recent rate models allow for a breakout between fixed and variable expenses and include the following general models:

a) $$\text{Indicated Rate Chg.} = \frac{\text{Rate Level Loss Ratio + Fixed Expense Ratio}}{\text{1 - Variable Expense Ratio}} - 1.000$$

b) $$\text{Indicated Rate Chg.} = \frac{\text{Rate Level Loss Ratio}}{\text{1 - Variable Expense Ratio}} + \left(\text{Fixed Expense Ratio} \times \text{Inflation Factor}\ \text{x}\right) - 1.000$$

It seems logical that a breakdown between fixed and variable costs would be appropriate. Certain expenses, such as taxes, underwriting profit and a portion of production costs are and should be computed as a percentage of premium. On the other hand, a portion of others, such as general overhead expenses (salaries and rent), are relatively fixed in nature and should be treated as such.

In addition to the above treatment of fixed and variable costs, it is important to point out that the present system of determination of expense provisions, as a ratio to premiums, leaves much to be desired. Normally an historical three year ratio of various expenses to written or earned premiums is examined as a provision selected for use in the future.

A more appropriate way to estimate this provision would be to project a budget for the coming year (the estimated dollars needed for various expense categories in a line for a state). This would result in a flat policy fee to be charged on each policy to be supplemented by other variable costs. The formula then for each class premium would be as follows:

$$\text{Gross Class Premium} = \frac{\text{Indicated Loss Cost For That Class}}{1 - \text{Variable Expense Ratio}} + \text{Policy Fee}$$

The above would remove the inconsistency of obtaining percentages of expenses to past calendar year premiums, which may be composites of many different rate levels for the particular year. For instance, if a company intends to increase its rates by 20% beginning next year and its general expense costs by only 10%, the general expense ratio should change from that estimated the year before.

The percentage used in the separating of expenses into the fixed and variable components can and should be quantified. These affect rate levels by a significant degree, as Exhibit III-A shows, using hypothetical data.

D. Reconciliation Between Rate Filings and Annual Statement Results

We indicated in Part 1 of this paper that if the methods used to estimate loss development factors had been followed in setting reserves on the financial level (i.e., for the Annual Statement) those liabilities might be much different than would presently be the case. In order to obviate this problem, a quantification of the various adjustments that are used in the calculation of Annual Statement reserves should be made and included as a separate memorandum in the rate filing.

Explicitly, then, the methodology used and in fact, the bottom-line reserves appearing in the Statement will tie in with those shown in the rate filing and a consistency would result. This consistency will serve to counter-balance the argument that consumer groups have in terms of the anomaly between Statement "profits" and losses claimed in rate filings. It will also allow for a more accurate representation of insurer results in the rate review process by reflecting the most likely more sophisticated techniques that were employed in the development of annual statement reserves.

In order to accomplish this goal, we suggest that a supplemental memorandum be included within each filing or rate review describing the reserving process used in the Statement and showing how the variables reflected therein were introduced as input in the filing or rate review under examination. Studies such as those explicitly set forth in Part 3 could be included in this analysis. Inclusion of such exhibits can ease approval of the filing and satisfy departmental inquiries regarding loss development and trend.

The issues of reflecting investment income in the ratemaking process has heretofore not been addressed because of the author's wish to keep within underwriting and actuarial areas. However, regardless of whether or not such income to whatever degree is reflected, it is important that insurers quantify this aspect correctly in rate filings or reviews. Exhibit X presents a brief description of the familiar cash flow approach and quantifies investment earnings on both loss and unearned premium reserves.

All of the above would, in the judgment of the writer, serve to ameliorate the relationship between the insurance company and the public and/or regulatory body.

The Casualty Actuarial Society in recent years has been blessed with numerous excellent studies regarding reserving methodologies. It is our belief that some part of these can and should be integrated in the rate-making process and in the actual filings to help estimate the insurer's liability for claims. This will provide a more accurate picture of ultimate losses and ultimately rate levels.

The report formats in Part 3 of this paper serve to highlight the information process which should, if possible, be included in rate filings and reviews and allow for reflection of these aforementioned studies.

Part 3: Recommended Report Formats To Be Included In Rate Filings

This section discusses tests which should be performed at every rate and reserve evaluation.

These two studies cannot be separated because, in most lines, losses normally comprise at least 60% of any insurer's gross rate and proper estimation of these losses requires production of an actuarially accurate reserve level by accident or policy year.

It is felt that these tests or reports could be used to help answer the following questions regarding an insurer's loss experience and the variables influencing this experience, particularly in the area of loss development, viz:

(1) Change in the company's reserving policy.

(2) Change in the company's claim settlement rates.

(3) Change in the company's policyholder profile (by class and territory).

(4) Change in the company's cause of loss for packaged policies.

(5) Change in the company's reporting patterns and trend factors (frequency & severity) by report year.

In order to be able to explicitly measure each of these variables and respond accurately to the changing conditions of the insurer in terms of the claim climate and policyholder mix, the following tests must be performed:

Test 1. Calculation of Reserve Adequacy

This test will estimate the effect of any changes in the filer's reserving policy affecting the adequacy of the case reserves. It is important, of course, that the data be presented separately by layer of losses (i.e., basic limits only, losses for amount of insurance range x to y, etc.) so that large claims do not distort the results.

The following format is suggested:

Average	Average Outstanding Cost For Limit at:				
Year	12 Mos.	24 Mos.	...	60 Mos.	72 Mos.
X					
X + 1					
.					
.					
.					
X + 9					

The average annual changes can be computed by dividing the "n"th average outstanding loss cost at each maturity by the "n-1"st as described in Exhibit IV of this paper to determine if a large change occurs at any one point.

Test 2. Calculation of Claim Disposal (or Settlement) Rate

As stated, this measures the filer's rate of settling claims and has an effect on both paid and incurred loss growth patterns. A calculation of disposal rate of claims to measure claim settlement practices of rate filers follow.

A. Definition of Disposal Rate (DR)

$$DR = \frac{NS}{NOB + NR}$$

where DR = Disposal Rate

NS = Number of Claims settled during calendar period

NOB = Number of Claims that were outstanding at the beginning of the period

NR = Number of Claims reported during the peirod

B. The following table would be prepared by the filer:

(1) Period (normally an accident year)	(2) Disposal Rates 0-12 Mos.	12-24 Mos.	24-36 Mos.
X			
X + 1			
X + 2			
X + 3			
.			
.			
.			

If report year data are unusable, then ratios of paid to total reported claims by maturity within accident years can be developed as described by Exhibits VII-A.

Test 3A. Calculation of Implicit Average Annual Change In Claim Costs

Implicit changes in claim costs could be obtained by examining the average incurred claim cost (by amount of insurance range x to y, basic limits losses, etc.) by maturity within accident year. The following table would be produced:

TABLE A

Accident Year	Average Loss Cost Reported (or closed) In 12 Mos. Ending With				
	12 Mos.	24 Mos.	...	60 Mos.	72 Mos.
X					
X + 1					
.					
.					
.					
X + 9					

The average cost at each maturity could be computed as a weighted arithmetic mean or by fitting a curve to the average costs at each maturity. Thus, claims reported or closed during the first 12 months may average an annual increase in cost of 5%, those closing the second 12 months, 10% and so on. An overall average annual change in cost would then be computed by obtaining a weighted average date of reporting or closure (i.e., payment) underlying the historical period studied. This "average" maturity in the past could be compared to that estimated presently on reflecting the current disposal rate to estimate the overall value of the trend factor to be used.

Hence, if Test 2 results indicate a speedup of about 12 months in the average date of settlement and historical data indicated a trend factor of 10% on an average 36 month closure and a 5% trend on a 24 month closure, then a trend factor closer to 5% might be indicated.

The above trend calculation only refers to claim severity and hence measures only one-half of the insurance inflation index. Claim frequency should also be considered and this could be determined by maturity within accident year in Test 3B.

Test 3B. Calculation of Average Claim Frequency Per Exposure

Accident Year	No. of Earned Exposures	Ratio of the No. of Claims Rptd. (or Settled) To Column (1) as of: 12 Mos.	24 Mos.	...60 Mos.	72 Mos.
X					
X + 1					
.					
.					
.					

The combination of the maturity changes of frequency with those of costs in Test 3A would determine the insurance inflation component separately for each maturity keeping in mind the basic laws of ratemaking: Pure Premium = Claim Severity x Claim Frequency and Total Insurance Inflation = Claim Cost Trend x Claim Frequency Trend.

Test 4. Compilation of Data Reflecting Changes In A Company's Book of Business

This is a very critical area which, unfortunately, has not received the attention it should have in the past.

If a filer has changed its book of business in the most recent year, the experience for years prior thereto becomes much less relevant to future situations and an adjustment (either qualitatively or quantitatively) must be made to reflect these changes. Some of the data required which would determine whether the filer has changed its book follows:

(1) Distribution By Class

(2) Distribution By Territory

(3) Distribution By Liability Limits or Deductible

The above are all very important in analyzing experience. When considering the situation by class, different types of insureds have different propensities to sue and if there has been a class shift, then the development factors obtained would be affected. The insurer should be prepared to correlate changes in its business mix with changes in its trend, development and other results. Following are types of formats which should be presented as part of any rate filing.

Compilation of Data Used To Determine A Shift In The Book of Business

Class	Number of Car Years Earned (or Other Exposure Measure) During Year X	X+1	X+2	X+3	X+4
1					
.					
.	A. Distribution By Class				
.					
n					

Territory	X	X+1	X+2	X+3	X+4
1					
.					
.	B. Distribution By Territory				
.					
n					

These are only "marginal" distributions in statistical terminology, hence would not disclose an interchange of classes of insureds between territories where territorial totals and class totals remained unchanged. If such a development is suspected, a two-way classification should be prepared for each year.

Year	Basic	50/100	100/300	50 Ded. Coll.	100 Ded. Coll.
X					
X+1					
X+2					
.	C. Distribution By Liability Limits or Deductible Coverage				
.					
.					
X+4					

Test 5. Interrogatories

The tests usually performed produce a number of questions, the answers to which should be made part of each report. Suppose the data seem to imply that a reserve policy change took place during the experience period. The company should have the opportunity to respond to the indications. It is possible that the data may be misleading and the company should have a chance to rebut.

The following questions may be included in an Interrogatory section:

1) Has there been any change in reserving policy during the experience period to make reserves more or less adquate?

2) Has there been a change in company's claim settling policy, either faster or more slowly?

3) Has there been any change in the company's system of reporting claims?

4) Has there been a change in the company's claim adjustment procedures, tactics, or policies?

5) Has there been any change in the book of business by territory, classification or by policy offering (higher or lower deductibles, policy limits)?

6) How have the above been reflected explicitly in the development of historical incurred losses to an ultimate settlement basis?

Exhibit I

Derivation of Reserves Using Historical Incurred Losses

PART A. Losses Incurred as of Maturity

Year	1	2	3	4	5
1974	250,000	375,000	487,500	560,625	588,656
1975	300,000	435,000	543,750	598,125	
1976	325,000	463,125	567,328		
1977	350,000	481,250			
1978	400,000				

Link Ratios

Year	1-2	2-3	3-4	4-5
1974	1.500	1.300	1.150	1.050
1975	1.450	1.250	1.100	
1976	1.425	1.225		
1977	1.375			

PART B. Average Link Ratios

	1-2	2-3	3-4	4-5
Selected:	1.438	1.258	1.125	1.050

Completion Ratios

1-Ult.	2-Ult.	3-Ult.	4-Ult.
2.137	1.486	1.181	1.050

PART C. Calculation of Ultimate Losses & Reserve Levels

Year	(1) Losses Incurred To Date	(2) Completion Ratio	(3) Ultimate Incurred Losses	(4) Losses Paid To Date	(5) Indicated Reserve
1974	588,656	1.000	588,656	588,656	0
1975	598,125	1.050	628,031	541,875	86,156
1976	567,328	1.181	670,014	450,328	219,686
1977	481,250	1.486	715,138	319,250	395,888
1978	400,000	2.137	854,800	188,950	665,850
TOTAL			3,456,639		1,367,580

Exhibit II

AUTOMOBILE LIABILITY INSURANCE
Private Passenger Cars

Average Paid Claim Cost Data

Year Ended	$10,000 Bodily Injury	Total Limits Property Damage	Total Limits Medical Pymts.
6/30/75	1,623	373	403
9/30/75	1,666	383	407
12/31/75	1,721	391	411
3/31/76	1,771	399	416
6/30/76	1,811	407	423
9/30/76	1,836	417	434
12/31/76	1,867	429	446
3/31/77	1,901	440	459
6/30/77	1,946	453	468
9/30/77	1,990	467	475
12/31/77	2,025	480	483
3/31/78	2,047	494	491
3/31/78 Claims	202,333	1,446,868	91,464
Avg. Annual Chg.	+8.7%	+10.7%	+8.1%

Average Paid Claim Frequency Data

(Claim Frequency Per 100 Cars)

Year Ended	Bodily Injury	Property Damage
6/30/72	1.9487	7.2151
9/30/72	1.9103	7.2084
12/31/72	1.8622	7.2010
3/31/73	1.7924	7.0722
6/30/73	1.8091	7.3311
9/30/73	1.7845	7.3780
12/31/73	1.7018	7.1910
3/31/74	1.6591	7.0924
6/30/74	1.5682	6.9167
9/30/74	1.5408	6.8727
12/31/74	1.5824	7.0670
3/31/75	1.5831	7.0202
6/30/75	1.6222	7.1884
9/30/75	1.6269	7.2716
12/31/75	1.6018	7.2865
3/31/76	1.5720	7.2697
6/30/76	1.5608	7.1284
9/30/76	1.5569	6.9747
12/31/76	1.5729	6.7731
3/31/77	1.5765	6.7320
6/30/77	1.5397	6.5212
9/30/77	1.5019	6.3103
12/31/77	1.4598	6.1057
3/31/78	1.4330	5.9851
3/31/78 Claims	202,333	1,446,868
Avg. Annual Chg.	-4.1%	-2.3%

Exhibit II-A

Accident Year	Average Paid* Claim Cost For Claims Closed In Maturity 1	2	3	4
1974	100	115	160	170
1975	110	135	210	290
1976	121	159	280	
1977	133	188		
1978	145			
Avg. Annual Chg.	10%	+20%	+30%	+60%

* or incurred

Exhibit II-B

Date 1	Insurance Company Claim Costs*			Government Index **
	Acc.Yr. 1977	Acc.Yr. 1978	Acc.Yr. 1979	
3/77				
6/77				
9/77				
12/77				
3/78				
6/78				
9/78				
12/78				
3/79				
6/79				
9/79				
12/79				

* Adjusted for changes in the deductible mix.

** Such as Modified Consumer Price Index, Construction Cost Index, etc.

Exhibit III

Historical Derivation of Expense Ratios and Reflection In The Traditional Loss Ratio Method of Ratemaking

	1977	1978	1979
Written Premium	1,000,000	1,200,000	1,500,000
Earned Premium	900,000	1,100,000	1,400,000
Total Production Costs	200,000	252,000	300,000
General Expense	100,000	110,000	130,000

Expense Ratio

				Mean
Production Costs of Written Premium	20.0%	21.0%	20.0%	20.3%*
**General Expense of Earned Premium	11.1%	10.0%	9.3%	10.1% Ø

* Use 20%

** Ø Use 10%

Exhibit III-A

	Percentage of Premium		
	Variable Data	Fixed Data	TOTAL
a) Commission & Brokerage	20%	0%	20%
b) Other Acquisition	1-1/2%	1-1/2%	3%
c) General Administration	5%	5%	10%
d) Taxes	2%	0%	2%
e) Profit	5%	0%	5%
f) Total Expense Ratio	33-1/2%	6-1/2%	40%
g) Expected Loss Ratio			60%

If the rate level loss ratio is 70%, the indicated change under the old loss ratio methods is +16.7%, viz:

$$\text{Indicated Rate Change} = \frac{.700}{.600} - 1.000 = .167$$

If we split between fixed and variable expense, we would obtain a +15.0% change, viz:

$$\text{Indicated Rate Change} = \frac{.700 + .065}{1.000 - .335} - 1.000 = .150$$

If we accept the idea that producers should receive at least a partially fixed commission for each policy (say 10% fixed, 10% variable instead of 20% variable as above), we obtain a +13.1% change, viz:

$$\text{Indicated Rate Change} = \frac{.700 + .065 + .100}{1.000 - .235} - 1.000 = .131$$

Exhibit IV

Derivation of Rate of Change of Outstanding Loss Cost

PART A. Losses Outstanding as of Maturity

Year	1	2	3	4	5
1974	75,000	80,000	75,000	45,000	0
1975	86,250	100,000	90,000	56,250	
1976	125,050	120,000	117,000		
1977	156,350	162,000			
1978	211,050				

PART B. Number of Losses Outstanding as of Maturity

Year	1	2	3	4	5
1974	50	40	30	15	0
1975	50	40	30	15	
1976	50	40	30		
1977	50	40			
1978	50				

PART C. Average Outstanding Loss Cost as of Maturity

Year	1	2	3	4
1974	1,500	2,000	2,500	3,000
1975	1,725	2,500	3,000	3,750
1976	2,501	3,000	3,900	
1977	3,127	4,050		
1978	4,221			

PART D. Rate of Change of Outstanding Loss Costs

Year	1	2	3	4
74-75	1.150	1.250	1.200	1.250
75-76	1.450	1.200	1.300	
76-77	1.250	1.350		
77-78	1.350			

Average Change

1	2	3	4
1.300	1.267	1.250	1.250

Exhibit V-A
Sheet 1

<u>Test For Change In Adequacy of Case Reserves</u>

Incurred losses are composed of both paid and outstanding losses. Hence, any change in the strength of case reserves will affect incurred losses and, therefore, will distort any analysis performed on them. The way to test for changes in the adequacy of case reserves is to examine trends in the size of average outstanding loss costs over time. Such changes can occur in two ways. First, there can be a slow increase in the strength of reserves over a number of years. If this is the case, then average outstanding loss costs will be increasing at a rate faster than total average loss cost. For instance, the former may be increasing at 25% a year while the latter increases at only 10% a year. In contrast to this, the reserves strengthening may be a one time phenomenon. This would show up as a large increase in the average outstanding loss costs for all accident years in one particular calendar year. The way to correct this is to adjust all case reserves to the same adequacy level. This is usually done by starting with the most recent average outstanding loss cost for each maturity and then trending back over time using an appropriate factor. When dealing with workmen's compensation insurance, the same procedure would be utilized except that law amendment benefit factors would be utilized in place of trend factors.

Exhibit V-A
Sheet 2

An example will help clarify these concepts. Exhibit I shows the derivation of loss reserves utilizing actual incurred losses. In Exhibit IV, the rate of change of average outstanding losses is determined. The change is in the area of 25% to 30%. Let us assume that external data show that costs are increasing only 10% a year. From this we can conclude that the company has adopted a policy of gradually increasing its reserve adequacy. In Sheet 3 of this exhibit a set of adjusted average outstanding loss costs are derived. This was done by trending the latest average outstanding loss cost back in time at 10% a year. These adjusted outstanding losses were then utilized to derive adjusted incurred losses. The latter are shown in Exhibit V-B. Application of the link ratio technique to the adjusted losses yields a reserve 19% lower (or incurred losses 10% lower) than that obtained in Exhibit I using the unadjusted losses.

Exhibit V-A
Sheet 3

Derivation of Adjusted Average Outstanding Losses

Adjusted Outstanding Loss Cost as of Maturity*

Year	1	2	3	4	5
1974	2,883	3,043	3,223	3,409	0
1975	3,171	3,347	3,545	3,750	
1976	3,488	3,682	3,900		
1977	3,837	4,050			
1978	4,221				

Adjusted Outstanding Losses as of Maturity**

Year	1	2	3	4	5
1974	144,150	121,720	96,690	51,135	0
1975	158,550	133,880	106,350	56,250	
1976	174,400	147,280	117,000		
1977	191,850	162,000			
1978	211,050				

* Using the latest calendar year's (i.e., last diagonal) average outstanding loss cost from Part C of Exhibit IV trended back by 10% per year by maturity within accident year. Thus, 4,221 ÷ 1.10 = 3,837, 3,837 ÷ 1.10 = 3,488, etc. for maturity 1. For maturity 2, 4,050 ÷ 1.10 = 3,682, etc.

** Adjusted average outstanding loss cost multiplied by corresponding outstanding claim count from Part B of Exhibit IV.

Exhibit V-B

Derivation of Reserves Using Adjusted Incurred Losses

Adjusted Losses Incurred as of Maturity

Year	1	2	3	4	5
1974	319,150	416,720	509,190	566,760	588,656
1975	372,300	468,880	560,100	598,125	
1976	374,350	490,405	567,328		
1977	385,500	481,250			
1978	400,000				

Adjusted Link Ratios

Year	1-2	2-3	3-4	4-5
1974	1.306	1.222	1.113	1.039
1975	1.259	1.195	1.068	
1976	1.310	1.157		
1977	1.248			

Adjusted Average Link Ratios

1-2	2-3	3-4	4-5
1.281	1.191	1.091	1.039

Adjusted Completion Ratios

1-Ult.	2-Ult.	3-Ult.	4-Ult.
1.729	1.350	1.134	1.039

Year	Losses Incurred To Date	Adjusted Completion Ratio	Adjusted Ultimate Incurred Losses	Losses Paid To Date	Adjusted Indicated Reserve
1974	588,656	1.000	588,656	588,656	0
1975	598,125	1.039	621,452	541,875	79,577
1976	567,328	1.134	643,350	450,328	193,022
1977	481,250	1.350	649,688	319,250	330,438
1978	400,000	1.729	691,600	188,950	502,650
Total			3,194,746		1,105,687

Exhibit VI

Derivation of Reserves Using Historical Paid Losses

PART A. Losses Paid as of Maturity

Year	1	2	3	4	5
1974	100,000	250,000	350,000	420,000	462,000
1975	150,000	330,000	429,000	493,350	
1976	175,000	350,000	437,500		
1977	200,000	390,000			
1978	250,000				

Link Ratios

Year	1-2	2-3	3-4	4-5
1974	2.500	1.400	1.200	1.100
1975	2.200	1.300	1.150	
1976	2.000	1.250		
1977	1.950			

PART B. Average Link Ratios

1-2	2-3	3-4	4-5
2.163	1.317	1.175	1.100

Completion Ratios

1-Ult.	2-Ult.	3-Ult.	4-Ult.
3.682	1.702	1.293	1.100

PART C. Calculation of Ultimate Losses and Reserves

Year	Losses Paid To Date	Completion Ratio	Indicated Reserve	Ultimate Losses
1974	462,000	1.000	0	462,000
1975	493,350	1.100	49,335	542,685
1976	437,500	1.293	128,188	565,688
1977	390,000	1.702	273,780	663,780
1978	250,000	3.682	670,500	920,500
TOTAL	2,032,850		1,121,803	3,154,653

Test For Change In Rate of Payment

In this method, the ratio of the number of paid claims to ultimate claims is measured by maturity within accident year. An upward trend in the data indicates that claims are being disposed of more rapidly and vice-versa. The way to correct for this is to adjust the paid loss data so that the same proportion of claims are paid for all accident years at each maturity stage. The procedure to be followed can be illustrated by a single example using the hypothetical paid loss development data set forth in Exhibit VI. These data were analyzed using the normal paid link ratio pattern described earlier for incurred losses. Use of the average growth factors yielded a reserve level in Exhibit VI of $1,121,803 and ultimate incurred losses of $3,154,653.

Sheet 4 of Exhibit VII-A displays the accident year/ maturity fractions of paid to incurred number of claims underlying the Exhibit VI data. Examination of this table illustrates that the insurer whose data are used is apparently paying claims at a more rapid rate than it has in the past.

Using these results, we produce an adjusted set of paid loss data in Exhibit VII-B by interpolation. For

example, the adjusted accident year 1974 paid losses as of maturity 1 are calculated as follows:

$$200{,}000 = 100{,}000 + \frac{.50 - .30}{.60 - .30} \times (250{,}000 - 100{,}000)$$

where

100,000 = losses paid for accident year 1974 as of maturity 1

250,000 = losses paid for accident year 1974 as of maturity 2

.50 = adjusted ratio of number of paid to ultimate losses as of maturity 1

.30 = ratio of number of paid to ultimate losses for accident year 1974 as of maturity 1

.60 = ratio of number of paid to ultimate losses for accident year 1974 as of maturity 2

Similarly, the adjusted losses paid as of maturity 2 for accident year 1975 is calculated as follows:

$$396{,}000 = 330{,}000 + \frac{.75 - .65}{.80 - .65} \times (429{,}000 - 330{,}000)$$

These adjusted losses are analyzed in Exhibit VII-B in developing an alternative reserve level. It should be noted that link ratios derived using the adjusted losses are much more stable than those calculated using historical losses. This is to be expected since the adjusted losses reflect the same rate of claim payment for all accident years. The reserve derived utilizing the adjusted data is only $521,559. This is a reduction of 54% below the figure obtained before in Exhibit VI.

It should not be construed from this example that all changes in payment rate will produce such dramatic reserve changes. However, it should be obvious that significant distortions can arise when historical data are employed without adjustment.

Hence, the general procedure that emerges when employing paid data is:

(1) Test to see if the rate of payment of claims has changed. If so, then

(2) Derive an adjusted loss payment data set.

(3) Use the adjusted figures to determine the reserve requirements.

A more exact way to determine claim settlement rates utilizes report year data. This technique involves measurement of the fraction of claims available for payment in a given time period that are actually paid. The number of claims available for payment is usually taken as the number of claims outstanding at the beginning of the period plus the number of claims reported during the period. Sheet 6 of Exhibit VII-A displays the calculation of disposal rates for the hypothetical company already used. As can be seen, this test confirms the fact that claims are currently being paid off faster than in the past. Therefore, this technique also implies that the historical paid losses should be adjusted.

Exhibit VII-A
Sheet 5

Test For Changes In Rate of Claim Payment

	Assumed Ratio of Number of Paid To Number of Ultimate Total Losses As of Maturity				
Year	1	2	3	4	5
1974	.30	.60	.75	.90	1.00
1975	.35	.65	.80	.95	
1976	.40	.70	.85		
1977	.45	.75			
1978	.50				

Exhibit VII-A
Sheet 6

Calculation of Disposal Rate

Number of Losses Paid as of Maturity

Year	1	2	3	4	5
1974	30	60	75	90	100
1975	35	65	80	95	
1976	40	70	85		
1977	45	75			
1978	50				

Number of Claims Outstanding as of Maturity

Year	1	2	3	4	5
1974	50	35	25	10	0
1975	45	30	20	5	
1976	40	25	15		
1977	35	20			
1978	30				

Number of Claims Reported During Maturity

Year	0-1	1-2	2-3	3-4	4-5
1974	80	15	5	0	0
1975	80	15	5	0	
1976	80	15	5		
1977	80	15			
1978	80				

Disposal Rate During Maturity

Year	0-1	1-2	2-3	3-4	4-5
1974	0.375	0.462	0.375	0.600	1.000
1975	0.438	0.500	0.429	0.750	
1976	0.500	0.545	0.500		
1977	0.563	0.600			
1978	0.625				

Exhibit VII-B

Derivation of Reserves Using Adjusted Paid Losses

Adjusted Losses Paid as of Maturity

Year	1	2	3	4	5
1974	200,000	350,000	396,667	441,000	462,000
1975	240,000	396,000	450,450	493,350	
1976	233,333	379,167	437,500		
1977	231,667	390,000			
1978	250,000				

Adjusted Link Ratios

Year	1-2	2-3	2-3	4-5
1974	1.750	1.133	1.112	1.048
1975	1.650	1.138	1.095	
1976	1.625	1.154		
1977	1.683			

Adjusted Average Link Ratios

1-2	2-3	3-4	4-5
1.677	1.142	1.104	1.048

Adjusted Completion Ratios

1-Ult.	2-Ult.	3-Ult.	4-Ult.
2.216	1.321	1.157	1.048

Year	Paid Losses To Date	Adjusted Completion Ratio	Adjusted Indicated Reserve	Ultimate Losses
1974	462,000	1.000	0	462,000
1975	493,350	1.048	23,681	517,031
1976	437,500	1.157	68,688	506,188
1977	390,000	1.321	125,190	515,190
1978	250,000	2.216	304,000	554,000
TOTAL ,	2,032,850		521,559	2,554,409

Exhibit VIII

Territory Groupings	(1) Policyholder or Premium Concentration			Changes In * Between Accident Years	
	1977	1978	1979	(a)	(b)
(a) Rural					
(b) Suburban					
(c) Urban					

Note: In this example, territories have been placed in one of three broad groups: Rural, Suburban, and Urban.

* Can either be average outstanding loss costs at comparable maturities, average settlement rates or other measures affecting loss growth patterns.

Cause of Loss	Percentage of Ultimate Losses For Accident Year 1976	1977	1978	Projected 1979	Selected Loss * Dev. Factor
1. Liability	10	15	20	25	1.60
2. Theft	40	30	35	35	1.00
3. Fire	40	35	35	30	1.00
4. Other	10	20	10	10	1.00

Overall Development Factor:

25% of 1.60 + 35% of 1.00 + 30% of 1.00 + 10% of 1.00 = 1.15

* After consideration of the variables discussed in Part 2 concerning claim settlement rates, reserving policy, policyholder mix, etc.

The cash flow discount model shown for the experience of a company in Sheet 3 of Exhibit X produces an investment income offset to the gross premium to reflect reserves for losses and unearned premiums.

Column (b) of Sheet 3 sets forth expected percentages of ultimate losses paid during each of the calendar year periods beginning with day 1 of the policy year studied. Thus, 3% of the ultimate losses are paid within 12 months after the start of the year, 7% from month 13-24, etc.

Column (c) estimates the amount of time from the start of the policy year the money was available for investment. Columns (d) and (e) represent the discounted payments at 9% and 10% rates of return, respectively.

Losses and loss expenses (Line (14)) comprise 86% of the total premium dollar for the client and thus, taxes and general expenses production and profit are considered separately in Lines (16) and (17).

Premiums are normally received between 60 and 90 days after inception of a policy. We have assumed 2/10ths of a year as the average. Regular commissions (zero, in our case) are deducted from premiums remitted by agents. Hence, the insurer never holds this money.

Taxes are paid following the end of the year and we have assumed March 1st of the following year as the date of payment.

In accordance with this schedule, all loss and expense payments and underwriting profit are discounted for interest back to the mid-point of the policy year to give us the Present Value of Outgo. Subtracting this from the correspondingly discounted value of premiums less commissions gives the Present Value of Income Less Present Value of Outgo.

Calculation of Potential Income Through Present Values as of the Midpoint of a Policy Year (July 1) of All Income and Outgo @9% and 10% Interest

Line		(a) Years From Start of Policy Year	(b) Yearly Percent of Losses Paid	(c) Years of Interest Discount	(d) Discount @9%	(e) Payments @10%
(1)		1	3%	.2	2.95	2.94
(2)		2	7	1.0	6.42	6.36
(3)		3	14	2.0	11.78	11.57
(4)		4	20	2.9	15.58	15.17
(5)		5	16	3.9	11.43	11.03
(6)		6	11	4.9	7.21	6.90
(7)		7	8	5.9	4.81	4.56
(8)		8	7	6.9	3.86	3.63
(9)		9	5	7.9	2.53	2.35
(10)		10	3	8.9	1.39	1.28
(11)		11	3	9.9	1.28	1.17
(12)		12	3	10.9	1.17	1.06
(13)	Total				70.41	68.02

(14) Expected loss and loss expense ratio			.860	.860
(15) Present value of payments (13) x (14)			60.55	58.50
(16) Taxes, as percent of Premium	2.0	.667	1.89	1.88
(17) General expenses, other production and profit	12.0	0	12.0	12.0
(18) Total present value of outgo (15) + (16) + (17)			74.44	72.38
(19) Premiums less commissions	100.0	.2	98.29	98.11
(20) Present value of income less present value of outgo (19) - (18)			23.85	25.73
(21) Line 20 as a percentage of losses (20) ÷ (14)			27.7%	29.9%

NOTE: $(d) = (b) \times \left(\frac{1}{1.09}\right)^{(c)}$ // $(e) = (b) \times \left(\frac{1}{1.10}\right)^{(c)}$

Part 4: Sample Explanation of Ratemaking Procedures

Insurance rates in general consist of two parts:

1) Expected Losses

2) Expenses

Unfortunately, although losses normally make up approximately 70% of the total premium dollar, they are not fully known by the time rates are set.

Rather, actuaries have to go through a projection process to estimate what the losses will be for a particular year of coverage. Such losses are an approximation of results from past years. Hence, actual losses from these past years are adjusted for the following items.

1) Development Factors

These are historically-derived ratios which adjust losses arising from claims reported to date to reflect losses from claims not yet reported and changes in the valuation of known claims. Insurance company claims personnel initially estimate the ultimate value of claims based on data which are not yet complete. The difference between these estimates and the first value of claims is referred in the insurance industry as "development on known claims".

Thus, loss development factors adjust historical losses for a particular year of coverage to reflect losses not yet reported and changes in the valuation of known claims.

2) Trend Factors

Trend factors are used to project ultimate losses to claim severity and frequency levels expected to prevail to the future. There are two types of insurance inflation:

a) claim frequency = the probability of having a claim

b) claim severity = the cost of the claim once it occurs

Each component varies and is projected separately to reflect these future conditions.

3) Expenses

After losses are adjusted using development and trend factors, expenses of an insurer's operation are added in to arrive at a gross premium value.

ACCTUARIAL ISSUES TO BE ADDRESSED
IN PRICING INSURANCE COVERAGES
By E. James Stergiou

REVIEWED BY Sheldon Rosenberg and Aaron Halpert

As most actuaries that have had an opportunity to prepare a rate filing will tell you, the ratemaker will generally have to convince three principals that the rate he is generating is a reasonable one. First, he must convince himself. This first step alone is, in many cases, a difficult and laborious task due to the many technical uncertainties with which an actuary must deal. However, only when this task is accomplished can one proceed to the next level of review.

Second, company management must be in agreement with the conclusions offered by the actuary. The questions posed by management will generally deal less with technical specifics of how the rate indications were developed and more often with whether due consideration has been given to past or proposed changes in all aspects of company policy as it affects the claims department, marketing department, underwriting department or others.

Finally, having gained the blessing of his management, the actuary must also receive approval for the filing from the respective regulatory agency. The regulator has the responsibility of seeing that rates promulgated by him are adequate and neither excessive nor unfairly discriminatory. At this point the actuary must be able to defend any judgement made within the rate filing; be it expense provisions, classification criteria, etc.

It is in dealing with this third level of review, that Mr. Stergiou provides us with much helpful advice in how to prepare a better rate filing.

He discusses and provides several examples of variables that he feels currently receive summary treatment at best and are more often perhaps totally ignored in the preparation of a rate filing. Yet, several of these factors may, in his opinion, have a direct and significant impact on the bottom line results.

The thrust of the author's points are that consideration of these variables will lead to better understanding on the part of regulators and the insureds they represent, and thus facilitate receiving approvals of needed rate revisions. One must wonder however to what extent action taken by regulators are based on considerations that are well beyond the scope of the technical arguments being presented.

We feel that Stergiou's paper would have been more effectively presented, had he keyed his remarks toward that important first level of review. Certainly the elements of ratemaking referenced in this paper relate to the technical soundness of a rate review; they should therefore be geared toward the ratemaker, to be used as a tool to convince himself that the answer he gets is a realistic and non-biased one.

Another issue to be raised in light of the additional exhibits the author wishes to see incorporated in rate reviews is that an actuary may never have the time <u>nor</u> the need to look at all the pieces of information that may impact a given rate indication. The key juncture therefore becomes the point at which the data base is defined and report specifications are prepared. The actuary must always leave himself the option of looking at data in a specific format, but may actually exercise that option only if and when it becomes necessary to do so. This issue will have particular meaning in the area of loss development as will be discussed later.

Another comment pertaining to the entire paper, is that the author's purpose would have been better served had he used examples incorporating actual data rather than hypothetical data. The latter, although designed to make a point, may not be true representations of the "real world." An example based on actual annual statement data would have been particularly effective in illustrating the authors contention that applying the traditional loss development approach to data in Part 2 of Schedule P of a company's statement will lead to projections of ultimate losses that are "far different than those predicted in the company's balance sheet."

Comments of a specific nature will now be addressed to several of the issues mentioned in Mr. Stergiou's paper.

Loss Development Factors

The method of developing losses by analyzing historical age-to-age valuations of incurred (or paid) losses is evaluated by Stergiou. He points out several instances where this approach may lead to a biased forecast of ultimate losses. But does it? Let's examine two of his examples carefully. His first example deals with the case of a company that has changed its reserving adequacy gradually over time. To show how a bias may result, he first employs the traditional loss development approach in Exhibit I to project the needed reserves for the years listed. He then adjusts the outstanding losses by assuming the latest diagonal to be indicative of the company's present reserving practice. These most recent reserves are "detrended" at 10% annually in order to estimate what loss reserves in prior years would have been had the company used its current reserve practices. After adjusting the outstanding losses he adds back the paid losses and recalculates the needed reserves on Exhibit V-B. These ultimate incurred losses based on the adjusted outstanding losses are 8% lower than the ultimate incurred losses on Exhibit I.

We do not agree though that this comparison between Exhibits I and V-B is proper. If one examines the link ratios on Stergiou's Exhibit I, it becomes evident that the chosen ratios should not be the average of the link ratios in each column. One chooses the average only when several elements are believed to be sample estimates of the same underlying value. In this example, there is clearly a downward trend in the link ratios over time. This by itself is fairly conclusive evidence that the company is becoming more accurate in setting its initial reserves. Thus rather than using an average of historical age-to-age loss development factors, an actuary faced with the figures on Exhibit I might use the link ratios based on the latest available information (the last link ratio in each column). Perhaps he would even project a trend in these ratios and use a ratio lower than that of the latest years. Had the last link ratio in each column of Exhibit I been used instead to project ultimate losses, the projected losses would have been 3,330,920. This number serves as a more reasonable comparison to Stergiou's result on Exhibit V-B.

The point is, a bias exists only if the traditional loss development procedure is used blindly without examining the numbers for trends. If a trend of the type in Exhibit I exists, then it can be incorporated into the procedure in conjunction with information derived from the more detailed exhibits presented by Stergiou.

The same comment applies to the second example presented in Exhibit VI of Mr. Stergiou's paper. In that exhibit, loss development factors are based on historical movements in paid losses. After analyzing disposal patterns on Exhibit VI-A, and realizing that the company is currently closing claims at a faster rate than during the earlier experience period, the author adjusts the historical paid losses in Exhibit VII-B so that the underlying pay-out pattern for all years is the same.

Once again, the downward trend in the link ratios on Exhibit VI would have yielded similar information. The average should not (and in most rate reviews would not) be chosen as the representative link ratio. Again, the main point is that prominent changes in the company's handling of claims or reserves are usually evident from the traditional loss development diagonals themselves. In these cases additional information may be required to aid in the selection of representative link ratios. In the examples cited, an actuary choosing the average link ratios would not just-risk losing credibility with regulators but would more importantly derive a wrong answer.

Regarding the examples themselves, we feel that the reader is left somewhat confused in proceeding from the example based on incurred losses (Exhibit I) to the example based on paid losses (Exhibit VI). Unless told otherwise, one believes they are both based on the same experience. The author would best serve the reader by stating clearly that they are not.

It should be mentioned that Stergiou's method for adjusting the outstanding losses derives no information from reserves prior to the latest diagonal. While the evidence of a 30% trend in average outstanding losses (when overall inflation is assumed to be 10%) may signal a change in reserve adequacy, some information may still be derived from prior diagonals. One way to do this would be to multiply each of the earlier average outstanding losses by $(1.3/1.1)^n$ where n is the number of years between the evaluation of the reserve and the latest evaluation date. In this way outstanding losses would be on the same "adequacy level" and yet yield independent pieces of information. The analogue of this is when one uses several policy years of data in reviewing liability rates. Because each year is at a different cost level, a trend factor is applied to each year's data. However, the earlier years are not set to be equal to the latest year divided by the trend factor. Instead all the years are used to derive a trend factor and after the trend is applied, each year's information is used in setting the rates.

The adjustment made on Exhibit VII-A to reflect the change in the company's settlement rate raises an interesting question. The author mentions that "although use of report year data....is preferable, many insurers do not have such data readily available. However, the Exhibit VII-A (Sheets 1-5) calculation is usually available and can be used for our purpose." The key question is, without any information regarding reported claims, how can the ratios on Exhibit VII-A, Sheet 5 be hypothesized. Specifically, how does one assume that 50% of all claims to ultimately be reported for Accident Year 1978 are paid as of the first maturity. It would seem therefore that an adjustment based on disposal ratios (which are perhaps difficult to retrieve within a company's data base, but are actual numbers rather than assumed ratios) would be preferable.

One must also be careful in defining cases where it would be proper to apply the author's adjustment to paid losses. For example had the numbers in his Exhibit VII-A, Sheet 5 been changed only slightly, the resulting adjustment would lead to questionable results. The author claims that the adjustment leads to more stable link ratios. If the numbers in column 2 of Exhibit VII-A, Sheet 5 were changed to read as follows:

Year	1	2	3	4	5
1974	.30	.40	.75	.90	1.00
1975	.35	.45	.80	.95	
1976	.40	.50	.85		
1977	.45	.55			
1978	.50				

then the general arguments in favor of making the author's adjustment would still hold (i.e. disposal ratios are still increasing over time). However, the derived first to second link ratios after the adjustment would be

	1-2
1974	.732
1975	.853
1976	1.036
1977	1.322

As can be seen, the author's adjustment does not lead to stable results in this case.

Trend Factors

Stergiou provides an interesting example of how calendar year paid claim cost trend factors may be distorted when the average settlement date is changing over time. While the assumption in his example that average paid claim cost increases with each maturity level within an accident year is certainly a familiar assumption, the idea that significantly different trends exist at each maturity is somewhat surprising. It is important to note though that the distortion in trend can occur even when the underlying severity trend is the same at each maturity level. This will happen when there is a sudden increase in claim frequency. To see this, consider the simple case where 50% of ultimate claims are closed in the year they are incurred, 50% in the following year. The underlying severity trend is 10% for both types of claims. Also assume that 100 claims are incurred each year except the latest year when 200 claims are incurred. The loss data may look as follows:

		Average Claim Cost For Claims Paid	
Accident Year	Number of Claims Incurred	During Year	In Following Year
1970	100	1000	2000
1971	100	1100	2200
1972	100	1210	2420
1973	100	1331	2662
1974	100	1464	2928
1975	100	1611	3222
1976	100	1772	3544
1977	100	1949	3898
1978	200	2144	4288

The corresponding calendar year paid claim cost data would then be as follows:

Calendar Year	Average Paid Claim Cost
1971	1550.00
1972	1705.00
1973	1875.50
1974	2063.00
1975	2269.50
1976	2497.00
1977	2746.50
1978	2728.67

Notice that the paid claim cost entry for the latest calendar year is distorted due to the fact that it contains an artificially large number of smaller claims (closed within the accident year). Thus, we see there are other good and sufficient reasons to exercise care when using paid claim cost data for trend, particularly in lines with a long payout pattern.

Expenses

Two models for calculating rate changes are presented by the author as alternatives to the loss ratio procedure currently used. The current procedure is:

$$\text{Indicated Rate Change} = \frac{\text{Rate Level (Indicated) Loss Ratio}}{\text{Expected Loss Ratio}} - 1$$

The alternative models presented are:

a) Indicated Rate Change =

$$\frac{\text{Rate level Loss Ratio + Fixed Expense Ratio}}{\text{1 - Variable Expense Ratio}} - 1$$

b) Indicated Rate Change =

$$\frac{\text{Rate Level Loss Ratio}}{\text{1 - Variable Expense Ratio}} + \text{Fixed Expense Ratio} \times \text{Inflation Factor} - 1$$

It is difficult to interpret the inflation factor included in model (b) above. Is this a relative trend factor to measure how fast fixed expenses are growing relative to premium? Are variable expenses loaded on this part of the premium?

Also, it should be noted that using model (a) above (as the author does in Exhibit III-A) implicitly makes the drastic assumption that fixed expense dollars will remain the same as during the experience period.

Perhaps a more appropriate model would be:

c) Indicated Rate Change =

$$\frac{\text{Rate Level Loss Ratio + (Fixed Expense Ratio)t}}{\text{1 - Variable Expense Ratio}} - 1$$

Where t is the rate at which fixed expense dollars will be increasing for the period for which rates are being set.

This formula can be derived as follows. Suppose L' is the rate level loss ratio (i.e. L' = l'/P where l' is the projected loss dollars, and P represents premium at present rates). Furthermore suppose f is the fixed expense dollars needed during the experience period, and f' is the fixed expense dollar that will be needed during the upcoming period (i.e. $f' = tf$). Also define $F' = f'/P$, $F = f/P$, and let V be the variable expense ratio. If $r - 1$ is the indicated rate level change, then:

$$(1-V)rP = PL' + PF'$$

$$= PL' + PFt$$

$$r = \frac{L' + Ft}{1 - V}$$

Again, model (a) above implicitly assumes t=1. The indicated rate change derived by the author in Exhibit III-A using the simple loss ratio method is +16.7%. This implicitly assumes that fixed expenses will also increase at 16.7% annually. His second calculation based on model (a) indicated a +15.0% change. Had he used model (c) with the assumption that fixed expense dollars are increasing at 10% a year the result would have been:

$$\frac{.700 + .065\ (1.10)}{1 - .335} - 1 \doteq .160$$

It is important to note that fixed expenses will be decreasing as a percentage of premium (i.e. after the rate change fixed expenses will constitute $.065(1.1)/1.16 = .062$ of premium) but the loading in the equation should be .065(1.1) rather than just .065.

Investment Income

While the author explicitly mentions in the text of his paper that he is not discussing the propriety of reflecting investment income in the ratemaking process, the reader may be easily misled by the words used in Exhibit X, Sheet 1. It states "the cash flow discount model.... produces an investment income offset to the gross premium to reflect reserves for losses and unearned premium". The numbers derived on line (20) of Exhibit X, Sheet 3, are only one source of input into the general equation to calculate the company's total return. The appropriateness of this return for the risk being assumed by the insurer, must be weighed in choosing the corresponding underwriting profit to be used in calculating rates.

The numbers calculated on line (20) of Exhibit X also seem unnecessarily high until various assumptions are recognized. An expected loss ratio of .86 is used and no commissions are contemplated. The results in line (20) are extremely sensitive to these assumptions. For example suppose commissions are 20% of premium and therefore the expected loss ratio is .66. Line (20) then becomes:

	@ 9%	@ 10%
(20) Present value of income less present value of outgo	18.27%	19.72%

FINAL COMMENTS

Mr. Stergiou's paper has made an important contribution in reinforcing our need to always test the assumptions incorporated in a filing. He has gone even further by sharing with us specific tests he uses to verify the accuracy and reasonableness of loss development factors. We do not believe his intent is to give us a method to use, by rote, to replace the methods we currently use. Rather, his goal is to get us to constantly reappraise our assumptions. This goal is as important as any in the ratemaking process and is well worth the author's efforts.

TITLE: THE PRICING OF MEDI GAP CONTRACTS

AUTHOR: Mr. Emil J. Strug

Mr. Strug is Assistant Vice President and Actuary for Blue Cross Blue Shield of Massachusetts. He received his FCAS designation in 1970. Emil has also authored a paper entitled "Joint Underwriting As A Reinsurance Problem" which appeared in the 1972 Procedings of the Casualty Actuarial Society.

REVIEWER: Mr. Robert F. Bartik

Mr. Bartik is an Assistant Actuary for Kemper Insurance Group. He earned a B.S. degree in Economics from the University of Illinois. Bob received his FCAS in 1972 and is also a member of the American Academy of Actuaries.

To set the stage for the current interest by the regulatory authorities in the pricing of and the benefit content of Medi Gap policies some analysis of the advent of Medicare and its subsequent impact on the economy might be helpful.

The seeds for providing health care to the aged were planted in 1935 in some of the initial versions of the Social Security Act. Under the study provisions of the Act, the Social Security Board was empowered to conduct research and investigations relative to national health insurance. During the intervening years, 1935 to 1965 (passage of Medicare), a series of bills dealing with national health insurance were presented to the Congress: 1939, the Wagner Bill; 1943, the Wagner, Murray, Dingell Bill; 1946, the Taft Bill. In the 1951 to 1964 era, most of the bills dealt with social insurance proposals for persons aged 65 and over. In 1960 the Kerr-Mills Act was passed establishing a program of medical assistance for the aged. Beginning in 1960 efforts to enact a social insurance program of hospital benefits were stepped up with a series of attempts to enact a sound insurance program of hospital benefits known as the King-Anderson Bills. Sufficient momentum was gained so that in 1964 the Senate passed an admendment providing hospital insurance benefits for the aged 65 and over. The House, however, would not agree on a compromise position and the legislation died in conference. In 1965, in addition to a King-Anderson Bill, other proposals were presented such as the Byrnes Bill (named after its author Representative Byrnes), the Eldicare Bill (sponsored by the American Medical Association and introduced by Representative Herlong and Curtis). Early in

1965, under the leadership of Chairman Mills of the House Ways & Means Committee, the Ways & Means Committee put together the Medicare program which was to become effective on July 1, 1967.

The social pressures brought about by the cost to the aged to provide for medical care was a major factor influencing the passage of Medicare. The aged were caught in the bind of fixed incomes and rising cost with medical care costs constantly consuming more of their available income. An examination of the Medicare benefits is in order to assess its impact on the covered individual as well as its impact upon the health care system and the group benefit package for the under 65.

The Medicare program for the 65 and over provides a most comprehensive package of benefits. On the hospital side inpatient room and board for a semiprivate accommodation (and where medically necessary private room) and all special services (general nursing, drugs, operating room, diagnostic services etc) were paid in full for the first 60 days after payment of a deductible. From the 61st to the 90th day the same benefit provisions prevailed but with a daily copayment equal to 25% of the initial deductible. In addition, there was coverage for care provided in a skilled nursing facility (SNF) plus home health services. Full outpatient diagnostic benefits were also provided to minimize use of inpatient usage for such services. Skilled nursing facility benefits were covered in full for the first 20 days, the next 80 days of benefits had a daily copayment equal to 1/8th of the initial inpatient deductible. All of these benefits were provided under the hospital insurance portion of Medicare and commonly referred to as Part A.

Physicians benefits, in addition to home health services, were provided under the Supplementary Medical Insurance (SMI) portion of Medicare generally referred to as Part B. The SMI portion had an annual deductible (as contrasted to a spell of illness deductible under Part A) with the patient and SMI sharing on a 20%-80% (20% patient payment - 80% SMI) basis. Physicians were to be reimbursed on a reasonable charge basis.

With the passage of Medicare the people 65 and over had available to them comprehensive benefits which equalled to and in many cases was greater than that held by the under 65 population. By removal of the financial constraints due to inadequate or no insurance and a backlog of medical need, the medicare population made full use of the program. Its impact upon the medical care system for the entire population has been well documented by health economist and is reflected in:

Table 1. Portion of Health Care Costs Paid By Individuals versus Third Party Payors

Table 2. Health Care Expenditures As % Of Gross National Product.

Table 3. Ratio of Personal Expenditures For Medical Care To Personal Income

Table 4. Annual Changes In Consumer Price Index and In Medical Components of the Index

The following tables present those variables from 1966 to 1977:

Table 1.—Portion of Health Care Costs Paid By Individuals versus Third Party Payors

	UNDER 65			65 AND OVER		
Fiscal Years Ending June 30	Total	Out of Pocket	Third Party	Total	Out of Pocket	Third Party
1966	100%	51%	49%	100%	53%	47%
1967	100%	48%	52%	100%	37%	63%
1970	100%	43%	57%	100%	33%	67%
1973	100%	38%	62%	100%	33%	67%
1976	100%	35%	65%	100%	27%	73%
1977 (Sept)	100%	32%	68%	100%	27%	73%

Table 2.—Health Care Expenditures As % of Gross National Product

Fiscal Years Ending	
1966	5.8%
1967	6.2%
1978	7.2%
1973	7.7%
1976	8.7%
1977 (Sept)	8.8%

Table 3.—Ratio of Personal Expenditures For Medical Care To Disposable Personal Income

Calendar Year	
1966	6.2%
1967	6.3%
1970	7.1%
1973	7.4%
1976	8.6%
1977	9.1%

Table 4.—Annual Changes In Consumer Price Index and In Medical Components of the Index

Calendar Year	All Items	All Medical Care Items	Physician Fees	Hospital Room	Prescriptions & Drugs
1966	2.9%	4.4%	5.8%	100%	1.3%
1967	2.9%	7.1%	7.1%	19.8%	- 0.5%
1970	5.9%	6.3%	7.5%	12.9%	2.3%
1973	6.2%	3.9%	3.3%	4.7%	0.3%
1976	5.8%	9.5%	11.3%	13.8%	6.1%
1977	6.5%	9.6%	9.3%	11.5%	6.4%

The results speak for themselves as to the rapid rise in medical care costs. Considering the limited and relatively fixed income for the 65 and over population one can see how the social pressures to provide relief in the form of medical care arose in the early 60's and have been aggravated in the latter half of the 70's.

A history of the movement of the medicare deductibles and the cost to purchase Part B (medical) benefits will also show how the increase in these elements have further impacted the standard of living of the aged.

MEDICARE DEDUCTIBLES, COPAYS & COINSURANCE AND PREMIUM

	PART A				PART B		
	Benefit Period Deductible	Daily CoPay 61st to 90th Hospital Days	21st to 100th SNF Days		Premium	Annual Deductible	Coin-surance
7/66	$40	$10	$5.00	7/66	$3.00	$50	20%
1/69	$44	$11	$5.50	4/68	$4.00		
1/70	$52	$13	$6.50	7/70	$5.30		
1/71	$60	$15	$7.50	7/71	$5.60		
1/72	$68	$17	$8.50	7/73	$5.80	$60	20%
1/73	$72	$18	$9.00	7/74	$6.30		
1/74	$84	$21	$10.50				
1/75	$92	$23	$11.00				
1/76	$104	$26	$13.00	7/76	$7.20		
1/77	$124	$31	$15.50	7/77	$7.70		
1/78	$144	$36	$18.00	7/78	$8.20		
1/79	$160	$40	$20.00	7/79	$8.70		

It should be noted that in 1972 the Medicare benefits were extended to the disabled under social security and those receiving treatment for chronic kidney disease. As was mentioned earlier in this treatise, deductibles were introduced to keep down the cost of the program to the government. The initial hospital deductible was set equal to the daily cost of care in a semiprivate room. The Part B deductible was set at $50 per calendar year with 20% paid by the recipient for the remaining balance with the first period being only 6 months to minimize the cost of the program to the government.

To meet the needs of the 65 and over population as to insuring the uncovered portions of the Medicare program, policies were designed which tended to duplicate in conjunction with Medicare comprehensive programs offered by the industry.

The major elements of cost to be met were:

1. The initial inpatient deductible for each spell of illness.
2. The copay days from the 61st to the 90th day.
3. Full coverage from the 91st day on.
4. The copay days in a skilled nursing facility from the 21st to the 100th day.
5. The deductible (currently $60) and coinsurance (20%) for services provided by physicians and surgeons which were routinely provided under a typical health insurance policy.
6. Prescription drugs not provided by the hospital.

More than a decade has past since the program began and along with it the availability of data particular to the insured medicare population. Data pertaining to the complimentary Part A deductible and copays is relatively clean as the benefits are for a spell of illness or benefit period. On the other hand the Part B presents some problems due to the status of the deductible being maintained by Social Security and the difficulty if not the inability to maintain appropriate service counts and distribution of losses by size that limits the ability to measure the impact of inflation upon the deductible and the truncation of service counts under the deductible. An additional ramification relative to inflation is the coverage provided under Part B for certain services provided in the outpatient area of a hospital similar to those provided on an inpatient basis.

For analytical purposes I've taken the pure premium calculation underlying the rate calculation for policies issued April 1 thru June 30, 1979, for a duration of 12 months.

The largest element, in terms of cost, is the inpatient hospital deductible. The estimation of the utilization for this benefit is relatively simple. A regression analysis is performed using 13 data points. These points represent 12 months of incurred or accident year data for successive quarters. The actual results and extrapolations considered and used are contained in Exhibit 1.

The estimation of the deductible amount, unfortunately, is not quite as simple since it is based upon nationwide data for a period of time which is incomplete as far as development. In the case of this example the 1980 deductible will be developed from 1978 data and as shown in Exhibit 2 (the rates were calculated during the last quarter of 1979).

The method employed to estimate the deductible is dependent upon two sources of data. The first is the calculation of the hospital deductible for the period prior to the year to be estimated (in this case 1979) as published in the Federal Register (see Appendix A) and the reports issued by OASDI displaying the experience for hospital insurance for various time periods which correspond to those used to calculate the hospital deductible (Appendix B). It should be noted that the deductible value serves as a basis to establish the daily copay payments (by the insured or insurer) for the 61st to the 90th day and in a skilled nursing facility from the 21st to the 100th day.

The use of such external indices as the hospital services component of the Consumer Price Index are not reliable as it does not reflect the mix of services used by the 65 and over population.

If one compared the change in the inpatient hospital deductible with the change in the hospital charge indices contained in the CPI, they would find no consistency even when the CPI change is adjusted to reflect its impact one year hence on the Part A deductible.

The pure premium calculation for the in hospital copay for the 61st to the 90th day and skilled nursing copay for the 21st to the 100th day present no unusual consideration except for the calculation of the copay value. The method of determining the liability, as previously mentioned, follows that of the inpatient deductible. Consideration must be given to any variance between results of insured programs and those of the total medicare population and the availability of institutions which provide certain levels of care. Medicare studies (Appendix C) indicate days of care in a short hospital stay decreasing as well as a decrease in the number of skilled nursing facilities. These factors were considered in the choice of anticipated utilization levels for in hospital copay days from the 61st to the 90th day and skilled nursing copay days from the 21st to the 100th (Exhibits 3 & 4 respectively).

Exhibit 3 shows the day utilization for cases with length of stays from the 61st to the 90th day increasing. Without a distribution of cases by length of stay from 1 to 90 or more days, it is difficult to compare the total day utilization (1 to 90 days) of the insured population to that of the medicare population. One can rationalize as being reasonable that the number of claims with length of stays of less than 61 days could be decreasing while cases with length of stays beyond 60 days could have either had an increase in volume or length of stay. In addition, an insured program might be more attractive to those who need or anticipate the need of medical care, thereby inducing higher utilization.

Exhibit 5 presents the development of pure premiums for in hospital benefits beyond the 90th day. Benefits for days beyond 90 are paid for in full by the insured carrier. Normally one would expect that this value would be determined by estimating the day utilization and the average daily costs. An analysis of these elements indicated erratic behavior in terms of utilization, length of stay and costs whereas the composite (i.e., pure premiums) produce stable as well as reasonable results.

The most difficult element of pure premium to calculate is that to cover the Part B annual deductible in whole or in part for physician and outpatient hospital services. As was previously mentioned, there are no available statistics by size of losses to determine the impact of inflation and utilization upon the deductible value as the status of the Part B deductible and the benefits applicable to satisfy the deductible are maintained by Medicare.

To obviate the problem, the choice of the regression curve was paramount. It had to not only show a high degree of correlation to historical data but also demonstrate a pattern of future development that was logical. With increasing unit cost one would expect that in successive years the average deductible would increase at a decreasing rate and become asymptotic as it approached the deductible limit.

The most recent observation would indicate that the values have become asymptotic, therefore, the last observed value was chosen as the expected pure premium for the rating period. The historic values and the projected pure premiums are shown in Exhibit 6.

For the coinsurance benefits which compliment the Part B 80% co-insurance payments, a return to the more traditional technique of using utilization (frequency) and average cost per service for calculating pure premiums was adopted. The physicians and hospital elements are separated as each is influenced differently by the inflation factors particular to each of the segments. The increase in physicians prevailing fees is controlled by the Department of Health, Education & Welfare. For 1979 this value was calculated to be 5.08% over 1978 values and this same value was assumed to continue in 1980. The increase in hospital charges would reflect the inflationary pressures of the local hospital area and are currently being controlled by competition amongst hospitals and the American Hospital Association voluntary effort. Appendices D1, D2, and D3 detail the allowable increases in physicians prevailing charges carried into the pure premium calculations.

Exhibit 7 develops the expected service utilization for physicians coinsurance benefits. The most recent observations indicate a leveling if not a moderation of usage. For projection purposes the last observed value was used. The average service cost is developed in Exhibit 8 and used the previously mentioned 5.08% and prior values as taken from the letters issued by Health, Education and Welfare to Part B intermediaries. The companion piece to the physicians' coinsurance is the outpatient hospital coinsurance benefit. The utilization and cost considerations are displayed in Exhibits 9 and 10.

To corroborate these trends and values (Physicians & Hospital) a review is made of the assumptions used by Health, Education and Welfare in developing Part B monthly actuarial and monthly premium rates. These calculations are contained in the Federal Register and are normally published each December.

Appendix E1 presents the various SSA assumptions underlying various SSA pricing and funding calculations. Table 5 presents a range of values for the projection factors for physicians' fees, utilization of physicians' services and outpatient hospital services. The projection factor used in the pure premium calculation to cover the uninsured coinsurance portion for these benefits are below those indicated by SSA for physicians' fees and utilization and within the high and low assumptions for outpatient hospital services.

The next and last benefit to be analyzed is prescription drugs. Prescription drugs, outside of those provided in a hospital setting, are not covered by Medicare. The benefit to be priced provides prescription drugs subject to a $25 quarterly deductible and 25% coinsurance. Pure premiums are developed by estimating the number of claims and the average number of prescriptions per claim and the average cost per prescription. The estimation of the number of claims and the average number of prescriptions present no unusual or unique considerations. Generally the number of claimants have increased over time with the number of prescriptions showing a continuing decline. The underlying data and projections for these two elements are shown in Exhibits 11 & 12. In order to develop the full prescription charge, the average prescription claim payment has to be adjusted to reflect the removal of the 25% coinsurance and the $25 deductible. Projecting the average prescription charge without modification would obviously produce erroneous and undefensible results. The conversion of the average prescription cost from a partial to a full basis is developed in Exhibit 13. The resultant values are then transferred to Exhibit 14 where the projected value is developed. To evaluate the reasonableness of this value the inherent annual trend from the last observed value to the projected value is compared to the trends observed for the most recent annual values in the Consumer Price Index and for those shown in the Lilly Digest. At the time of preparation of the filing, the Consumer Price Index trend, as of October 1978, was 7.5%, while the Lilly Digest (1977) showed 9.4%. The 5.5% trend in the pure premium projections used was therefore considered to be reasonable. The estimated pure premiums for the benefit was calculated by developing the estimated full charge per claim and then reducing this value by the deductible amount and 20% coinsurance. Exhibit 15, Item H, details the calculation.

The pure premium for each of the benefit categories previously described and their detailed calculations are contained in Exhibit 15.

ANNUAL CLAIMS INCIDENCE

PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL 3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	PROJECTED 6/14/80
Inpatient Hospital Deductible	24.932	24.966	25.001	25.026	25.346	25.513	25.750	25.910	25.771	25.917	25.889	26.215	26.968*

* The projected values resulting from the three projection methods indicated below were initially considered. Despite the significantly high indexes of determination and the reasonability of the values, it was determined to be appropriate to calculate the projected claim incidence value using the most recently observed annual rate of increase (1.2%) which is somewhat lower than the annual trends underlying the aforementioned projected values. [$(26.215)(1.012)^{28.5/12} = 26.968$].

Projection Method	Form of Equation	Index of Determination	Projected Value
Linear	Y = A + BX	.928	27.329
Hyperbolic	Y = 1/(A + BX)	.927	27.462
Exponential	Y = A(EXP(BX))	.926	27.392

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.928	24.735696	.120598	27.329	1.8%
2. Y=1/(A+B*X)	HYPERBOLIC	5	.927	.040403	-.000186	27.462	2.0%
3. Y=A*EXP(B*X)	EXPONENTIAL	2	.926	24.743391	.004730	27.392	1.9%
4. Y=A*(X^B)	POWER	3	.827	24.625612	.021323	26.290	-.1%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.826	24.615411	.542850	26.201	.1%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.540	.001940	.038695	25.703	-.7%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.537	25.843334	-.049305	25.784	-.7%
8. Y=A+(B/X)	HYPERBOLIC	4	.534	25.843736	-1.253488	25.785	-.7%

EXHIBIT 1

EXHIBIT 2

Estimate of 1980 Medicare Inpatient Hospital Deductible*

ITEM	AMOUNT	SOURCE
A. Average hospital charge per day for the period January 1, 1977 to December 31, 1977	$197.07	Appendix B
B. Average per diem rate for the period January 1, 1977 to December 31, 1977	$160.69	Page 44891, Federal Register, Vol. 43 No. 190, dated 9/29/78. Appendix A
C. Ratio of per diem rate to average hospital charge per day for the period January 1, 1977 to December 31, 1977	.815	Item B ÷ Item A
D. Average hospital charge per day for the period January 1, 1977 to June 30, 1977	$190.77	Appendix B
E. Average hospital charges per day for the period January 1, 1978 to June 30, 1978	$217.21	Appendix B
F. Estimated average hospital charge per day for the period January 1, 1978 to December 31, 1978	$224.38	(Item E ÷ Item D)(Item A)
G. Estimated ratio of per diem rate to average hospital charge per day for the period January 1, 1978 to December 31, 1978	.815	Based on 1977 experience. Item C.
H. Estimated average per diem rate for the period January 1, 1978 to December 31, 1978	$183.68	(Item F)(Item G)
I. Average per diem rate for the period January 1, 1966 to December 31, 1966	$ 40.01	Page 44891, Federal Register, Vol. 43 No. 190, dated 9/29/78. Attachment I
J. Estimated 1980 inpatient hospital deductible	$184.00	(Item H - Item I)($40) rounded to the nearest multiple of $4.00

*The law provides that for spells of illness beginning in calendar years after 1968 the inpatient hospital deductible shall be equal to $40 multiplied by the ratio of (1) the current average per diem rate for inpatient hospital services for the calendar year preceding the year in which the promulgation is made to (2) the current average per diem rate for such services for 1966. Changes in the amount of the inpatient hospital deductible also affect certain other cost-sharing provisions under the Medicare hospital insurance program, the patient co-payment for the 61st to 90th inpatient day which equals 25 percent of the inpatient hospital deductible, and the skilled nursing home daily co-payment which is equal to 12.5 percent of the inpatient hospital deductible.

EXHIBIT 2A

Actual and Estimated Deductible and
Coinsurance Amounts for Medicare Beneficiaries

Item	ACTUAL 1974	1975	1976	1977	1978	1979	ESTIMATED 1980
A. Hospital Inpatient Deductible Per Admission	$84.00	$92.00	$104.00	$124.00	$144.00	$160.00	$184.00
B. Patient Co-Payment from the 61st to the 90th Inpatient Day (25% of Item (A) values)	$21.00	$23.00	$ 26.00	$ 31.00	$ 36.00	$ 40.00	$ 46.00
C. Skilled Nursing Facility Daily Co-Payment (12.5% of Item (A) values)	$10.50	$11.50	$ 13.00	$ 15.50	$ 18.00	$ 20.00	$ 23.00
D. Physicians' Services and Outpatient Services Annual Deductible	$60.00	$60.00	$ 60.00	$ 60.00	$ 60.00	$ 60.00	$ 60.00
E. Patient Coinsurance for Physicians' Services and Outpatient Services	20%	20%	20%	20%	20%	20%	20%

Calculation of the Liability for the
Period of these Rates for the Benefit Categories
Impacted by the Expected Increase in the
1980 Medicare Inpatient Hospital Deductible

ITEM	AMOUNT	SOURCE
A. 1979 Medicare inpatient hospital deductible	$160.00	Page 44891, Federal Register, Vol. 43 No. 190, dated 9/29/78. Attachment I
B. Estimated 1980 Medicare inpatient hospital deductible	$184.00	Exhibit 2A, Item K
C. Medicare inpatient hospital deductible for the period 5/15/79 to 5/14/80	$169.00	[(7.5/12)(Item A) + (4.5/12)(Item B)]
D. Co-payment for the 61st to the 90th inpatient hospital day for the period 5/15/79 to 5/14/80	$ 42.25	(Item C)(.250)
E. Skilled nursing facility daily co-payment for the period 5/15/79 to 5/14/80	$ 21.13	(Item C)(.125)

ANNUAL DAY INCIDENCE
PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL 3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	PROJECTED 5/14/80
Co-payment for the 61st to the 90th Inpatient Hospital Day	15.732	16.504	16.633	17.384	17.995	18.137	18.420	18.407	18.453	18.484	18.443	18.644	19.225*

* The projected values resulting from the two projection methods indicated below were initially considered. Despite the significantly high indexes of determination and the reasonability of the values, it was determined to be appropriate to calculate the projected day incidence value using the most recently observed annual rate of increase (1.3%) which is somewhat lower than the annual trends underlying the aforementioned projected values [$(18.644)(1.013)^{28.5/12} = 19.225$].

Projection Method	Form of Equation	Index of Determination	Projected Value
Logarithmic	$Y = A + B(\ln X)$	.951	19.529
Power	$Y = AX^B$	.951	19.648

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+B*LOG(X)	LOGARITHMIC	7	.951	15.679929	1.254644	19.529	2.0%
2. Y=A*(X^B)	POWER	3	.951	15.721701	.072661	19.648	2.2%
3. Y=X/(A+B*X)	HYPERBOLIC	6	.867	.011565	.053449	18.523	-.3%
4. Y=A*EXP(B/X)	EXPONENTIAL	8	.855	18.675771	-.197837	18.505	-.3%
5. Y=A+(B/X)	HYPERBOLIC	4	.843	18.646558	-3.390909	18.489	-.4%
6. Y=A+(B*X)	LINEAR	1	.807	16.194574	.242322	21.404	6.0%
7. Y=A*EXP(B*X)	EXPONENTIAL	2	.797	16.206697	.013945	21.873	7.0%
8. Y=1/(A+B*X)	HYPERBOLIC	5	.786	.061665	-.000804	22.532	8.3%

EXHIBIT 3

ANNUAL DAY INCIDENCE

PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL 3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	PROJECTED 6/14/80
Skilled Nursing Facility Copayment from 21st to 100th day	38.222	37.110	36.874	36.101	34.642	34.094	32.028	29.945	27.113	23.493	20.663	18.111	11.257*

* The projected value is the result of an exponential projection [Y = A(EXP(BX))] which has an index of determination of .879. This value is considered to be appropriate for inclusion in the rate calculation in view of the acceptable index of determination as well as the fact that the annual trend underlying the projected value is consistent with the expectation that day incidence for Skilled Nursing Facilities will continue to decrease, but at a somewhat lesser rate than has been historically observed. A linear projection [Y = A + BX] has a higher index of determination (i.e., .926), however the resulting projected value of 3.161 was considered to be clearly inadequate and therefore rejected. Tho remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.926	42.621060	-1.835343	3.161	-52.0%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.879	45.623749	-.065091	11.257	-18.1%
3. Y=1/(A+B*X)	HYPERBOLIC	5	.821	.018952	.002392	14.207	-9.7%
4. Y=A+B*LOG(X)	LOGARITHMIC	7	.697	43.344607	-7.596821	20.037	4.3%
5. Y=A*(X^B)	POWER	3	.631	46.313503	-.263028	20.665	5.7%
6. Y=A+(B/X)	HYPERBOLIC	4	.395	26.444535	16.422209	27.208	18.7%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.340	25.892833	.554400	26.569	17.5%
8. Y=X/(A+B*X)	HYPERBOLIC	6	.288	-.019405	.039521	25.894	16.2%

EXHIBIT 4

MONTHLY PURE PREMIUM

PER CONTRACT FOR FISCAL YEARS ENDING

Benefit Category	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	-PROJECTED- 5/14/80
	ACTUAL												
Payment from the 91st Inpatient Hospital Day On.	$1.210	$1.324	$1.433	$1.525	$1.606	$1.601	$1.631	$1.643	$1.653	$1.643	$1.661	$1.650	$1.678*

* The projected value is the result of a hyperbolic projection [Y = X/(A + BX)] which has an index of determination of .944, the highest index of determination of the projection methods employed. A logarithmic projection [Y = A + B(ln X)] has virtually the same index of determination (i.e., .943), however the resulting projected value of $1.816 was considered to be excessive in view of the relative stability of the recent actual experience. The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=X/(A+B*X)	HYPERBOLIC	6	.944	.265492	.583515	1.678	.7%
2. Y=A+B*LOG(X)	LOGARITHMIC	7	.943	1.229987	.191130	1.816	4.1%
3. Y=A*(X^B)	POWER	3	.936	1.237339	.131828	1.854	5.0%
4. Y=A*EXP(B/X)	EXPONENTIAL	8	.928	1.698744	-.376911	1.669	.5%
5. Y=A+(B/X)	HYPERBOLIC	4	.910	1.687731	-.539045	1.663	.3%
6. Y=A+(B*X)	LINEAR	1	.736	1.318242	.035399	2.079	10.2%
7. Y=A*EXP(B*X)	EXPONENTIAL	2	.713	1.317406	.024125	2.213	13.2%
8. Y=1/(A+B*X)	HYPERBOLIC	5	.609	.759790	-.016557	2.476	18.6%

EXHIBIT 5

MONTHLY PURE PREMIUM
PER CONTRACT FOR FISCAL YEARS ENDING

	ACTUAL												PROJECTED
Benefit Category	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	5/14/80
Physicians' Services and Outpatient Services Annual Deductible	$1.822	$1.851	$1.852	$1.837	$1.975	$2.065	$2.109	$2.134	$2.238	$2.235	$2.242	$2.234	$2.234*

* The most recent observation (i.e., the year ending 12/31/77) has been carried forward to the period of the rates. The three projection methods indicated below have significantly high indexes of determination; however due to the relative stability of the four most recent observations, the projected values were judged to be excessive and therefore rejected.

Projection Method	Form of Equation	Index of Determination	Projected Value
Linear	Y = A + BX	.923	$2.745
Exponential	Y = A(EXP(BX))	.919	$2.876
Hyperbolic	Y = 1/(A + BX)	.914	$3.102

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.923	1.748182	.046357	2.745	9.1%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.919	1.761175	.022810	2.876	11.2%
3. Y=1/(A+B*X)	HYPERBOLIC	5	.914	.564407	-.011259	3.102	14.8%
4. Y=A*(X^B)	POWER	3	.828	1.719763	.103302	2.361	2.4%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.822	1.701871	.208711	2.342	2.0%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.550	.119619	.460290	2.147	-1.7%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.540	2.173125	-.239441	2.149	-1.6%
8. Y=A+(B/X)	HYPERBOLIC	4	.530	2.173832	-.480788	2.151	-1.6%

EXHIBIT 6

ANNUAL SERVICE INCIDENCE

PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	-PROJECTED- 5/14/80
	ACTUAL												
Physicians' Services Coinsurance	349.034	361.880	379.235	397.626	405.828	419.269	434.288	447.282	448.633	451.196	445.098	444.293	444.293*

* The most recent observation (i.e., the year ending 12/31/77) has been carried forward to the period of the rates. The two projection methods indicated below have significantly high indexes of determination; however, due to the relative stability of the five most recent observations, the projected values, which represent upward trends, were judged to be inappropriate and therefore rejected.

Projection Method	Form of Equation	Index of Determination	Projected Value
Power	$Y = AX^B$	.956	487.301
Logarithmic	$Y = A + B(\ln X)$	.947	480.797

- 623 -

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A*(X^B)	POWER	3	.955	340.755047	.116595	487.301	4.0%
2. Y=A+B*LOG(X)	LOGARITHMIC	7	.947	337.525093	46.697891	480.797	3.4%
3. Y=A+(B*X)	LINEAR	1	.877	354.056190	9.422917	556.649	10.0%
4. Y=A*EXP(B*X)	EXPONENTIAL	2	.868	355.636978	.023301	586.911	12.4%
5. Y=1/(A+B*X)	HYPERBOLIC	5	.856	.002802	-.000058	641.861	16.8%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.806	.000769	.002227	441.929	-.2%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.783	447.533321	-.303118	441.268	-.3%
8. Y=A+(B/X)	HYPERBOLIC	4	.759	446.350307	-120.050450	440.76[illegible]	-.3%

EXHIBIT 7

EXHIBIT 8

Calculation of the Average Cost Per Service
for the Period of these Rates for the Physicians'
Coinsurance Benefit Category

ITEM	AMOUNT	SOURCE
A. Calculation of the cost trend factor to project the average cost per service for physicians' coinsurance benefit category from the year ending 12/31/77 to the year ending 5/14/80.		
1. The economic index applicable to physicians' services announced by the Social Security Administration for the period July 1, 1976 through June 30, 1977.	1.276	Part B Intermediary Letter No. 76-34 from Department of Health, Education and Welfare, dated August 1976. Appendix D1
2. The economic index applicable to physicians' services announced by the Social Security Administration for the period July 1, 1977 through June 30, 1978.	1.357	Part B Intermediary Letter No. 77-24 from Department of Health, Education and Welfare, dated June 1977. Appendix D2
3. Percent of increase for fiscal year 1978 over fiscal year 1977	6.35%	Item A.2. ÷ Item A.1.
4. The economic index applicable to physicians' services announced by the Social Security Administration for the period July 1, 1978 through June 30, 1979	1.426	Part B Intermediary Letter No. 78-23 from Department of Health, Education and Welfare, dated June 1978. Appendix D3
5. Percent of increase for fiscal year 1979 over fiscal year 1978	5.08%	Item A.4. ÷ Item A.2.
6. Expected percent of increase for fiscal year 1980 over fiscal year 1979	5.08%	Judgement
B. Cost trend factor to project the year ending 12/31/77 to the year ending 5/14/80.	1.132	$(1.0635)^{6/12}(1.0508)(1.0508)^{10.5/12}$
C. Cost per service for the physicians' coinsurance benefit category for the year ending 12/31/77.	$7.85	Medicare Complimentary Rate Study Tabulation
D. Expected average cost per service for physicians' coinsurance benefit category for the year ending 5/14/80	$8.89	(Item B) (Item C)

ANNUAL SERVICE INCIDENCE
PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL												-PROJECTED-
	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	5/14/80
Outpatient Hospital Services Coinsurance	70.307	74.164	78.924	83.151	85.813	90.761	95.921	99.602	102.056	105.553	108.745	113.426	150.742*

* The projected value is the result of a linear projection [Y = A + BX] which has an index of determination of .996, the highest index of determination of the projection methods employed. This value is considered to be appropriate for inclusion in the rate calculation in view of the extremely high index of determination as well as the fact that the annual trend underlying the projected value is consistent with the decelerating annual rates of increase observed in the recent historical experience. An exponential projection [Y = A(EXP(BX))] and a hyperbolic projection [Y = 1/(A + BX)] also have extremely high indexes of determination (i.e., .987 and .970, respectively), however the resulting projected values (i.e., 173.859 and 257.653, respectively) were considered to be excessive and therefore rejected. The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.996	67.072222	3.891619	150.742	12.7%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.987	69.133700	.042893	173.859	19.7%
3. Y=1/(A+B*X)	HYPERBOLIC	5	.970	.014183	-.000479	257.553	41.2%
4. Y=A*(X^B)	POWER	3	.942	65.492350	.199877	120.924	2.7%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.912	62.790400	17.757761	117.272	1.4%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.739	.005724	.009589	101.474	-4.6%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.688	103.716866	-.490394	101.378	-4.6%
8. Y=A+(B/X)	HYPERBOLIC	4	.636	103.382814	-42.594857	101.402	-4.6%

EXHIBIT 9

AVERAGE COST PER SERVICE

FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL 3/31/76	6/30/75	9/30/75	12/31/76	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	-PROJECTED- 6/14/80
Outpatient Hospital Services Coinsurance	$6.66	$6.98	$7.23	$7.36	$7.56	$7.72	$7.92	$8.16	$8.32	$8.47	$8.68	$8.89	$11.03*

* The three projection methods indicated below result in extremely high and nearly equal indexes of determination. The projected value produced by the hyperbolic projection was rejected as being clearly excessive. It was determined to be appropriate to use the mean of the linear projection and the exponential projection [($10.73 + $11.33) ÷ 2 = $11.03] in the rate calculation in consideration of the nearly equal validity of the linear and exponential projection methods, as well as the fact that the annual trend underlying the mean value is consistent with both recent historical experience and reasonable expectations of future hospital cost increases for outpatient services.

Projection Method	Form of Equation	Index of Determination	Projected Value
Linear	Y = A + BX	.996	$10.73
Exponential	Y = A(EXP(BX))	.991	$11.33
Hyperbolic	Y = 1/(A + BX)	.983	$12.43

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.996	6.572121	.193392	10.730	8.2%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.991	6.635739	.024875	11.328	10.7%
3. Y=1/(A+B*X)	HYPERBOLIC	5	.983	.149577	-.003215	12.429	15.2%
4. Y=A*(X^B)	POWER	3	.940	6.434750	.115542	9.172	1.3%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.919	6.353128	.886190	9.072	.9%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.730	.037955	.118866	8.290	-2.9%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.695	8.397684	-.285364	8.287	-2.9%
8. Y=A+(B/X)	HYPERBOLIC	4	.660	8.386769	-2.156228	8.286	-2.9%

EXHIBIT 15

ANNUAL CLAIM INCIDENCE

PER 100 CONTRACTS FOR FISCAL YEARS ENDING

Benefit Category	ACTUAL 3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	PROJECTED 5/14/80
Prescription Drugs	45.596	46.638	47.320	48.467	49.514	51.017	53.018	54.695	56.173	57.436	58.618	59.663	72.772*

* The projected value is the result of a linear projection [Y = A + BX] which has an index of determination of .991. This value is considered to be appropriate for inclusion in the rate calculation in view of the extremely high index of determination as well as the fact that the annual trend underlying the projected value is consistent with the decelerating annual rates of increase observed in the recent historical experience. An exponential projection [Y = A(EXP(BX))] and a hyperbolic projection [Y = 1/(A + BX)] have slightly higher indexes of determination (i.e., .992), however the resulting projected values (i.e., 77.042 and 85.039, respectively) were considered to be excessive and therefore rejected. The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A*EXP(B*X)	EXPONENTIAL	2	.992	44.016928	.026036	77.042	11.4%
2. Y=1/(A+B*X)	HYPERBOLIC	5	.992	.022510	-.000500	85.039	16.1%
3. Y=A+(B*X)	LINEAR	1	.991	43.495270	1.361689	72.772	8.7%
4. Y=A*(X^B)	POWER	3	.866	42.970726	.116049	61.340	1.2%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.847	42.342538	6.006066	60.769	.8%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.595	.005300	.017889	55.141	-3.3%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.569	55.903593	-.269983	55.206	-3.2%
8. Y=A+(B/X)	HYPERBOLIC	4	.544	55.918910	-13.815348	55.276	-3.2%

EXHIBIT 11

AVERAGE NUMBER OF PRESCRIPTIONS

PER PRESCRIPTION DRUG CLAIM FOR FISCAL YEARS ENDING

	ACTUAL												-PROJECTED-
Benefit Category	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	5/14/80
Prescription Drugs	9.875	9.760	9.642	9.402	9.277	9.149	9.081	9.011	8.925	8.866	8.788	8.712	8.054*

* The three projection methods indicated below result in extremely high and nearly equal indexes of determination. It was determined to be appropriate to use a mean of the linear, exponential, and hyperbolic projections [(7.877 + 7.778 + 8.508) ÷ 3 = 8.054] in the rate calculation in consideration of nearly equal validity of these three projection methods as well as the fact that the annual trend underlying the mean value is equal to the most recently observed annual rate of decrease (-3.3%).

Projection Method	Form of Equation	Index of Determination	Projected Value
Hyperbolic	Y = 1/(A + BX)	.976	7.877
Exponential	Y = A(EXP(BX))	.971	7.778
Logarithmic	Y = A + B(ln X)	.970	8.508

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=1/(A+B*X)	HYPERBOLIC	5	.976	.101054	.001204	7.877	-4.2%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.971	9.880607	-.011128	7.778	-4.7%
3. Y=A+B*LOG(X)	LOGARITHMIC	7	.970	10.017760	-.492071	8.508	-1.0%
4. Y=A+(B*X)	LINEAR	1	.965	9.867120	-.102916	7.654	-5.3%
5. Y=A*(X^B)	POWER	3	.964	10.037782	-.052904	8.534	-.9%
6. Y=A+(B/X)	HYPERBOLIC	4	.745	8.877872	1.238568	8.935	1.1%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.733	8.881842	.132370	8.937	1.1%
8. Y=X/(A+B*X)	HYPERBOLIC	6	.720	-.014160	.112545	8.938	1.1%

EXHIBIT 12

CALCULATION OF THE AVERAGE CHARGE PER PRESCRIPTION DRUG CLAIM
FOR FISCAL YEARS ENDING

	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	SOURCE
1. Average cost per claim	$32.00	$32.47	$32.71	$32.97	$32.93	$32.86	$32.86	$33.16	$33.35	$33.68	$33.92	$34.37	
2. Average charge per claim	$65.00	$65.59	$65.89	$66.21	$66.16	$66.08	$66.08	$66.45	$66.69	$67.10	$67.40	$67.96	Drug benefits provide for 80% coinsurance after satisfaction of a $25.00 deductible (Item 1 ÷ .80) + $25.00.
3. Average number of prescription per claim	9.875	9.750	9.542	9.402	9.277	9.149	9.081	9.011	8.925	8.866	8.788	8.712	
4. Average charge per prescription	$6.58	$6.73	$6.91	$7.04	$7.13	$7.22	$7.28	$7.37	$7.47	$7.57	$7.67	$7.80	Item 2 ÷ Item 3

EXHIBIT 13

AVERAGE CHARGE PER PERSCRIPTION
FOR FISCAL YEARS ENDING

Benefit Category	3/31/75	6/30/75	9/30/75	12/31/75	3/31/76	6/30/76	9/30/76	12/31/76	3/31/77	6/30/77	9/30/77	12/31/77	-PROJECTED- 5/14/80
	ACTUAL												
Prescription Drugs	$6.58	$6.73	$6.91	$7.04	$7.13	$7.22	$7.28	$7.37	$7.47	$7.57	$7.67	$7.80	$8.86*

* The three projection methods indicated below have extremely high and nearly equal indexes of determination. The value produced by the hyperbolic projection was rejected as being excessive in view of the historical rates of increase. It was determined to be appropriate to use the mean of the linear projection and the exponential projection [($8.77 + $8.95) ÷ 2 = $8.86] in the rate calculation in consideration of the nearly equal validity of the linear and exponential projection methods, as well as the fact that the annual trend underlying the mean value is consistent with recent historical experience.

Projection Method	Form of Equation	Index of Determination	Projected Value
Linear	Y = A + BX	.987	$8.77
Exponential	Y = A(EXP(BX))	.982	$8.95
Hyperbolic	Y = 1/(A + BX)	.976	$9.19

The remaining projection methods employed produce values and/or indexes of determination that were judged to be inappropriate for consideration.

FORM OF EQUATION	TYPE OF FUNCTION	EQUATION NUMBER	INDEX OF DETERMINATION	A	B	PROJ. VALUE	ANN. TREND
1. Y=A+(B*X)	LINEAR	1	.987	6.561969	.102902	8.774	5.1%
2. Y=A*EXP(B*X)	EXPONENTIAL	2	.982	6.580818	.014302	8.950	6.0%
3. Y=1/(A+B*X)	HYPERBOLIC	5	.976	.151582	-.001991	9.193	7.2%
4. Y=A*(X^B)	POWER	3	.952	6.457253	.067192	7.936	.7%
5. Y=A+B*LOG(X)	LOGARITHMIC	7	.941	6.432230	.479468	7.903	.6%
6. Y=X/(A+B*X)	HYPERBOLIC	6	.742	.023777	.132492	7.485	-1.7%
7. Y=A*EXP(B/X)	EXPONENTIAL	8	.722	7.542393	-.167905	7.484	-1.7%
8. Y=A+(B/X)	HYPERBOLIC	4	.701	7.537976	-1.187710	7.483	-1.7%

EXHIBIT 14

Calculation of the Expected Monthly
Pure Premium Increments
for the Period 5/15/79 to 5/14/80

ITEM	AMOUNT	SOURCE
A. Inpatient hospital deductible per admission		
1. Annual claim incidence per 100 contracts	26.968	Exhibit 1
2. Average payment per inpatient hospital deductible	$169.00	Exhibit 2B, Item C
3. Expected monthly pure premium	$ 3.798	[(Item A1)(Item A2) ÷ 1200]
B. Co-payment for the 61st to the 90th inpatient hospital day		
1. Annual day incidence per 100 contracts	19.225	Exhibit 3
2. Average payment per day	$ 42.25	Schedule 2B, Item D
3. Expected monthly pure premium	$ .677	[(Item B1)(Item B2) ÷ 1200]
C. Expected monthly pure premium for the 91st to the 120th inpatient hospital day	$ 1.678	Exhibit 5
D. Expected monthly pure premium for the joint physicians' services and outpatient services annual deductible	$ 2.234	Exhibit 6
E. Physicians' services coinsurance		
1. Annual services incidence per 100 contracts	444.293	Exhibit 7
2. Average payment per service	$ 8.89	Exhibit 8, Item D
3. Expected monthly pure premium	$ 3.291	[(Item E1)(Item E2 ÷ 1200]
F. Outpatient hospital service coinsurance		
1. Annual service incidence per 100 contracts	150.742	Exhibit 9
2. Average payment per service	$ 11.03	Exhibit 10
3. Expected monthly pure premium	$ 1.386	[(Item F1)(Item F2) ÷ 1200]

ITEM	AMOUNT	SOURCE
G. Skilled Nursing Facility		
1. Annual day incidence per 100 contracts	11.257	Exhibit 4
2. Average payment per day	$ 21.13	Exhibit 2, Item E
3. Expected monthly pure premium	$ .198	[(Item G1) (Item G2) ÷ 1200]
H. Prescription Drugs		
1. Average number of prescriptions per claim	8.054	Exhibit 12
2. Average charge per prescription	$ 8.86	Exhibit 14
3. Average charge per claim	$ 71.36	(Item H1) (Item H2)
4. Expected average payment per claim	$ 37.09	[$71.36 - $25.00][.80] = $37.09
5. Annual claim incidence per 100 contracts	72.772	Exhibit 11
6. Expected monthly pure premium	$ 2.249	[(Item H4) (Item H5) ÷ 1200]

Office of the Secretary

MEDICARE PROGRAM

Inpatient Hospital Deductible for 1979

Under the authority in section 1813(b)(2) of the Social Security Act (42 U.S.C. 1395e(b)(2)), I have determined and hereby announce that the medicare inpatient hospital deductible for 1979 shall be $160.

Section 1813 of the Social Security Act provides for an inpatient hospital deductible and certain coinsurance amounts to be deducted from the amount payable for inpatient hospital services and post-hospital extended care services furnished an individual during a spell of illness. Section 1813(b)(2) of the act requires the Secretary to determine and publish, between July 1 and October 1 of each year, the amount of the inpatient hospital deductible for the following calendar year.

Under a formula in the law, the deductible for calendar year 1979 must be equal to $40 multiplied by the ratio of: (1) The current average rate for a day of inpatient hospital services for calendar year 1977 to (2) the average daily rate for such services in 1966. The amount so determined is rounded to the nearest multiple of $4. The average daily rates are determined by the Secretary based on the amounts paid on behalf of insured individuals to the hospitals participating in the medicare program plus the amounts withheld because of the deductible and coinsurance provisions.

Because the applicable coinsurance amounts in section 813 of the Social Security Act are fixed percentages of the inpatient deductible for services furnished in the same spell of illness, the increase in the deductible has the effect of also increasing the amount of coinsurance the Medicare beneficiary must pay. Thus, for spells of illness beginning in 1979, the daily coinsurance for the 61st through 90th days of hospitalization (one-fourth of the inpatient hospital deductible) will be $40; the daily coinsurance for lifetime reserve days (one-half the inpatient hospital deductible) will be $80; and the daily coinsurance for the 21st through the 100th days of extended care services (one-eighth of the inpatient hospital deductible) will be $20.

The data used to make the necessary computations of the current average daily rate for calendar years 1966 and 1977 are derived from individual inpatient hospital bills that are recorded for all beneficiaries in the records of the program. These records show, for each bill, the number of inpatient days of care and the interim cost (the sum of interim reimbursement, deductible, and coinsurance). Tabulations are prepared which summarize the data from these bills by the year in which the care was provided. The resulting average interim daily rate accurately reflects interim costs on an accrual basis.

In order to properly reflect the change in the average daily hospital cost under the program, the average interim cost (as shown in the tabulations) must be adjusted for the effect of final cost settlements made with each provider of services after the end of its accounting year to adjust the reimbursement to that provider from the amount paid during that year on an interim basis to the actual full cost of providing covered services to beneficiaries. To the extent that the ratio of final cost to interim cost for 1977 differs from the ratio of final cost to interim cost for 1966, the increase in average interim daily costs will not coincide with the increase in actual cost that has occurred.

The current average interim daily rate for inpatient hospital services for calendar year 1977, based on tabulated interim costs, is $155.26; the corresponding amount for 1966 is $37.92. These averages are based on approximately 93 million days of hospitalization in 1977 and 30 million days in 1966 (last 6 months of the year). The ratio of final cost to interim cost is approximately 1.035 for 1977 and 1.055 for 1966. Thus, the inpatient hospital deductible is $\$40 \times [(155.26 \times 1.035)/(37.92 \times 1.055)] = \$160.67$, which is rounded to $160.

Dated: September 25, 1978.

JOSEPH A. CALIFANO, Jr.,
Secretary.

[FR Doc. 78-27363 Filed 9-28-78 8:45 am]

Table H-19.--OASDHI hospital insurance: Number of bills for inpatient short-stay hospital[1] care approved for payment, covered days, total charges, and amount reimbursed, by type of beneficiary and period approved, as of September 30, 1978 2/

Period approved 3/	Approved bills: Number (in thousands)	Covered days of care: Total (in thousands)	Covered days of care: Average per bill	Hospital charges: Total (in thousands)	Hospital charges: Per bill	Hospital charges: Per day	Amount reimbursed 4/: Total (in thousands)	Amount reimbursed 4/: Percent of total
Total 5/								
Jan. - Dec. 1973	7,083	78,037	11.3	$8,417,723	$1,188	$106	$6,339,512	75.3
Jan. - Dec. 1974	8,063	87,814	10.9	10,500,666	1,303	120	7,926,619	75.4
Jan. - Dec. 1975	8,530	90,104	10.6	13,084,084	1,534	145	9,831,171	75.1
Jan. - Dec. 1976	9,008	94,077	10.4	16,169,672	1,794	171	12,048,820	74.5
* Jan. - Dec. 1977	9,366	95,820	10.2	18,883,288	2,016	197	13,816,711	73.2
Jan. - June 1976	4,626	49,041	10.6	8,064,818	1,743	164	6,026,681	74.7
July - Dec. 1976	4,443	45,838	10.3	8,115,754	1,827	177	6,022,138	74.2
** Jan. - June 1977	4,761	49,479	10.4	9,439,281	1,983	191	6,944,770	73.6
July - Dec. 1977	4,605	46,341	10.1	9,443,987	2,051	204	6,871,911	72.8
*** Jan. - June 1978	4,869	50,054	10.3	10,872,413	2,234	217	7,823,683	72.6
Persons aged 65 and over 6/								
Jan. - June 1976	4,169	44,478	10.7	7,254,946	1,740	163	5,437,615	75.0
July - Dec. 1976	3,979	41,337	10.4	7,256,906	1,824	176	5,399,740	74.4
Jan. - June 1977	4,255	44,611	10.5	8,418,544	1,978	189	6,209,390	73.8
July - Dec. 1977	4,098	41,497	10.1	8,384,726	2,047	202	6,115,247	72.9
Jan. - June 1978	4,314	44,779	10.3	9,649,837	2,221	215	7,013,249	72.7
Disability beneficiaries 7/								
Jan. - June 1976	457	4,563	10.0	809,872	1,772	177	589,069	72.7
July - Dec. 1976	463	4,499	9.7	858,848	1,856	191	622,538	72.6
Jan. - June 1977	506	4,968	9.8	1,020,747	2,017	205	735,380	72.0
July - Dec. 1977	509	4,844	9.5	1,059,271	2,081	219	755,894	71.4
Jan. - June 1978	545	5,275	9.7	1,222,606	2,243	232	870,434	71.2

1/ General and special hospitals reporting average stays of less than 30 days.
2/ See table H-18, footnote 1.
3/ See table H-18, footnote 2.
4/ See table H-18, footnote 4.
5/ See table H-18, footnote 5.
6/ See table H-18, footnote 6.
7/ See table H-18, footnote 7.

* Average Hospital Charge Per Day for the Period January 1, 1977 to December 31, 1977.
($18,883,288) ÷ 95,820 = $197.07.

** Average Hospital Charge Per Day for the Period January 1, 1977 to June 30, 1977.
($9,439,291) ÷ 49,479 = $190.77.

*** Average Hospital Charge Per Day for the Period January 1, 1978 to June 30, 1978
($10,872,413) ÷ 50,054 = $217.21.

SELECTED DATA FROM THE MEDICARE PROGRAM

Item	1971	1972	1973	1974	1975
Persons enrolled as of January 1 for:					
Hospital insurance (HI)-aged	20,388,456	20,966,267	21,374,693	21,612,003	22,066,910
Hospital insurance (HI)-disabled	NA	NA	NA	1,830,852	2,069,744
Supplementary medical insurance (SMI)-aged	19,738,504	20,145,286	20,544,688	21,105,223	21,620,376
Supplementary medical insurance (SMI)-disabled	NA	NA	NA	1,657,497	1,855,301
HI and/or SMI(2)- aged	20,732,693	21,154,498	21,601,315	21,868,142	22,362,847
Amounts reimbursed during the fiscal year:					
HI: Total (in thousands)	$5,442,971	$6,109,139	$6,749,000	$7,805,000	$10,415,000
SMI: Total (in thousands)	$2,034,999	$2,255,069	$2,439,000	$2,865,600	$3,780,000
HI: Amount per HI enrollee	$264	$291	$316	$333	$432
SMI: Amount per SMI enrollee	$103	$112	$119	$126	$161
Participating facilities as of July1					
Number:					
All hospitals	6,745	6,726	6,757	6,733	6,773
Short-stay	6,153	6,131	6,132	6,102	6,107
Other	592	595	625	631	666
Skilled nursing facilities	4,287	4,041	3,977	3,952	3,932
Home health agencies	2,284	2,222	2,211	2,248	2,242
Independent laboratories	2,751	2,873	2,929	3,029	3,048
Beds:					
All hospitals	1,188,013	1,155,982	1,148,428	1,143,664	1,140,395
Short-stay	834,514	850,070	864,786	882,496	901,757
Other	353,499	305,912	283,642	261,168	238,638
Skilled nursing facilities	307,548	291,636	287,606	294,000	287,479
Beds per 1,000 HI enrollees:					
Short-stay hospitals	40.5	40.5	41.6	37.6	37.4
Skilled nursing facilities	15.2	13.9	13.8	12.5	11.9
Admissions (in thousands) during the fiscal year:					
All hospital inpatient admissions-aged	6,243	6,495	6,781	6,996	7,305
All hospital inpatient admissions-disabled	NA	NA	NA	663	787
Skilled nursing facility admissions-aged	421	397	405	425	436
Skilled nursing facility admissions-disabled	NA	NA	NA	13	15
Admission rate per 1,000 HI enrollees:					
All hospital inpatient admissions-aged	305	313	320	324	331
All hospital inpatient admissions-disabled	NA	NA	NA	362	384
Skilled nursing facility admissions-aged	20	19	19	20	20
Skilled nursing facility admissions-disabled	NA	NA	NA	7	7
Average charges per day (covered):					
Short-stay hospitals-aged	$84	$94	$102	$108	$130
Short-stay hospitals-disabled	NA	NA	NA	$117	$142
Skilled nursing facilities-aged	$30	$32	NA	$34	$39
Skilled nursing facilities-disabled	NA	NA	NA	$38	$45
Average length of stay (covered):					
Short-stay hospitals-aged	12.6	12.1	11.7	11.2	10.7
Short-stay hospitals-disabled	NA	NA	NA	10.3	10.1

(1) Includes U.S. and all outlying areas such as Puerto Rico, Guam and the Virgin Islands.
(2) Equals HI for disabled.
NA Not available.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
SOCIAL SECURITY ADMINISTRATION
BALTIMORE, MARYLAND 21235

REFER TO:
IHI-312

August 1976

PART B INTERMEDIARY LETTER NO. 76- 34

SUBJECT: Announcement of the Economic Index Applicable to Prevailing Charges for Physicians' Services for the Period July 1976 Through June 1977

In accordance with Public Law 94-368, the annual update of Medicare reasonable charge screens will no longer be related to the Federal Government fiscal year (FY), but will continue to be made on July 1 of each year. We will refer to this 12-month period beginning on July 1 as the fee screen year (FSY). This is to inform you that the economic index applicable to prevailing charges for physicians' services for the period July 1976 through June 1977 is 1.276. Accordingly, carrier prevailing charge screens for physicians' services will be permitted to increase for fee screen year 1977 in accordance with established reasonable charge methodology, but not more than 27.6 percent above fiscal year 1973 levels. Pursuant to section 2 of Public Law 94-368, the no-rollback provision of Public Law 94-182, which provides that prevailing charges will not be reduced below FY 1975 levels because of the application of the economic index, will remain in force for FSY 1977 and subsequent years. An announcement of the applicable index has been approved by the Secretary of Health, Education, and Welfare for publication in the Federal Register.

Public Law 94-368, enacted into law on July 16, 1976, besides establishing the July 1 through June 30 fee screen year and continuing the no-rollback provision (section 101(a) of Public Law 94-182), also provides that, for the 12-month period beginning on July 1, 1976, the annual update of prevailing charge levels shall apply to claims filed after June 30, 1976, with a carrier and processed by the carrier after it has made the appropriate changes in the prevailing charge levels. Hence, adjustments retroactive to July 1 will not be made. The economic index for FSY 1977 will also be applicable in the same manner, i.e., from the time of the carrier's update forward.

As you know, the economic index calculated each year consists of two components reflecting (on a cumulative basis) the changes that have taken place since calendar year 1971 in physicians' practice expenses and in general earnings levels. With the exception of malpractice insurance premium data, the data that have been used to calculate the economic index (see attached chart) were derived from The Monthly Labor Review published by the U.S. Department of Labor. For example, the Bureau of Labor Statistics index for nonsupervisory workers in finance, insurance, and real estate was used as a reasonable approximation of wage trends for persons employed by physicians. For office space, the housing component of the Consumer Price Index (which includes data on rentals as well as costs of home ownership, data on utilities, and other corresponding data) was used. For drugs and supplies, the drugs and pharmaceuticals component of the Wholesale Price Index was used. For physicians' automobile expenses, the private transportation component of the Consumer Price Index was used. For miscellaneous "other expenses," which include attorneys' fees, travel, food and lodging while away from home, and many other items, the entire Consumer Price Index was used. The weights assigned to the various components of the index were derived from Medical Economics (December 8, 1975) and from the Profile of Medical Practice (1974 edition).

When the economic index limitation on increases in prevailing charges for physicians' services was implemented under Medicare in fiscal year 1976, it was expected that the methodology for constructing the index would be refined over time. The changes considered in this regard have included adjustments for regional differences in cost increases, and adjustments for differential practice costs among specialties. However, lack of a sufficiently refined data base on physicians' practice costs has, so far, precluded these changes.

The only substantive change in the methodology for computing the economic index for the 12-month period beginning in July 1976 is the inclusion of a separate element to reflect the effect of malpractice insurance premium increases on physician office expenses. (Previously, malpractice insurance costs were included in the miscellaneous expense category of office practice costs.) The component of the index which measures the rise in malpractice insurance premiums is based on a survey of the premiums charged in 46 States by six major insurers who, collectively, write about 70 percent of all malpractice insurance in the United States and thus provide a representative sample of malpractice premium rates nationwide. It provides a measure of the percentage increase in the premiums in calendar year 1975 over 1974. (Reliable separate data on malpractice insurance costs for earlier periods are not available.)

To accommodate the lack of prior (separate) data on malpractice insurance costs, the other components of the index have been computed on an annual basis to reflect the changes in these components in 1975 over 1974. The calendar year 1974 data used for these components in the calculation of the economic index, in turn, reflect the cumulative increases since calendar year 1971. Therefore (except for the malpractice insurance data), the economic index (1.276) for the 12-month period beginning in July 1976 reflects the cumulative change in the components of the index since calendar year 1971, as is intended by section 1842(b)(3), as amended by Public Law 94-368, of the Medicare statute and by section 405.504(a)(3)(i) of the regulations.

Also, some of the calendar year 1974 data used reflect information that became available from the Bureau of Labor Statistics after the economic index for fiscal year 1976 (1.179) was calculated last year and put into effect. The economic index for a particular period must necessarily be calculated on the basis of the best information that is available at the time the calculation is made and put into effect. Therefore, the adjusted data have been used to calculate the economic index for the period July 1976 through June 1977 in order to provide the most accurate calculation that is possible at this time of the changes that have taken place in the components of the index since the base year (calendar year 1971).

In view of the urgency of this activity, we request that you take all necessary actions, including appropriate regional office approval, to update the reasonable charge screens no later than September 27, 1976. Please note that the updating of the screens must be in accord with previously issued instructions, including Part B Intermediary Letters No. 76-30 and No. 76-31.

Thomas M. Tierney, Director
Bureau of Health Insurance

Attachment

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
HEALTH CARE FINANCING ADMINISTRATION
BALTIMORE, MARYLAND 21235

REFER TO:
IHI-312

June 1977

PART B INTERMEDIARY LETTER NO. 77-24

SUBJECT: Annual Reasonable Charge Update - Economic Index Applicable to Prevailing Charges for Physicians' Services for the Period July 1, 1977, Through June 30, 1978

This intermediary letter is to inform you that the economic index applicable to prevailing charges for physicians' services for the period July 1977 through June 1978 will be 1.357 (i.e., 35.7 percent above fiscal year 1973 levels). This economic index for the 12 months beginning July 1, 1977, represents a 6.35 percent increase over the economic index (1.276) used for the previous 12 months. Carriers will therefore use a 1.0635 figure where an annualized index is applied in accordance with Part B Medicare Carriers Manual section 5020.3C3. All carriers should, in accordance with the established reasonable charge methodology, continue to develop updated customary and prevailing charge screens for fee screen year 1978 based on calendar year 1976 charge data, and implement the indicated economic index limitation on prevailing charge increases. We request that you take all necessary actions, including regional office approval, to update the reasonable charge screens on July 1, 1977.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
HEALTH CARE FINANCING ADMINISTRATION
BALTIMORE, MARYLAND 21235

REFER TO:
HH-312

ADVANCE COPY

June 1978

PART B INTERMEDIARY LETTER NO. 78-23

SUBJECT: Annual Reasonable Charge Update - Economic Index Applicable to Prevailing Charges for Physicians' Services for the Period July 1, 1978, Through June 30, 1979

This intermediary letter is to inform you that the economic index applicable to prevailing charges for physicians' services for the period July 1978 through June 1979 will be 1.426 (i.e., 42.6 percent above fiscal year 1973 levels). This economic index for the 12 months beginning July 1, 1978, represents a 5.08 percent increase over the economic index (1.357) used for the previous 12 months. Carriers will therefore use a 1.0508 figure where an annualized index is applied in accordance with Part B Medicare Carriers Manual section 5020.3C3. All carriers should, in accordance with the established reasonable charge methodology, continue to develop updated customary and prevailing charge screens for fee screen year 1979 based on calendar year 1977 charge data, and implement the indicated economic index limitation on prevailing charge increases. We request that you take all necessary actions, including regional office approval, to assure a timely update of reasonable charge screens.

Thomas M. Tierney, Director
Medicare Bureau

SELECTED BIBLIOGRAPHY

Books

Abernathy, David S.,; Pearson, David A. "Regulating Hospital Costs - The Development of Public Policy" Aupha Press, 1979

Myers, Robert J. "Social Security" McCahan Foundation, 1975

Reports

Health Insurance Institute, "Source Book of Health Insurance Data" 1978-1979

Employee Benefit Plan Review Research Reports, Medicare

Federal Register

SSA Medicare Bulletins

THE PRICING OF MEDI GAP CONTRACTS

By Emil J. Strug

REVIEWED BY Robert F. Bartik

I feel the approach used and described in the paper to be very useful. The author covered the specific subject of using primarily external data to produce a composite claim cost for a mixture of coverages for an older age population very well.

I will make a few comments regarding individual items directly related to the paper and then continue to pursue the general subject of rating a Medicare supplement type plan, extending into a few areas not covered by this paper. As we approached this problem ourselves a few years ago, we also tackled many of the items referred to in the paper. At that time some of the external data referred to was not available, but I would feel a strong inclination at this time to pursue an adoption of much of the procedure presented here.

However, as we advanced further in our research on the subject, we were soon led to the conclusion that for the type of contract we finally decided upon, that the subsequent considerations proved to be of such a significant character, in regard to determination of final rate level, that further sophistication at that time in regard to this aspect of the detail would not have great impact on that final determination.

The extreme costs of medical care for the aged strongly influenced employers when they tried to include retirees or even active employees over 65 into group health plans. Therefore, what could have been a natural relief mechanism for the problem did not hold and increased the bind the aged were caught in, producing even greater pressure for the passage of Medicare.

The tables showing the impact of the Medicare Program on the population are horrendous in themselves, but an even more comprehensive study (showing more horrendous results) may be found in "Ten Years of Medicare: Impact on the Covered Population" (SSB July 1976, pp 3-21). Table 20 (page 18) of that source puts Table 1 of this paper in better perspective in that it gives not only percentage, but per capita dollars, and in that the record of the insurance industry is set out by itself.

The third of the major elements of cost should take into account the "Reserve Days" under Medicare. It may be that this is accounted for elsewhere, but I was not able to discover it.

The Part B coinsurance in the plan depends on actual Medicare allowances. There has been a definite trend toward Medicare allowing a lesser and lesser percentage of actual charges by a process of charge screens, the adjustment of which have not even come close in recent years to matching the increase in level of actual charges of the medical industry. Accordingly, our trend on this portion of the coverage is not based on the factors in the intermediary letters (Appendix D) but on either the CPI or on the SSB tabulations of Part B charges.

The author has introduced a significant variety of curve fitting of the data for purposes of projection and has selected the process of averaging these results for the final answer. I suspect that if one were to attempt a more sophisticated analysis of these procedures over time, significant refinements would become apparent, such as a weighting of the curves under various circumstances and the total elimination of some, with even the possible emergence of a best single curve, producing better results than any composite.

Once the basic data has been determined, the process of using it to produce the estimated parameters entering into the subsequent steps is very well displayed, explained and easy to follow. The subsequent steps of modifying and tieing together the estimated parameters to produce the separate pieces of the total claim costs is also handled very well and easy to follow. The final summary of the separate pieces into a consistent whole total is a neatly laid out summary, totally contained within the single Exhibit 15. Actually there are many different policies on the market offering some degree of supplementation of medicare benefits and even those that tend to offer in effect full supplementation, there are differences. There are also some differences between the typical contract offered by the "Blues" as opposed to those offered by other comapnies. A very significant difference between policies is referred to later in this review. Therefore, I would like to have had included a more detailed outline of the policy coverages and provisions referred to in the paper.

One additional major factor involved in the final rate structure is the decision regarding rate relativities by age. Although the claim costs for the senior citizen classifications is significantly higher than under 65, there are significant differences by age within the classification itself, easily ranging up to two and one half to three times the cost in the higher age brackets (depending upon where one decides to establish an "all over age" bracket) vs. the 65 to 69 bracket. Therefore, a rate derived for the composite of the classification derived from the population date of the classification will reflect the age distribution of that population, but can conceivably attract a much different distribution of insureds by age, very likely including most of the older individuals but a significantly lesser proportion of those from 65 to 69.

Another major factor in the rating of this block of business involves a high proportion of business being marketed with level rates even when separate classifications by age are used. Therefore, significant attention must be paid to the persistency expectations and although one can take from a claim standpoint at least a conservative approach, and assume only normal mortality, it is highly likely that a significant portion of insureds do not continue with the policy selected for their entire remaining life span, and as a result, such a conservative approach, although commendable from a safety standpoint, will undoubtedly produce a totally non-competitive rate in the marketplace and produce zero business from which to be even able to determine its adequacy. Interestingly, although interest assumptions could also prove a significant factor, the net effect tends to be significantly reduced because of the dramatic drop off in persistency, if for no reason than mortality, thereby lessening the major effect of the increasing claim cost by age and its effect upon the investment income.

One significant aspect of this type of policy marketed to senior citizens who are already at the age where they have or soon will have significant continuing medical problems, is the fact that it is usually written with minimal or no underwriting, that is, in effect a mass marketing type operation. An attempt is made to offset the effects of this through the use of preexisting condition clauses with time limitations, ranging from a few months to two years. The variation of expected results from this clause can vary dramatically. The primary effects of the clause are to set up a screening device which will deter the poorest risks from selecting this particular policy and (especially in this particular marketplace) from transferring from an existing policy to a new one.

In addition, there is a direct claim cost savings from the non payment of expenses during the preexisting period. However, there is a significant limitation upon this since the preexisting clause does not totally and indefinitely exclude continuing conditions. That is, once the time interval has passed, medical expenses from that point on for an already existing disability become payable. Therefore, there is a significant question regarding the degree of effectiveness of the provision and the competitive situation has produced a relatively short time interval common on this particular form, normally three to five months. Since little data is available regarding the value of such clauses even of longer duration and on younger populations, for the circumstances of short duration and older populations it is virtually non-existent. The expected claim costs can easily be double that of an underwritten or even population block with zero preexisting, which will hopefully range down to equality of expected claim cost at some point near two years.

The significant question is, what is such a value for say, five months preexisting. It is my feeling that this is still a rather significant and elusive figure, but in any event, the primary need is for the proper factor to use in determining the premium to charge for this coverage. In any event, we are in the process of accumulating experience data under this form which will answer many of these questions directly at least as a composite of the structure we are using. At that time a procedure such as this will be essential to keep pace with the changing conditions, both of the economy and the legislation, as well as its implementation.

TITLE: IS ECONOMETRIC MODELING OBSOLETE?

AUTHOR: Mr. Oakley E. Van Slyke

Mr. Van Slyke is Consulting Actuary with Warren, McVeigh and Griffin. He receives his FCAS in 1980 and is a member of the American Academy of Actuaries. Lee holds a B. S. degree from the Massachusetts Institute of Technology.

REVIEWER: Mr. Michael Fusco

Mr. Fusco is Vice President-Actuary, Personal Lines for Insurance Services Office. Mike has a B. S. in Mathematics and a J.D. from Fordham Law School. He received his FCAS designation in 1976 and is a member of the American Academy of Actuaries. He currently serves on the CAS Examination Committee and is editor of the CAS yearbook.

IS ECONOMETRIC MODELING OBSOLETE?

Econometric models are widely used to forecast economic events. A number of macroeconometric models are well known, including those of the Wharton School, Chase Econometrics, Data Resources, Inc., and the Federal Reserve Bank of St. Louis. Less imposing models are cropping up in all walks of life. The Insurance Services Office (ISO) is studying the application of econometric models to actuarial problems. Because of the effect these models have on our lives through government planning, and because of the possible effect they may have on our livelihoods if they become standard actuarial tools, it is wise for us to understand what econometric models are, why they are increasingly popular, and why they may be only a precursor of even more dramatic changes to come.

For the purposes of this paper an econometric model is a mathematical representation of economic relationships using linear equations. A model may consist of one equation or of many. As an example of a model with one equation, one might represent the relationship between bodily injury loss costs, wages and medical prices using

$$Y_t = C \cdot W_t + C \cdot M_t$$

where

Y_t = Average Claim Size for Bodily Injury Claims at Time t

W_t = Wage Index at Time t

M_t = Medical Index at Time t

C_1, C_2 are constants, usually set to reflect historical data about Y, W and M

In an econometric model the dependent variable (such as average claim size) and the independent variables (such as wages and medical costs) are represented by time series. These time series are usually observed values taken at regular intervals over time. The variables may also be the changes in time series, in either absolute or percentage terms.

Econometric modeling, therefore, refers to a particular way of describing economic relationships. Econometrics, on the other hand, refers to the broader arena in which quantitative methods are applied to economic problems. All of the techniques discussed here are within the field of econometrics, and all are models. The term "econometric model," however, commonly refers only to the type of model described above. We will continue that common usage here.

HISTORY

Paul Samuelson (12) has traced the history of macromodels for over 40 years. Macromodels are models which attempt to encompass all of the major relationships in a particular economy. Although the earliest model cited by Samuelson was the effort in 1932 by Ragnar Frisch, Samuelson says that Jan Tinbergen's model of the Dutch economy in 1935 was the "fountainhead and source." Macromodels were used in the 1940's to describe and prescribe wartime and post-war development and planning. Today's major macromodels were begun in the 1950's and 1960's and revised as theory and conditions changed.

Simpler models, like the model of average claim size given above, have an even longer history. These models are applications of linear regression. Linear regression methods date back to the turn of the century.

Quite often these simple econometric models rely heavily on the relationship between the dependent variable and a measure of time. This is done because the dependent variable cannot be forecast until the independent variables have been forecast and any errors in the forecasts of the independent variables will affect the forecast given by the model. This error problem doesn't exist if time is used as the independent variable. On the other hand, forecasts of various economic indices have become more readily available during the 1960's and 1970's. As a result, simple models based on these indices have become more popular in the past few years.

Simple econometric models have been advocated and employed by casualty actuaries during the 1960's and 1970's. Masterson (8) set forth a number of claim cost indices in 1968. He has periodically written about these indices to keep them up to date. They are weighted averages of various published cost indices. The weights are set by judgment, not by regression (least-squares) techniques.

Finger (1) has suggested various mathematical models of loss costs for which the parameters were to be found by a method of least squares. He suggested as an independent variable a sort of operational time (fraction of claims closed) as well as time itself.

In 1974 Lommele and Sturgis (6) published models of aggregate premium and loss statistics for workers' compensation. They used a conceptual framework in which the independent variables were time series taken from macroeconomic models and insurance industry statistics. Time was not one of the independent variables. Parameters (the c_i's in the average claim cost model above) were found by a least squares method. This approach is being pursued by the ISO as a possible ratemaking step for some lines of insurance. The papers by Masterson and Lommele and Sturgis are now listed in the Recommendations for Study for the examinations given by the Casualty Actuarial Society.

ADVANTAGES

Econometric models have a number of advantages over less mathematical forecasting methods. Arthur M. Okun (10) listed the following advantages in his discussion of macromodels:

1. The objective framework permits the organization of a team effort with a division of labor.
2. The mathematical interrelationships in the model result in a consistency among the component elements of the forecast.
3. The reproducibility of the forecast permits the model user to conduct a post mortem analysis to identify the causes of poor predictions.
4. The objectivity of the forecast is itself desirable.
5. Because the steps leading to the forecast are documented, the model provides for a cumulating of knowledge.

6. The models are labor-efficient because a computer can be used for the routine calculations.

The ISO suggests that its simple models have at least the following additional advantages:

1. They should produce better forecasts than other techniques because they are more sophisticated.
2. The objectivity of the forecasts should make them more acceptable to regulators.
3. The use of non-industry data should lead to:
   - More credibility to the layman
   - A more defensible explanation of cost changes
   - Earlier warning of turning points in the time series being projected
   - Greater accuracy because the non-industry data is generally more current than data about losses.

SUCCESSES AND FAILURES

Econometric models have disadvantages as well as advantages. The history of econometric models includes both successes and failures.

Sophisticated macromodels have had their share of successes and failures. As Samuelson noted, "The famous consensus forecast by government economists of great post-war unemployment did not advance the prestige of the method." The simultaneous increases in inflation and unemployment in 1975 and 1976 (and perhaps late 1979) are not explained by the major macromodels. Samuelson cites a study by Robert Adams of the accuracy of different methods of forecasting the national economy. In Adams's words, "being Sumner Slichter" was apparently better than using any econometric model.

The simpler models now being advocated in actuarial circles have had an even more dismal past. Discussing the models of the U.K. Treasury and Britain's National Institute of Economic and Social Research, Ramsey (11) noted that the models showed "no tendency to improve over time." These models were characterized by their simple assumptions, by stable patterns in the data over time, and by the use of simple trend relationships. (Ramsey contrasted them to macromodels in the U.S. and the Netherlands. These, he said, tended to improve over time as they became more complex and began to reflect dynamic interrelationships.) One prediction by a member of the Academy will illustrate the dramatic way in which these models have sometimes failed.

To preserve the anonymity of the actuary, hypothetical data will be used here. While the actuary forecasted several time series, we shall forecast only one in our example. The actuary had a time series (dependent variable) over a period of 24 months. He assumed that the dependent variable grew exponentially over time. He used a linearized regression model to project the values for a total period of fifteen years. (A linearized regression model has the form $\ln Y = a + bt$.) The regression is shown graphically using hypothetical data in figure I-A. Note that the statistics of the regression indicate a good fit.

The alternative models shown in Exhibit I also fit well. Unfortunately, they lead to radically different forecasts. The indicated range for the predicted values at the end of 15 years is from about 200 to about 200 billion. This is because of the length of the forecast period. Clearly all three models are inappropriate because the month-to-month changes in the dependent variable tell us practically nothing about the changes that will take place in future years.

Similar poor results have been noted in forecasts of medical malpractice costs. I was once part of a three-man team making actuarial projections of the average loss cost per doctor (pure premiums) in a particular state. The projections were made in late 1975; accident year 1976 was being forecast. Sixteen policy years of data were used. Models similar to those described by Finger were applied. The various models projected pure premiums of $7,000, $14,000 and various amounts in between. Again modeling failed to provide a useful prediction. At least in this case the indicated range was useful.

The track record of econometric models suggests that they cannot be relied upon to produce useful predictions; some expertise must be applied to the results of the model. As Okun put it, "In fact, virtually nobody takes economic forecasts straight out of a model really seriously as the sole guide to a forecast of the near-term future. Quite apart from filling in the exogeneous variables, every model builder or model user has to adjust equations."

Both Okun and Samuelson regard the forecasts from models as references. Again in Okun's words, ". . . the model as a forecasting device is not an alternative to judgment. It is not a product in and of itself. It is a tool in the hands of a trained economist."

DISADVANTAGES

There are several reasons why econometric models can produce poor forecasts. First, they require accurate projection of the exogeneous variables. Exogeneous variables are those which the model itself does not attempt to predict. In our simple example the exogeneous variables were the wage index and the medical cost index. The projection of average claim cost can be no more accurate than the projection of wages and medical costs.

Second is the index-number problem. Practicable econometric models must incorporate data about the real world in summary fashion. The price of every type of consumer good cannot be fed into the model because the number of consumer goods that can be distinguished from one another is practically beyond enumeration. Instead, the prices of a few representative goods are measured. These components are then combined using some weighting scheme. All the components of an index must be moving the same direction at the same rate, and the weights must be constant and appropriate at all times, or the index will not accurately measure the "average" it purports to represent. It should be clear that no index can in practice pass these tests. This tautology - that index numbers are necessary abstracts that cannot be accurate representations of costs - is called the index number problem.

The third source of error is that the wrong variables might be included in the model. This includes the possibility of leaving out the right variables. In the example of the 15-year forecast described above, it is clear that the actuary left out some important considerations about long-term changes when he made his 15-year forecast. In ratemaking problems it is especially hard to know what variables to include. For example, should the Consumer Price Index be used in rating Homeowners insurance because some Homeowners losses involve consumer goods? Statistical tests can tell us the probability that we will err by removing a certain variable, but no test can tell us if we have included the right variables.

The fourth source of error is that the variables might be interrelated in the wrong way. We may assume that simple relationships are stable, when in fact they are changing. Or we may choose the wrong relationship. There are many to choose from, even if there are only two variables and the dependent variable (Y) is an increasing function of the independent variable (X). For example, the following relationships (and others) are sometimes used:

$$Y_t = a + bX_t \qquad Y_t = aX_t^{\,b}$$

$$Y_t = a + b^{X_t} \qquad Y_t = a + X_t^{\,b}$$

$$Y_t = ab^{X_t} \qquad Y_t = a + bX_t^2$$

$$Y_t = a + bX_{t-1} \quad - \quad \text{etc.}$$

$$\Delta Y_t = a + b\Delta X_t \quad - \quad \text{etc.}$$

$$\frac{\Delta Y_t}{Y_t} = a + b\,\frac{\Delta X_t}{X_t} \quad - \quad \text{etc.}$$

This picture is further complicated when there are two or more independent variables.

In this respect, econometric models have been criticized sharply for their inability to deal with the interrelationships of the real world. According to Forrester (3):

> "Our social systems belong to the class called multi-loop nonlinear feedback systems . . . .
>
> "A great computer model is distinguished from a poor one by the degree to which it captures more of the essence of the social system that it presumes to represent. Many mathematical models are limited because they are formulated by techniques and according to a conceptual structure that will not accept the multiple-feedback-loop and nonlinear nature of real systems."

System dynamics may provide a way to make models more realistic.

Also, econometric models rarely predict the sudden changes that can sometimes occur. It would be difficult to conceive of an econometric model that would have accurately predicted Iran's GNP in 1979, even from a vantage point in early 1978. Another approach, called catastrophe theory, may be useful for predicting sudden changes.

SYSTEM DYNAMICS

System dynamics is a way of mathematically describing the components of a complex system so as to focus attention on the pressures that build up in the relationship between the components. As an example, consider a simple process of exponential growth in, say, premium. An econometric model would view exponential growth as a relationship between premium and time:

$$P_t = ab^t$$

System dynamics would view this as a relationship between premium at a particular time t and premium at an earlier time:

$$P_t = b \cdot P_{t-1}$$

In this elementary example the econometric model and the system dynamics model would both give the same answer. In practical problems this will not generally be the case.

Once we have removed the limitations on the model relationships that are inherent in econometric models, many economic and social systems can be modeled more realistically. System dynamics provides for non-linear feedback mechanisms in the model. We shall refer the reader to Forrester (3, 4, 5) for a large number of examples and a detailed explanation of why this is so, but a simple example from insurance will illustrate the point.

We are all familiar with the presence of underwriting cycles in our business. Most students of the cycles have observed that losses are growing at a relatively steady rate, while fluctuations in premiums produce most of the cyclical effect. Stewart (13) has explained the mechanisms involved:

> "Like farmers, insurers meet a fairly constant demand for what they sell. Even more than farmers, they can vary the amount they sell rather finely and quickly. Later on they may not like what was done with prices, underwriting and so forth - any more than farmers like what happens to their prices when they all plant fencepost to fencepost. But the decision to change supply can be carried out . . . .
>
> "For the main lines of insurance and for the industry as a whole, we can call the turns in the underwriting cycle quite reliably two years in advance . . . .
>
> "Even when warned, the individual insurer is trapped. He can only lower prices in advance if willing to smooth the cycle by giving up profits before the top. He can only raise prices in advance if willing to give up customers before the bottom. Either one is asking a lot of human nature and even of good business sense."

Econometric models (in the common use of the term) cannot deal with this behavior. System dynamics is specifically designed to deal with feedback mechanisms like this. The propensity to raise supply when reports of profits are received and reduce the supply when reports of losses are received is a feedback mechanism.

The feedback loop for the underwriting cycle is shown in Exhibit II. As the supply of insurance is increased faster than the demand grows, profit falls. When current profits are falling from black ink to red, insurers have the maximum accumulated profit. This is one pressure to reduce rates. As losses cut into accumulated profits, the insurer's capacity is reduced and it begins to restrict the supply of insurance. Restrictions are tightest when accumulated profits are at a minimum. An accurate model would also reflect the effect of financial reports showing unprofitable underwriting results, and perhaps the current practice of using five years' of trended data in a standard rate filing.

Models that allow for negative feedback predict that cycles will occur. In general, the response to a stimulus will be greater than is needed over the long term, and the system will overshoot its best long-term values. It will then respond to this error by overshooting in the other direction. The results will be patterns like those in Exhibit III. Exhibit III shows the patterns in theory and an example from medical malpractice.

It is important to contrast this type of model with econometric models. The model for workers' compensation written premium suggested by Lommele and Sturgis was:

$$WPREM_t = 289{,}184 + 5{,}687.23(WAGE_t)(PC_t)(RATE_t)(WO_t)$$

where

$WPREM_t$ = Written premium in year t.

$WAGE_t$ = Wages and salaries disbursed in billions of dollars in year t.

PC_t = Percent of the work force covered by workers' compensation in year t.

$RATE_t$ = Average countrywide rate level index in year t for workers' compensation including law amendments.

WO_t = A wage offset calculated to reflect the effect of payroll limitations for year t.

Clearly this model does not explicitly include any provision for changes in the supply of workers' compensation insurance. This is not merely a peculiarity of the model suggested by Lommele and Sturgis, but is a characteristic of the type of model commonly meant by the term "econometric model." As Lommele and Sturgis point out, future values of $RATE_t$ must be supplied by the analyst and are not produced by the model itself. The analyst can reflect in his estimate the effect of past values of $RATE_t$, but this is beyond the scope of the model. System dynamics is designed to bring these considerations within the scope of the model so they can be made explicit.

The history of system dynamics illustrates the major advantages and disadvantages of the approach. The first applications were in the physical sciences. The distance from the Earth to the Sun is the result of the effect of a feedback mechanism (the law of conservation of energy) on the movement of the Earth. This distance varies regularly as in curve C of Figure II-A. Radio squeal, the high-pitched sound one hears when changing stations, is the result of explosive negative feedback as in Curve A of Figure II-A. The Automatic Frequency Control on FM radios is an example of a damped cycle, like that of Curve B. Fluctuations in the radio's tuning are damped out by this circuit. System dynamics obviously works well when the real world can be modeled accurately.

The first applications to economic problems were for manufacturing firms. Inventories, employment levels, orders in process, rates of delivery and other variables were successfully modeled using system dynamics. The success was less complete than it had been for physical systems, of course. It was more difficult to identify the correct interrelationships in the industrial firm. Nonetheless, interviews often developed the necessary information about why orders were placed, why people were hired for overtime or laid off, and so forth. According to Forrester (4), this application, called industrial dynamics, has been generally successful.

System dynamics has also been applied to the economies of several cities, to the production sector of the U.S. economy, and to the world economy [c.f. Forrester (3), Mass (7) and Forrester (5)]. These applications have been useful in identifying the counter-intuitive behavior of social systems. For example, in a discussion of urban dynamics Forrester (3) observed, "To try to raise quality of life without intentionally creating compensating pressures to prevent a rise in population density will be self-defeating."

Nonetheless, system dynamics has been much less useful in producing practical recommendations for such social systems than for industrial systems.

The major reason for this lack of success appears to be that the predictions of the models are sensitive to the assumptions of the model. Also, the limited experience of the builders of systems dynamics models has not been enough to develop a set of assumptions with which most planners will agree. Forrester (3), for example, appears to assume that an increase in the quality of life will lead to an increase in population, if all else is equal. Yet many demographers have observed social systems in which a rise in the quality of life was associated with a decline in birth-rates to replacement levels or below. Another assumption that Forrester (3) made in his study of world dynamics was to include medicine and public health as a part of industrialization. At the same time, he assumed that increased industrialization would lead to increased pollution and, in turn, to a decline in public health. It is unlikely that an increase in medicine and public health would directly increase pollution and ill health.

Second, the model framework for system dynamics predicts only a few types of sudden responses. Other types of sudden responses may take place that cannot be modeled using system dynamics. These have been more accurately modeled using catastrophe theory.

In spite of the shortcomings of several recent applications, system dynamics models are better than econometric models in certain circumstances. One major area of use is in modeling parts of the insurance business that are characterized by negative feedback mechanisms. Underwriting is one example; insurers tend to increase the supply of insurance when they are receiving feedback that their capacity is at an unusually high level even though the resulting new business may be unprofitable. This happens for individual lines such as medical malpractice as well as for insurance in total.

Also, system dynamics models may be more useful than econometric models if they provide a more accurate abstraction of the real world. Models should teach as well as predict. If the limited model structures of econometric models are not instructive, the more flexible structures of dynamic models may provide the desired insight. For example, the model $Y_t = ab^t$ may give the same prediction as the model $Y_t = bY_{t-1}$, but the latter may make the growth process more clear.

CATASTROPHE THEORY

Catastrophe theory is a mathematical model of some common types of catastrophes. For the purposes of this theory, a catastrophe is a special kind of event or result: an abruptly changing effect resulting from a continuously changing force. There is a catastrophe in the making whenever the straw can break the camel's back. An example from Zeeman (15), with due credit to Conrad Z. Lorenz, is of aggression in a dog. As Van Slyke (14) wrote:

> ". . . . It can be observed that gradually increasing fear in the emotional make-up of a slightly angered dog will result in only a slight change in that dog's behavior. (We assume here that the dog is not angry enough to attack.) This gradual change will continue until at some level of fear the dog will suddenly turn and flee; that is, the increasing fear will at some point cause a sudden change in the behavior of the dog. This special type of catastrophe is only roughly similar to our usual uses of the word. For example, bridge collapses and buffalo stampedes are catastrophes in either sense of the word; an outbreak of a contagious disease would not be a catastrophe covered by this theory."

In catastrophe theory the dependent variable may be abruptly changing. The independent variables, on the other hand, are changing smoothly. In the case above, the independent variables were fear and rage or anger. An attractor is analogous to a dependent variable, but can include a whole set of behavior attributes. It is a stable or equilibrium state of behavior. For example, at a certain level of both fear and rage, the dog had one stable pattern of behavior: to stand snarling; or to flee; or to do something else.

Van Slyke cites as an even clearer example of an attractor an example Zeeman gives of tropical fish:

> "Some tropical fish exhibit territorial behavior, building nests, defending these nests from foes and using them as sanctuaries. A fish of this type, if foraging away from its nest, would flee from a larger fish. It would continue to flee until it reached an unseen boundary near its nest that we call its defense perimeter. Upon reaching the defense perimeter, the fish would turn and defend its nest. Similarly, a fish near its nest would defend that nest out to what we might call an attack perimeter. There is a pattern of behavior that causes the fish to turn and defend when it reaches the defense

perimeter and that causes the same fish to advance and attack so long as it stays within its attack perimeter. That pattern of behavior is an attractor. Although other behavior might be exhibited by the fish, the attractor is far and away the behavior that is most likely."

Catastrophe theory can be illustrated with an insurance example. Consider the insurance or self-insurance of losses. A move to self-insurance often results in a rapid reduction of insurance premiums by 50% of more. As the costs of using insurance to provide for losses increase with the growth of a business, the business is more and more likely to establish a self-insurance program. Usually the business will not establish the self-insurance program until well after the time that self-insurance becomes financially advantageous. Then it will keep the self-insurance program, even if the financial advantages diminish (perhaps because of a softening of insurance markets or a reduction in the size of the business).

Catastrophe theory can be useful in describing situations having five particular qualities. First, a catastrophe exhibits behavior that has two likely states. In the case of self-insurance, the likely states are insurance and self-insurance. Second, a catastrophe exhibits sudden transitions between these states. The transition to self-insurance takes place on one particular day when the amount of insurance is reduced. Third, in a catastrophe the place of the transition between the states depends upon the direction that the behavior is changing. For this reason, the financial advantage required to begin a self-insurance program is less than that required to continue it. A fourth quality of catastrophes is that they lack a middle ground of behavior. Usually a significant self-insurance retention is taken, if any. The fifth and last quality of catastrophes is that a very small change in the initial conditions can result in a very large change in behavior. For example, if a business felt that the costs of insurance were great, a slight rate increase could trigger a move to self-insurance. If, on the other hand, the business had been satisfied with its insurance, the same rate increase (or the same resulting rate level) could produce no change at all.

Econometric models, system dynamics models, and catastrophe models can be contrasted by imagining the possible values of the dependent variable as points on a surface. Exhibit IV attempts to illustrate the major differences between the three approaches, without attempting to provide any further explanation of theory.

Exhibit IV-A illustrates the basic premise of econometric models: that things will continue to change according to some preordained pattern. Every movement in the independent variable produces a change in the dependent variable according to a preset relationship. The relationship is embodied in the surface shown in the exhibit.

In system dynamics models, all of the variables are functions of time and of one another. Imagine a marble rolling along a trough (see Exhibit IV-B). The height and sideways displacement of the marble, and the speed of the marble in each direction, are tied together in a relationship that does not change over time. In the case of a marble rolling in a trough, these factors are related by physical laws dictating that the total energy of the system is constant. If energy is removed from the system (perhaps by friction), the marble's path will be a damped cycle. If energy is added (as a child pumps a swing), the marble's path will be an explosive cycle. Of course, economic models are much more complicated than this example.

In catastrophe models, the dependent variable depends on the values of the independent variables and the past history of the system. The interrelationships are visualized as a folded surface. In the path shown in Exhibit IV-C, a catastrophic drop in the dependent variable has occurred when changes in the independent variables have moved the system over the edge of the fold (solid path). Had the independent variables changed in a different way (dotted line), the same final values would have been reached for all variables without a catastrophic change. Also, the same values of indepedent variables can be associated with different values of the dependent variable, as illustrated by points A and B. Whether the dependent variable will be at A or B depends on the history of the independent variables.

Zeeman mentions uses of catastrophe theory in the fields of behavioral science, as indicated by these examples, and biology, physics, engineering and the development of a science of language.

Catastrophe theory is new. Although it hasn't been used in insurance, it should be useful whenever the five qualities of a catastrophe are present. The field is so wide that examples come easily, e.g., the formation of captives and doctor-owned insurance companies.

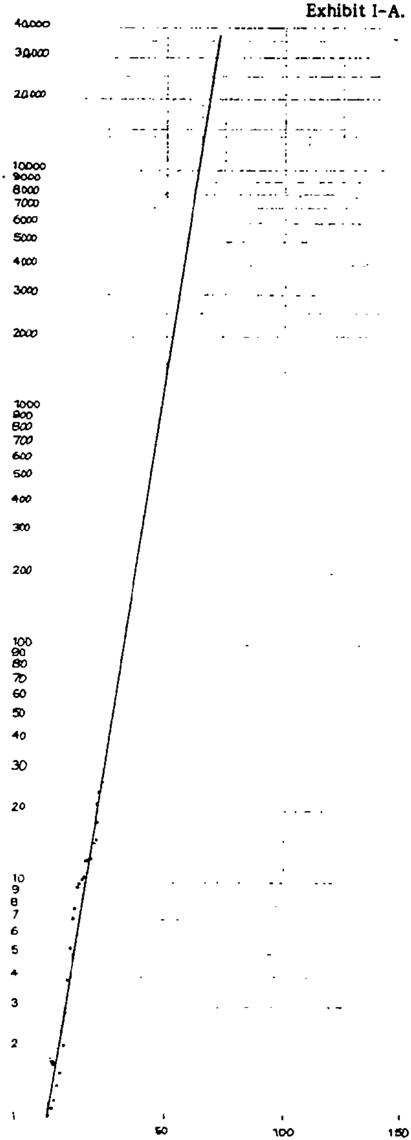
CONCLUSION

Econometric models are useful tools for actuaries. They can offer many advantages, especially when used as a tool in short-term forecasting. These advantages include their objectivity, which permits a division of labor, greater credibility with regulators, and cumulating knowledge; mathematical explicitness, which allows the analyst to identify the causes of poor predictions, efficient use of computers and a consistency among the elements of the forecast; and the use of non-insurance data, which provides more credibility with laymen, a more defensible explanation of cost changes, possibly earlier warning of turning points, and greater accuracy by reducing the analyst's reliance on immature loss data. Econometrics is not obsolete.

Nonetheless, it is not the most sophisticated forecasting tool available. The best model is the one that best represents the relevant qualities and relationships in the real world. This may be an econometric model. But in some problems it is important to recognize that the variables are all interrelated, and that a change in one causes feedback to the others. In other problems it is important to recognize that catastrophic change can occur, and that the effect of the economic environment may depend on the history of that environment. In these cases, the more sophisticated models of system dynamics or catastrophe theory may be better than econometric models.

Most important of all, the models are just tools. Because they will always fail to recognize the complexities of the real world, they must be just a part of the forecasting process, not a replacement for it.

Exhibit I-A.



Regression Equation:

$$Y = 1.050 \cdot 1.156^{x}$$

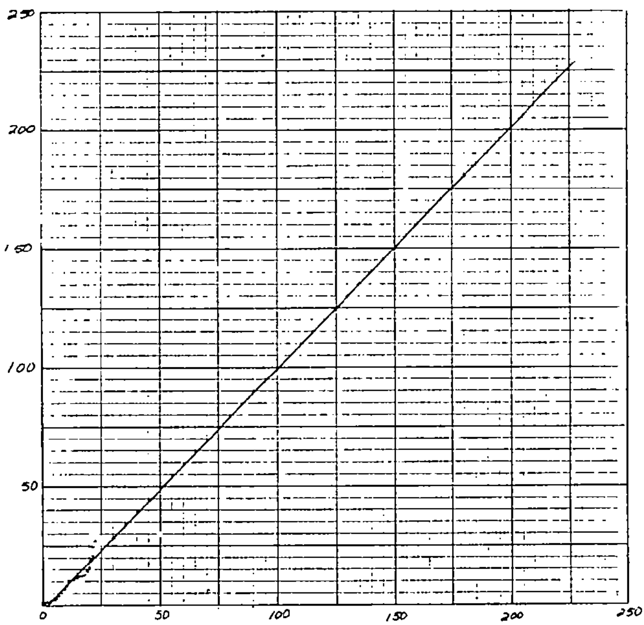
$R^2 = .938$

$s^2 = .07$

F - score = 332.7

Projection for x = 180:
y = 2.2574×10^{11}

Exhibit I-B.



Regression Equation:

$Y = -3.155 + 1.030x$

$R^2 = .934$

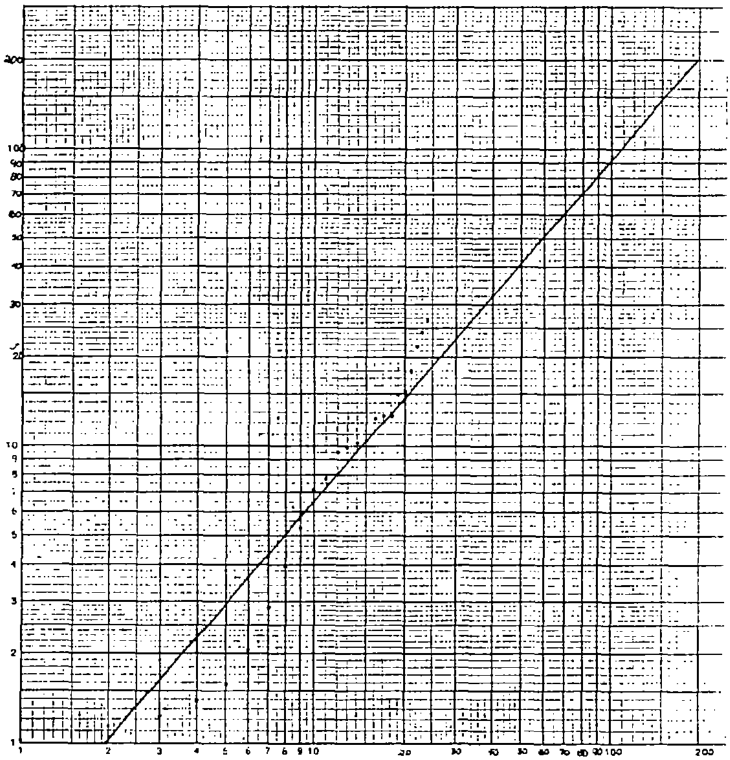
$s^2 = 3.92$

F - score = 311.7

Projection for x = 180:

y = 182.32

Exhibit I-C.



Regression Equation:

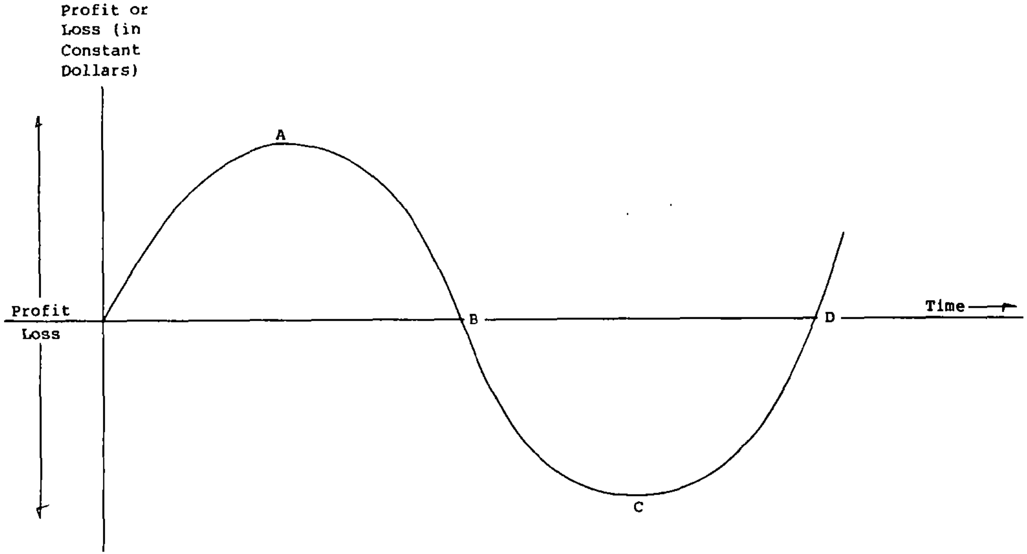
$Y = .467 \cdot x^{1.1475}$

$R^2 = .860$

$s^2 = .16$

F - score = 134.6

Projection for x = 180:
y = 180.82



A Maximum Profit

B Maximum Rate of
Rate Cutting at Time of
Highest Accumulated Profit

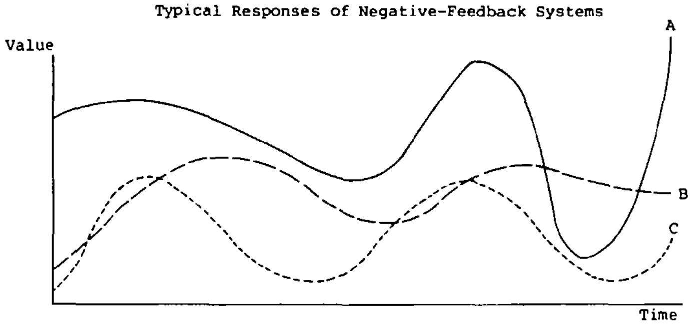
C Minimum Profit

D Maximum Rate of Increase
in Rates at Time of Lowest
Remaining Accumulated Profit

Exhibit II

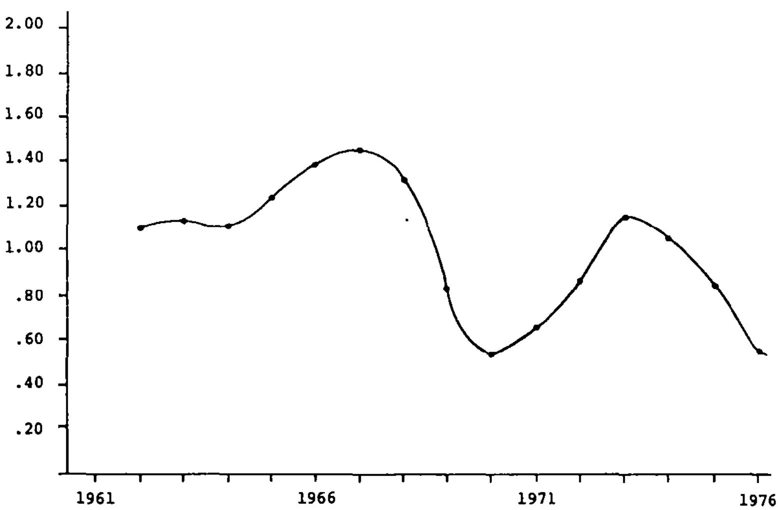
Exhibit III

Exhibit III-A.
Typical Responses of Negative-Feedback Systems



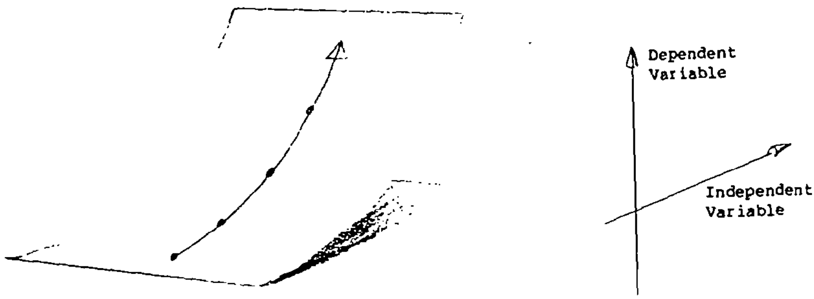
Cycles Can be Explosive (A), Damped (B) or Steady (C).

Exhibit III-B.
3-Year Average Loss Ratios for Medical Malpractice, 1961-1977
(One Carrier in One State)

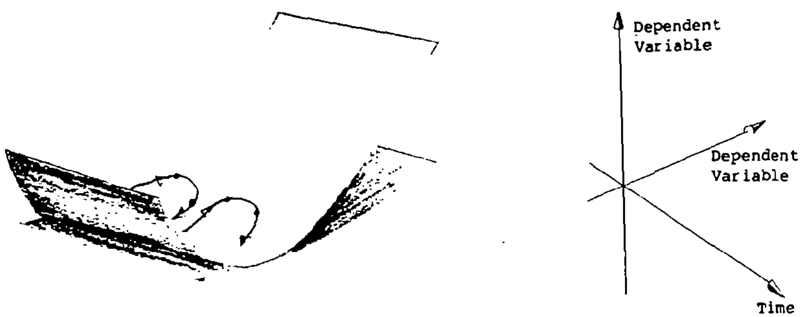


Three-year averages are shown because the insurer's underwriting policy did not change as often as annually. Also, the small volume of losses masks this pattern if individual years are considered.

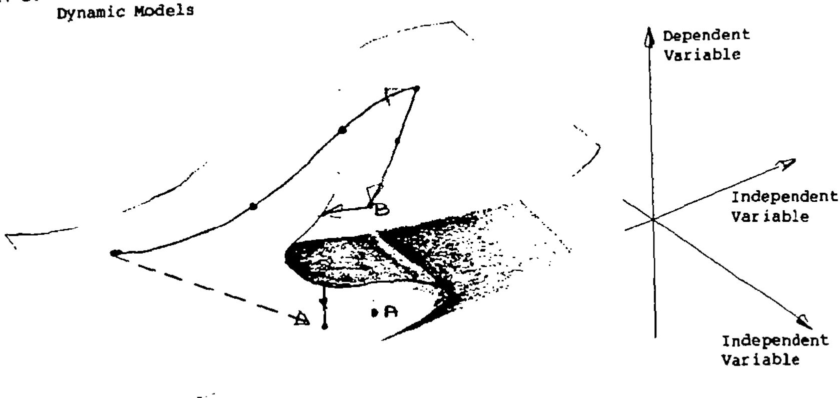
Exhibit IV



IV-A. Surface of Opportunities in Econometric Models



IV-B. Surface of Opportunities in System Dynamic Models



IV-C. Surface of Opportunities in Catastrophe Theory

REFERENCES

1. Robert J. Finger, "Modeling Loss Reserve Developments," PCAS LXIII, 1976, p. 90.

2. Jay W. Forrester, "Changing Economic Patterns," Technology Review, August/September, 1978.

3. Jay W. Forrester, "Counterintuitive Behavior of Social Systems," Technology Review, January, 1971.

4. Jay W. Forrester, Principals of Systems, Second Preliminary Edition, Wright-Allen Press, Inc., Cambridge, Mass., 1968. (Distributed by M.I.T. Press, Cambridge, Mass.)

5. Jay W. Forrester, World Dynamics, M.I.T. Press, Cambridge, Mass., 1971.

6. Jan A. Lommele and Robert W. Sturgis, "An Econometric Model of Workmen's Compensation," PCAS LXI, 1974, p. 170.

7. Nathaniel J. Mass, "Modeling Cycles in the National Economy," Technology Review, March/April, 1976.

8. Norton E. Masterson, "Economic Factors in Liability and Property Insurance Claim Costs 1935-1967," PCAS LV, 1968, p. 61.

9. Charles L. McClenahan, "A Mathematical Model for Loss Reserve Analysis," PCAS LXII, 1975, p. 134.

10. Arthur M. Okun, "Uses of Models for Policy Formulation," in The Brookings Model: Perspective and Recent Developments, Gary Fromm and Lawrence R. Klein, eds., American Elsevier Publishing Co., Inc., New York, 1975.

11. James B. Ramsey, Economic Forecasting - Models or Markets, Hobart Paper #74, Institute of Economic Affairs, London, 1977.

12. Paul A Samuelson, "The Art and Science of Macromodels Over 50 Years," in The Brookings Model: Perspective and Recent Developments, Gary Fromm and Lawrence R. Klein, eds., American Elsevier Publishing Co., Inc., New York, 1975.

13. Richard E. Stewart, "Profit, Time and Cycles," address to the C.A.S., May 21, 1979.

14. O. E. Van Slyke, "Catastrophe Theory," published in John B. Reid, Jr., "Fits and Starts," Best's Review, Life and Health edition, approximately February, 1977.

15. E. C. Zeeman, "Catastrophe Theory," Scientific American, April, 1976.

IS ECONOMETRIC MODELING OBSOLETE?

BY OAKLEY E. VAN SLYKE

REVIEWED BY MICHAEL FUSCO

Mr. Van Slyke's paper presents a discussion of econometric modeling in a fairly general way. I would have preferred to see more on possible specific applications to insurance pricing, especially with regard to the more sophisticated techniques of systems dynamics and catastrophe theory.

But one can hardly disagree with Mr. Van Slyke's conclusions; namely:

1. Econometrics is not obsolete.

2. The more sophisticated models of systems dynamics or catastrophe theory may be better than econometric models.

3. Models are just tools, to be used to enhance the forecasting process, not to replace it.

I would like to discuss these conclusions one at a time.

ECONOMETRIC MODELS

Mr. Van Slyke's definition of an econometric model is "a mathematical representation of economic relationships using linear equations." This is accurate and the equation he cites relating wages and medical costs to Bodily Injury Claim Costs is a good example of such a model.

However, later he refers to an example where the independent variable is related only to a measure of time. This, which we might recognize to be historical trend data fitted to a least-squares line, is too simple an example and I believe would not be considered by an econometrician to be an econometric model. The key element that is missing is an economic independent variable. Time is often a parameter of the equation because we hope to use the model to forecast a value for a certain time period, but time cannot stand alone as the economic independent variable.

Mr. Van Slyke cites several advantages to the use of econometric models, but neglects to cite disadvantages. His sub-heading "Disadvantages" should really be termed "Shortcomings of Econometric Modeling Techniques." Perhaps there are no real disadvantages, but I would hope one day an analysis could be performed to determine if the benefits derived from econometric models have been worth the cost and/or whether the laymen really finds a model to be a more understandable explanation for why his insurance rates are increased.

Mr. Van Slyke appropriately lists the reasons why models can produce poor forecasts - bad forecasts of independent variables, the index-number problem, wrong variables included, wrong equations assumed. There is no question that errors can and will occur; hopefully, by continually updating data and testing models, these errors can be minimized.

I can't resist relating what the defenders of the Consumer Price Index have said on the index-number problem. It is not a problem with the index for the index is exactly what it purports to be. It measures changes in cost of a fixed market basket of goods. Rather, it is a problem with those (actuaries, econometricians, etc.) who choose to misinterpret the index.

While the tone of Mr. Van Slyke's remarks seems to imply that he is going to conclude that econometric modeling is obsolete, he does not and cites it as a valuable tool. I am not surprised by this conclusion, nor do I disagree with it, but almost wish he had rendered it obsolete to see what reaction this would have generated within the Casualty Actuarial Society.

SYSTEMS DYNAMICS AND CATASTROPHE THEORY

Mr. Van Slyke concludes that these techniques may be better than econometric models. I can agree with this conclusion but put strong emphasis on the word "may."

The description of situations that lend themselves to the application of catastrophe theory was clearer to me than that given for systems dynamics. However, in neither case was I convinced that there is a real property/casualty pricing problem that can be solved through these techniques. Perhaps the reader can provide examples.

Also, I wonder if the dividing line between econometric modeling and systems dynamics is a clear one. Econometricians who are making predictions of underwriting results are generally starting with given loss and expense ratios and making various assumptions on the future changes in losses, expenses and premiums. However, the econometrician may use a statistical model of rate level changes based upon the loss and expense ratios of prior years. As a result, an interactive system is developed; underwriting ratios are used to predict rate level changes which are used to predict underwriting ratios, etc.

This interactive set of models is common in large econometric models. Would Mr. Van Slyke consider these types of models econometrics or systems dynamics?

Does the econometrician give any recognition to more or less restrictive underwriting patterns in his assumptions on loss changes? Does he include any consideration of a changing regulatory environment in his assumptions on price changes? Perhaps he does so more in a judgemental manner and less in a systematic manner than Mr. Van Slyke would want in order to call this systems dynamics. Nevertheless, I do not find the two techniques to be mutually exclusive.

It would be a worthwhile exercise to check the advantages Mr. Van Slyke listed for econometric models against systems dynamics and catastrophe theory to see if they still apply. Credibility to the laymen seems to be a tougher one to justify. Otherwise, we are forced to accept Mr. Van Slyke's conclusion on faith alone. I am willing to accept the surface area configurations on Exhibit IV, but have not accepted that there exists a property/casualty insurance product that looks like figure IV-C.

A TOOL TO ENHANCE FORECASTING PROCESS

I am in complete agreement with Mr. Van Slyke in this area.

He hopes that whatever models are used, they be instructive as well as predictive in nature. At times, plugging different assumptions into our equations tells us something new. The range of predictions can often reveal how sensitive our dependent variable is to a particular independent variable, of course assuming that our model itself is reasonably accurate.

But as Mr. Van Slyke notes, the model itself often must be adjusted. This is not to say the model is bad, but rather that it is an imperfect tool. In an econometric model ISO developed, the number of small cars on the road was used to reflect the magnitude of a collision. The first measure of this we adopted was an imported car ratio. However, as the number of small domestic cars increased, this measure became inappropriate and we switched to a measure of compact cars. This measure, too, was eventually discarded as the definition of compact cars changed.

The fact that models must be adjusted is not a reason for rejecting modeling techniques. Rather, it points out the need for careful construction of models and monitoring of their effectiveness. No one is suggesting that the model be used without the application of judgement.

CONCLUSION

In conclusion, I enjoyed reading Mr. Van Slyke's paper and recommend it to anyone wanting a description of econometric modeling techniques. His contribution to the CAS literature should motivate readers to delve into specific pricing applications.

TITLE: RISK CLASSIFICATION STANDARDS

AUTHOR: Mr. Michael A. Walters

Mr. Walters is Senior Vice President and Actuary for Insurance Services Office. He received his B. S. degree from Fordham University, his M. S. degree from the University of Notre Dame and his FCAS in 1972. Mike is Treasurer of the CAS and is on the CAS Board of Directors. He has served on the Education Committee (Chairman), Finance Committee, Committee on Government Statistics and Joint Committee on Review of Education and Examination. He has authored a paper entitled "Homeowners Insurance Ratemaking" which appeared in the 1974 PCAS. Mike is also a member of the American Academy of Actuaries.

REVIEWER: Mr. Robert A. Bailey

Mr. Bailey is an Actuary for the National Association of Insurance Commissioners. He has earned his B.S. and M.S. degrees in Actuarial Science from State University of Iowa. Bob received his FCAS in 1955 and is a member of the American Academy of Actuaries. Perhaps no one has authored more papers or reviews of papers for the CAS than Mr. Bailey. Over the years, he has written numerous articles dealing primarily with insurance finance and pricing. Bob has also served on many CAS committees including the Board of Directors of both the CAS and the American Academy of Actuaries.

INTRODUCTION

The escalating inflation of the past decade spawned complaints about more than just overall insurance rate increases. Unlike most other products, insurance costs depend upon buyer characteristics, so questions of fairness have naturally arisen as some insureds were confronted with four digit auto insurance prices along with double digit inflation. "Affordability", "availability", and "social acceptability" all became buzz-words of the late seventies.
In particular, regulators, legislators, and other consumer advocates have focussed increasing concern on the third requisite of virtually every state's mandate on insurance rates, that they "not be unfairly discriminatory".

Some critics have claimed that insurance rating methods, and classifications specifically, should be sensitive to consumer perceptions about what is fair. They suggest that classifications possess qualities of reliability, causality, controllability, separation and incentive value. Some of these proposals might be essential to the insurance process, while others may be merely sound business advice, and still others might only be consumerist rhetoric.

A search through insurance and actuarial literature does not find an abundance of historical resource material relevant to, or in the language of, these current issues. Some of the more persuasive reformers have, in fact, coined new phrases and fashioned new literature as the basis for change. From a social standpoint, some of the espoused changes may be genuine attempts to solve affordability problems in what is intended to be a "fair" manner, but if the resulting mechanism violates the principles of insurance, it is not an insurance program. Therefore, it might not be under the jurisdiction of a state's insurance regulation.

A recent insurance monograph by Professor John Long elaborates on the problem.[1]

> "It is fashionable to be critical of insurance theory and to blame the ills of the insurance marketplace on the shortcomings of insurance theory. For example, one point of view is that the purpose of the insurance industry is to serve the needs of the public and that any inability of the industry to do so means that something is wrong with the underlying insurance theory... This point of view is questionable, whatever concept one holds about the nature of theory...
>
> "A case in point has to do with exposure to flood loss... The Congress has seen fit to provide a subsidy to eligible people who participate in what is called the national flood insurance program. This program raises the question of how much 'non-fortuitous' transfer of funds can occur in a transaction without causing the transaction to be something other than insurance... In the author's judgment, the federal flood program exceeds such limit and, therefore, is a type of welfare rather than a type of insurance. This classification is not to imply that because the flood program is not insurance it is 'bad'. The only point being made is that the subsidy for all participants by the taxpayers as a whole is so large that the arrangement is not insurance. Calling something insurance does not necessarily imbue it with the characteristics associated with insurance."

It is important therefore to distinguish those qualities which some would like to see an insurance classifiction system possess to achieve alternative goals, from those which are necessary and sufficient conditions, or standards, which flow from the nature of insurance. The purpose of this paper is to develop a set of these standards for insurance classifications, which have been implicitly used, or should be used, to evaluate compliance with insurance statutes.

NATURE OF INSURANCE

The purpose of insurance is to protect an insured from a large and fortuitous financial loss. It is achieved by contractually transferring the insured's uncertainty of loss to the insurer for the certainty of a smaller payment called the premium. This uncertainty of loss is called risk.

[1] John D. Long, "Soft Spots in Insurance Theory", Issues in Insurance, Vol. II, 1978, P. 444.

Since the insurer assumes the individual insured's risk of loss, the premium should be fundamentally based upon the expected value of the insured's loss. The expected loss for an insured is the probability of his having an accident or a claim times the average cost of that claim. The premium should also include the expense of servicing the policy plus a margin for profit and contingency as a reward for the risk taking. The amount of this profit margin should depend upon two basic factors, the ability of the insurer to estimate the expected (or average) loss of the individuals insured;[2] and, second, the amount of overall reduction of uncertainty accomplished by the pooling process.

Insurers are not, of course, trying to predict the actual losses of each insured, only the expected loss. It is the variation of an individual's actual losses from his expected loss that motivates his purchase of insurance, while the variation of expected losses from individual to individual that motivates insurers to price insureds differently.

Although from an insured's standpoint, the essence of insurance is the transfer of risk, a further value of insurance for society is the reduction of overall risk or uncertainty by pooling many insureds independently exposed to loss.

Now, these risks in the pool do not have to be exactly the same types of risks for insurance to work, as witnessed by the success of Lloyd's of London, with a multiplicity of risks no two of whom may have been the same over the years. And certainly, insureds who are inherently different risks should not have to pay the same for the insurance

[2] There is obviously more risk involved to the insurer than distinguishing one insured from another. The uncertainty of next year's inflation level, for example, affects the expected cost of individuals, but more or less to the same degree.

process to work. But pooling works especially well within a given line of insurance, like private passenger auto insurance, when enough independent risks are pooled such that it is virtually impossible that they all will have accidents in the same year. In fact, the more risks that are written, the closer reality comes to the expected. This intuitively expresses the "law of large numbers".[3] Its first and perhaps best known application allows insurers to have more confidence that, once each risk has been reasonably priced, the actual losses on all those risks combined or pooled will come reasonably close to the combined expected losses at the end of the year.

This does not say that the pooling of risks is the same as pooling of losses. This latter term somehow may connote that everyone should share the costs equally. Insurance can work just as well even if every risk had a different expected loss, as long as you can reasonably estimate the expected losses.

Likewise insurance does not require that each classification must be large enough to stand on its own. This fallacy says that individual classes cannot share the risk among other classes.[4] It would also deny the ability to summarize across classes to gain additional information about other classes, such as pooling classification information within territory to determine territory rates, or territories within state to determine statewide rate levels.

[3] See D. B. Houston, "Risk, Insurance, and Sampling", The Journal of Risk and Insurance, XXXI No. 4, 526-530.

[4] See Stanford Research Institute, The Role of Risk Classifications in Property and Casualty Insurance, Final Report, May 1976, p. 63: "Confusion surrounding the term 'classification' stems also from an association with the concept of pooling of risks to reduce the aggregate risk. Many people feel that the essence of classification lies in having large classes, the members of which share the total risk of the class (and supposedly do not share the risk of any other class). According to this incorrect view, classes must each have many members to pool risks; classes with too few members are therefore not 'credible' and are assumed to violate the basic principle of risk sharing."

Furthermore, some may believe that insurance is an instrument of social policy to compensate victims. This view treats the premiums as merely a means of accumulating funds to pay out losses in ways possibly fundamentally different from the relative risk that each insured presents to the pool. But trying to do something noble via the premium collection facilities of insurers does not make the resultant mechanism insurance, (as earlier cited). Insurance is what it is - the transfer and reduction of risk; it is not a tax to redistribute wealth.

Thus, the expected loss of the individual is important to the pricing of insurance. But, being inherently unknowable, even by the insured himself, how do insurers infer this vital quantity? There are three basic methods.

First, they may use wisdom and experience as an underwriter in exercising informed judgment about the nature of the insured and the exposure to loss and attendant hazards. This is not the most accurate method but it is sometimes the only one available. From an insured's standpoint, with a complicated risk desired to be transferred, as long as both parties agree on a price, the insurance mechanism is working.

The second method of inferring individual expected loss is to observe the insured's actual losses over a long period of time. This gains certain additional information, picking up more of the subtleties of the risk that could not be obtained by logical, informed judgment. (This is analogous to experience rating versus schedule rating.) However, once obtained this information may be outdated, as the risk to be insured next year may have changed substantially. Furthermore, depending on the frequency of accidents, it may take twenty to thirty years of observation to infer correctly, given the dominance of randomness in the accident occurrence process.

The third method of inferring expected losses is to observe the experience of a group of similar risks over a much shorter and more recent period of time. These groups of similar risks are called classifications. Furthermore, the group observation process also involves the second use of the Law of Large Numbers. The first use was that if you know the expected losses in advance, then the actual losses will tend to approximate the expected at the end of the year for the insurance enterprise as a whole.

However, by observing a smaller number of similar risks over a short period of time you have more confidence you have closely estimated the expected losses of the individuals in advance. This is especially important if the set of insureds can change from one year to the next. (This process of classifying is analogous to using stratified random sampling to gain more information when the size of the total sample is limited.[5])

There are some who feel that group inference for an individual member of a group is unfair per se, no matter how the groups are defined. This would seem to prohibit the use of any statistical based knowledge throughout society, and is contradicted by all

5 Houston, p. 534. Author's Note: If the classes are fairly stable over time, they do not even need to have similar expected losses for the individuals within in order to gain a good estimate of the class average expected losses. Merely the variance of actual losses from the mean for each individual insured in the class should be similar. This results from the fact that insurance classification reviews use all the risks insured in each calss.

insurance statutes which allow, or even mandate, the use of classifications. The SRI also clearly addressed this:

> "...the opinion that distinctions based on sex, or any other group variable, necessarily violate individual rights reflects ignorance of the basic rules of logical inference in that it would arbitrarily forbid the use of relevant information. It would be equally fallacious to reject a classification system based on socially acceptable variables because the results appear discriminatory. For example, a classification system may be built on use of car, mileage, merit rating, and other variables, excluding sex. However, when verifying the average rates according to sex one may discover significant differences between males and females. Refusing to allow such differences would be attempting to distort reality by choosing to be selectively blind."[6]

CLASSIFICATION STANDARDS

So insurance classifications are seen as needed in the pricing of many kinds of insurance, helping to reduce overall risk, as well as enabling insureds to pay in proportion to their relative hazard of loss. If there were no reflection of these relative costs by an insurer, it could risk insolvency if the distribution of exposures changed substantially. At a minimum, such an insurer will require a larger margin for profit and contingency to offset the much greater chance of adverse underwriting results.

At this point, it is important to distinguish risk classification from risk selection. Risk selection determines both the general (via marketing) and a more specific (via underwriting) set of insureds with whom the insurer decides to enter into a contractual relationship [7] and whom the classification system must price according to its predetermined criteria.

[6] SRI (1976), p. 91.

[7] In some lines and states, a shared (or so called involuntary) market exists which requires participation by insurers in order to write voluntary business. This helps solve an availability problem for those not "selected" by insurers under usual markets.

Given the preceding, the variables comprising a classification system should be chosen so that a set of general standards or conditions are met (in addition, of course, to any expressed statutory requirements regarding fair discrimination).

1) Similar risks should be assigned to the same class with respect to each variable; coversely, dissimilar risks should be assigned to different classes, so that there are no clearly identifiable subsets with a significantly different loss potential or expected loss in the same class.[8]

2) The common characteristics used to identify insureds as similar should reasonably relate to the potential for, or hazard of, loss.[9]

3) Next, the classes should be exhaustive and mutually exclusive; that is, an individual should belong to at least one, but only one class with respect to each rating variable.

4) There should be clear and objective phraseology in the definition of classes, so that there exists no ambiguity as to what class an individual insured belongs.

5) An insured should not be able to easily misrepresent or manipulate his classification.

[8] It is important to stress the words "clearly identifiable" when dealing with the alleged overlap or heterogeneity of certain classes.

[9] This is different from, and yet related to, what some others have used as the notion of causality, and will be covered in the section on Non-Standards.

6) The cost of administering a rating variable should be reasonable in relation to the benefits received.

7) And finally, to the extent possible, the class rating factors should be susceptible to measurement by actual insurance data.

These seven standards actually fall into three broader categories which can describe a set of necessary and sufficient conditions for insurance classifications: i.e., homogeneous, well-defined, and practical.

Homogeneous

More homogeneous classes will take fewer risks to obtain reasonable estimates of expected costs, and will minimize the ability of competition to skim off better than average risks thus changing the ultimate costs.

The "reasonable relationship" standard is also a way of avoiding spurious measures which likely have potentially identifiable subsets. Of course, if a strong statistical correlation persists over time, with no emergence of practical subdivisions, then the degree of perceived reasonableness may be enhanced over time, as well.

Homogeneity is also undergoing some current debate as to the possibility of statistical measurement.[10] While the scope of this paper precludes entering that debate, it is helpful to recall that one of the reasons for classification is the impossiblility of knowing a risk's true expected loss or accident likelihood. Given

[10] See Richard G. Woll, "A Study of Risk Assessment", PCAS LXVI, 1979.

the randomness of accident and loss occureence, and the fact that statistical tests must use actual loss distributions for individuals, it may be difficult to gain more than a glimpse or an insight into possible distributions of accident likelihoods within class. This is especially true since assumptions must also be made about the functional form of the accident likelihood model (as well as of the loss severity model).

Furthermore, the real test of homogeneity is in the most refined classification cell, not in the separate variables used in combination to classify the risk. It is also not necessary (or even likely) for a classification to have identical expected losses for all risks within the class, even if true individual risk accident likelihood were "knowable". Finally, even if inferences can be made about a possible distribution of expected losses within a classification, the lower expected loss insureds deduced to exist are not in any way identified (or identifiable) to the insurer or even known by the risks themselves. Therefore, it is bordering on a philosphical game to assert that such a class is too heterogeneous, and is therefore not permissible.

The SRI spoke to that fallacy as follows:

> "Indeed, the rationale that proscribing the use of certain rating variables is in the public interest because, under imperfect risk assessment systems, actuarial fairness is not achieved for some -- albeit unidentifiable - individuals is fundamentally contradictory. It promotes a remedy for unfairness to some that increases the unfairness overall (by the same actuarial yardstick) and redistributes it."[11]

[11] SRI International, Choice of a Regulatory Environment for Automobile Insurance, May 1979, p. 58.

Well-Defined

The second broad standard is that of being well-defined, and helps to ensure that each risk is actually placed in the right classification and to avoid unequal application of the classification system. The "exhaustive" quality allows more risks to be accepted and, once accepted, gives a complete method of rating them. "Exclusivity" precludes two different rates for the exact same risk. "No ambiguity" also prevents unequal treatment of the same risk, while protection from misrepresentation by insureds will keep the statistical data consistent as well as enhancing the equal treatment of insureds.

Practical

The dictionary definition of practical refers to "workable, useable, and sensible" and the final two standards deal with these goals. Being cost-effective is important because an inefficient system (or even attempts to be too precise) could increase total costs beyond the value of the information to be obtained. If, for example, it costs an insurer ten dollars on each policy to find only a small portion of risks who could save twenty dollars, it is not worth the effort.

In final perspective, one of the advantages of classifying was to use the Law of Large Numbers on actual observed experience of the past instead of relying on pure business judgment. If there is no method or attempt to test class average prices by actual data, the system is tantamount to schedule rating. Of course, whether or not a classification rating factor is tested frequently depends upon the likelihood of change in a short period of time, and the relative size and importance of the rating factor.

NON-STANDARDS

In this paper, the word "standards" has been used to denote a set of necessary and sufficient conditions for insurance classifications, consistent with the nature of insurance as well as insurance statutes. However, the dictionary definition also includes "a basis of comparison in measuring or judging... quality." It is possible or indeed likely that other characteristics of classification may be desirable. Failure to include these in the basic standards means that it is felt that their presence is not required to render the classification system valid and appropriate.

Two different qualities that have been recently espoused are actually correlatives - <u>controllability</u> and <u>incentive value</u>. By controllability is meant the ability of an insured to determine by his own efforts (presumably consciously) the class to which he is assigned. If that quality is present, it is argued, the insured will have the incentive to change to a lower rated class and thus reduce his own losses as well as the losses of the overall system.

One can sympathize with a risk that presents a much higher hazard, over which it has little or no control, but to deny use of that criterion, and make others with lower inherent risk subsidize the higher risk is, in effect, a denial of reality. In workers' compensation insurance, for example, the logging or lumbering industry has an inherently higher risk of injury to workers than clerical office type work. Not to charge for that difference would be to contradict the essence of classification. Similarly, age in life insurance is an essential classification, yet is obviously uncontrollable. Controllability therefore is an extraneous add-on, which has benefits primarily in the area of public understanding.

Incentive value also has public appeal, and in its obverse may be important to the overall insuring process. Whether it be classifications or exposure base, or indeed the existence of insurance, the presence of an insurance contract should not encourage a laxity towards loss control or create a moral hazard of exaggerated or false claims.[12]

While incentive value could be a noble addition to a rating system, it is not a necessary one, nor should classification plans be judged by it as a standard. Personal lines risks, for example, cannot be easily subjected to loss prevention measures like large commercial risks. Even so-called "merit rating" in automobile insurance may be nothing more than a theoretical incentive to prevent accidents. Few drivers wear seat belts despite the life saving evidence, so the prospect of saving a few dollars of insurance surcharge certainly will not induce the modification of driving behavior. In a DOT Study, a major conclusion in this area was also reached: "As long as deterrent measures concentrate on a punitive approach to the correction of 'driver error,' they are likely to remain relatively ineffective."[13] (Of course, once an accident occurs, the fear of a surcharge may affect the reporting of accidents and submission of collision claims, but that may be in conflict with the liability insurance policy "condition" requiring notification of accidents).

[12] C.A. Williams et.al., Principles of Risk Management and Insurance, Vol. I, 1978, p. 128.

[13] U.S. Department of Transportation, Causality, Culpability and Deterrence in Highway Crashes, 1970, p. 245.

Causality is also recently cited as a desired quality for classifications to possess, defined as follows: "the actual or implied behavioral relationship between a particular rating factor and loss potential."[14] The use of the term "behavioral" makes this difficult to accept as a standard, because living in the river valley does not cause the river to flood, yet certainly increases the hazard involved in flood insurance.

Merit rating in auto insurance is almost totally non-causal. The fact that an insured has been involved in a past accident does not behaviorally cause him to get in the next one or even to have become a worse driver. And yet the same critics of current rating cite past accident record as an ideal rating variable.

Instead, a reasonable relationship to the hazard of loss, without such a rigid chain of causality or behavior, is more appropriate. As the earlier mentioned DOT Study concluded: "...driver responsibility for crashes is rarely unilateral and is often impossible to isolate from the multiplicity of causes involved in almost every crash."[15]

By classifying risks, an insurer does not seek to determine the cause of the accidents. To the extent high risk insureds are identified, society may benefit by focussing attention on the need for possible remedies.

[14] "Final Report of the Rates and Rating Procedures Task Force" of the (NAIC) Automobile Insurance (D-3) Subcommittee, November 1978, p. 5.

[15] DOT, p. 209.

Separation has been defined as "a measure of whether classes are sufficiently different in their expected losses to warrant the setting of different premium rates."[16] This deals with the so-called "overlap" question where it is felt that if one class rate were close to another, some insureds in the first class would have accident likelihoods close to those in the second class, and therefore may be mis-classified.

This is related to the homogeneity question. If the insureds who supposedly deserve to be in the second class are not identifiable, then it is questionable whether you can call them mis-classified. Secondly, classifications with mean rates close together are not undesirable, if the hazard being reflected is a gradual one. Finally, even if some insureds in a $300 rated class truly deserve to be in a $305 class, the system is still working well from a cost/benefit standpoint. Therefore, the concept of separation does not appear very useful in the context of classification standards.

Reliability has also been a term which includes qualities that are objective, clearly defined, and easy to verify,[17] all of which are consistent with the standards earlier mentioned, and about which there is little or no controversy.

However, social acceptability and admissibility are terms which connote a variety of meanings and contexts regarding the use of insurance classifications. By way of perspective, it is one thing to give advice as to the public's view of certain rating variables among alternatives of equal value. It is quite something else

[16] Division of Insurance, Commonwealth of Massachusetts, Automobile Insurance Risk Classification: Equity and Accuracy, 1978, p. 3.

[17] Massachusetts, p. 3.

to say that the unpopularity of some variables, as perceived subjectively by some, or even through public opinion polls, pre their use. Rate adequacy and public acceptability are often in conflict.

The earlier cited SRI Report suggested that insurers choose vari among the set of possible ones, without loss of precision, that a clearly explainable to the public, provide incentives for loss prevention, and are adjusted to social mores.[18] That this was meant as sound business advice, rather than a set of necessary conditions, is illustrated by their comments on the very next page:

> "On the other hand, the opinion that distinctions based on sex, or any other group variable, necessarily violate individual rights reflects ingnorance of the basic rules of logical inference in that it would arbitrarily forbid the use of relevant information. It would be equally fallacious to reject a classification system based on socially acceptable variables because the results appear discriminatory. For example, a classification system may be built on use of car, mileage, merit rating, and other variables, excluding sex. However, when verifying the average rates according to sex one may discover significant differences between males and females. Refusing to allow such differences would be attempting to distort reality by choosing to be selectively blind.
>
> "The use of rating territories is a case in point. Geographical divisions, however designed, are often correlated with sociodemographic factors such as income level and race because of natural aggregation or forced segregation according to these factors. Again we conclude that insurance companies should be free to delineate territories and assess territorial differences as well as they can. At the same time, insurance companies should recognize that it is in their best interest to be objective and use clearly relevant factors to define territories lest they be accused of invidious discrimination by the public".[19]

Moreover, in a later work, the SRI clearly stated: "The regulator's determination of what is unfairly discriminatory should relate only to the use of variables whose predictive validity cannot be

18 SRI Report, 1976, pp. 89-90.
19 SRI Report, 1976, p. 91.

substantiated and to unequal application of a classification system."[20] Furthermore, they put the context of extreme social intolerability in the legislative arena:

> "One possible standard does exist for exception to the counsel that particular rating variables should not be proscribed. What we have called 'equal treatment' standard of fairness may precipitate a societal decision that the process of differentiating among individuals on the basis of certain variables is discriminatory and intolerable. This type of decision should be made on a specific, statutory basis. Once taken, it must be adhered to in private and public transactions alike and enforced by the insurance regulator. This is, in effect, a standard for conduct that by design transcends and preempts economic considerations. Because it is not applied without economic cost, however, insurance regulators and the industry should participate in and inform legislative deliberations that would ban the use of particular rating variables as discriminatory."[21]

Admissibility, as per the Massachusetts definition, begins with federal and state statutory requirements regarding discrimination and privacy, but continues in the social acceptability vein:

> "There are also distinctions that, while not clearly illegal, are being increasingly questioned. These include sex, income, and marital status. Clearly, it is preferable to avoid such distinctions. Distinctions are best able to meet the text of admissibility if they are within an individual's ability to control and are causally related to the probability of loss. It would be undesirable, for example, to charge higher rates for redheads than brunettes even if it could be shown statistically that people with red hair have more accidents than those with brown hair."[22]

Use of the words "preference" and "desirability", from a perception of the public's view and using popular intuition about controllability and causality, again confirms that this characteristic is in the form of business marketplace advice. Insurers who can combine

[20] SRI International, Choice of a Regulatory Environment for Automobile Insurance, May 1979, p. 93.

[21] SRI, 1979, p. 94.

[22] Massachusetts, 1978, p. 4.

sound and relevant rating variables with the public's view of what is better will obviously be more successful. However, unless or until possible substitute variables are found which do not sacrifice accuracy and do not create subsidies, the failure to use appropriate, though unpopular, variables will only cause some individuals availability problems and still others to be overcharged relative to their risk.

REGULATION VERSUS COMPETITION

Given that insurance regulators must enforce the rate regulatory laws, a logical question to be asked is whether natural competitive forces will reinforce or conflict with the standards for insurance classifications.

Regarding homogeneity, it is obvious that the essence of competition will be to try to find rateable subsets of existing classifications to price more accurately and equitably (prices matching costs).

If classes are too broad, underwriters will tend to select risks out. However, it takes more discipline to define objective and practical new classifications to maximize the number of risks to be written voluntarily. If several different companies are licensed in a group under the same management control, the competitive drive for more homogeneity can be partially met by a different set of underwriting standards for each company in the group.

If there is only a strong statistical correlation for a particular variable, without an obvious relationship to hazard of loss, competitive forces will definitely strive to find a closer link. If no closer link is found over an extended period of time, as mentioned earlier, the reasonableness of the relationship becomes much more established.

There is an analogy here with the statistical correlation between lung cancer and cigarette smoking which for many years was not held to be a health hazard. In fact, there has yet to be found in human medicine a cause and effect link showing lung cancer resulting from tobacco smoking. Conceivably (but unlikely), cigarette smokers could have other characteristics related with carcinogens that are also less prevalent in non-smokers. The answer, of course, is not to avoid the use of statistical information until better data is found. Indeed, the U.S. Surgeon General and others have taken strong steps based mainly (and reasonably) on the statistical evidence. Even though the actual risk of death from lung cancer among the heaviest smokers is very small, it is many times that of non-smokers. Stated another way, most heavy smokers will not contract lung cancer; yet all of them have had certain privileges revoked and rights modified.

One can normally expect marketplace rewards for those who use well-defined class plans allowing equal treatment for all risks. However, there is a temptation to allow some ambiguity or subjectivity as a trade-off for additional costs needed to gain consistent information.

Regarding practicality, competitive forces will place a natural restraint on overspending to attain rating information. However, part of the workability of classifications involves testing the rating factors with actual data to minimize the subjectivity of pricing. There is a potentially conflicting instinct, however, to rely on judgment and assumptions to avoid the cost of truly testing for the appropriate price relationships. Of course, to the extent that other insurers find cost-effective ways of better measuring class relativities, then as long as there is the ability to exchange information, any pricing inequities will be short term.

Some examples of potentially unfair discrimination in insurance classifications might include the following:

The use of occupation as a rating variable for auto liability insurance may be a problem with regard to ambiguity in splitting the population into exhaustive categories, as well as not all cells likely being reasonably related to the hazard of loss.

Similarly, national origin (if not already proscribed by law) would have problems with the mutually exclusive and exhaustive categories.

Use of unverifiable criteria or too subjective wording, such as with psychological profiles would also present major problems. The use of characteristics which are easily circumvented by some insureds and not others can favor the pricing of some to the detriment of others.

Another example of possible unfair discrimination would be the failure to reflect premium differences for identifiable and rateably different subsets of broader classifications, unless some overriding reason existed such as the cost of determining the necessary information being too high for the overall system.

The pricing impact of not subdividing depends upon the size of the subsets and the resulting differences in price for each of the subclasses. It may be that only a small amount of premium can be saved by refinement, if one of the subclasses is very large and also the lowest priced (such as rating by past accident record in auto insurance where accident-free or claim-free drivers usually save at most five percent over the cost of not having such a program).

If, however, lower risk insureds were identified in a system and were useable as a rating classification, the failure to reflect those differences would constitute a subsidy. But, if the set of insureds are not identifiable in advance, then there is no subsidy. For example, some have alleged that all of insurance is a subsidy since, as the reasoning goes, those who do not have accidents are subsidizing those who do. This is fallacious because you cannot identify in advance those who will have accidents. That is why people buy insurance. However, you can identify those with a higher likelihood of an accident which is what classification is all about. Failure to classify would therefore be a subsidy by those with a lower loss likelihood of those with higher loss expectancy.

Some also allege that it is a cruel disservice to identify the high risk insureds in advance through refined classification plans. However, insurers should not be blamed for the existence of high risks in society. In a report from the Federal Trade Commission to the U.S. Department of Transportation in 1970, it concluded: "Regardless of law and underwriting systems, high risk drivers exist. The present system identifies them; it does not create them." [23] In fact what insurers do by keeping track of the sources of accidents is to help identify those segments of the population

[23] Report of the Division of Industry Analysis, Bureau of Economics, Federal Trade Commission to the Department of Transportation, Price Variablility in the Automobile Insurance Market, August 1970, p. 144.

when loss prevention may be the answer rather than risk pooling. "In the interests of loss control and prevention, this high-risk group must be identified and treated before the accidents occur."[24] In other words, if high risk driving in high density areas produces an inordinate amount of loss, perhaps more stringent licensing should be considered, or mass transportation improvements, or other alternatives, but do not hide the information. Until such time as the source of the problem is solved, to paraphrase the SRI Report on Risk Classification, society should not legislate against the use of knowledge in a free society. [25]

SUMMARY

The purpose of this paper was to view the issue of reasonable classifications from the perspective of the nature of insurance itself. In this way perhaps the qualities that many have felt classifications ought to possess could be distinguished between the essential and the non-essential.

Much has been written in the past few years about what is fair or unfair, but this evaluation should not take place without an understanding of what classifications are designed to do in insurance. Affordability is one example of a quality which society might like insurance rates to have, but the essence of classifications serves to highlight high-risk, high-cost segments of the population. Unfortunately in that instance and in possibly others the solution to the problem may lie outside the scope of insurance classifications or even the insurance mechanism itself.

[24] DOT, p. 144.

[25] SRI, The Role of Risk Classification in Property and Casualty Insurance, 1976, Executive Summary Report, p. 25.

REFERENCES

Division of Insurance, Commonwealth of Massachusetts. Automobile Insurance Risk Classification: Equity and Accuracy, 1978.

"Final Report of the Rates and Rating Procedures Task Force" to the (NAIC) Automobile (D-3) Subcommittee, November 1978.

Houston, D.B. "Risk, Insurance, and Sampling." The Journal of Risk and Insurance, XXXI No. 4, pp. 526-530.

Long, John D. "Soft Spots in Insurance Theory." Issues in Insurance, Vol. II. Malverne, Pennsylvania: American Institute for Property and Liability Underwriters, 1978.

Report of the Division of Industry Analysis, Bureau of Economics, Federal Trade Commission to the Department of Transportation. Price Variability in the Automobile Insurance Market, August 1970.

SRI International. Choice of a Regulatory Environment for Automobile Insurance, May 1979.

Stanford Research Institute. The Role of Risk Classification in Property and Casualty Insurance. Final Report, May 1976.

U.S. Department of Transportation. Causality, Culpability and Deterrence in Highway Crashes, 1970.

Williams, C.A., Head, G.L., Glendenning, G. William. Principles of Risk Management and Insurance, Vol. I, Malverne, Pennsylvania: American Institute for Property and Liability Underwriters, 1978.

Woll, Richard G. "A Study of Risk Assessment." PCAS LXVI, 1979.

RISK CLASSIFICATION STANDARDS

By Michael A. Walters

Reviewed by Robert A. Bailey

SUMMARY OF THE PAPER

The author discusses the current issues related to risk classification in insurance. He distinguishes those qualities which flow from the nature of insurance, from those that attempt to redistribute wealth via a subsidy. From a description of the nature of insurance he derives a set of necessary and sufficient standards for insurance classifications which he summarizes into three broad categories: homogeneous, well-defined, and practical.

The author also discusses other characteristics which may be desirable but which are not necessary to make a classification system valid and appropriate. These include controllability, incentive value, causality, separation, social acceptability, and admissibility.

The author concludes with a discussion of regulation versus competition. He suggests that natural competition will enforce the standards he defined as necessary and that regulation, to the extent it interferes with industry pricing practices, would create a subsidy.

SUMMARY OF THE REVIEW

Insurance pricing is a combination of rating classification and underwriting selection. The author defines necessary and sufficient standards for rating classification which are not met by underwriting selection or insurance pricing as a whole. It is unreasonable to assert that it is necessary for rating classifications to meet certain standards when the pricing structure as a whole does not meet those same standards. The author's acknowledgment that pricing standards are necessary and appropriate, and his focus on only that portion of the pricing structure which the industry chooses to make visible, while ignoring the entire pricing structure, illustrates the inconsistent standards practiced by the insurance industry, and indicates why many believe that governmental intervention into the insurance pricing structure is necessary and appropriate.

SUBSIDY

The author suggests that regulators should not act in a manner so as to turn the insurance mechanism into a subsidy program. By defining what constitutes a subsidy in a manner justifying industry practices, the unsurprising conclusion is that regulators should keep their hands off. The author's definition of "subsidy" lacks a certain degree of persuasiveness and clarity.

His definition of subsidy is: "If, however, lower risk insureds were identified in a system and were useable as a rating classification, the failure to reflect those differences would constitute a subsidy." (page 709) This definition depends on which lower risk insureds are identified by a system, without qualifications as to whether the system is truly competitive or truly efficient. None of our present systems are ideally competitive, ideally efficient, or free of manipulation and control. If a system is controlled in some way then those who control it are in a position, under this definition, to define as subsidy anything that interferes with their objectives. Two examples of a controlled system that might not have identified the low risk insureds to the same degree as a more competitive system are the systems in effect (1) when the industry conspired together to reduce competition among themselves in the South Eastern Underwriters era and used a simple uniform classification system, and (2) when Massachusetts set uniform classifications for all insurers. Accordingly it is not surprising that the author, using this definition, concludes that regulatory interference with industry practices would create a subsidy.

Inasmuch as the author acknowledges "the impossibility of knowing a risk's true expected loss" (page 697) and that it is "not necessary (or even likely) for a classification to have identical expected losses for all risks within the class," (page 698) it appears that subsidy is always present to some degree and the question is not an objective "Is there subsidy?" but an inquisition, "Who caused the subsidy?" and a subjective and an endlessly debatable "How much subsidy is too much?"

CLASSIFICATION STANDARDS

The standards the author advocates in defense of the industry's present rate classification practices are equally applicable in condemning the industry's present underwriting practices, which are just as integral a part of the pricing system as rate classifications are.

For example, the author advocates that risk classifications be "well defined, . . . to ensure that each risk is actually placed in the right classification and to avoid unequal application of the classification system." (page 699) This is precisely what underwriting practices are not. It is unreasonable to advocate that rate classifications be well-defined when they are modified by an underwriting selection process which is not defined and which invites unequal application of the classification system.

To illustrate how an underwriting selection process could result in an unequal application of a classification system: a driver who meets the classification definition for a class offered by the Unfair Insurance Company, might be declined by the underwriter if he is from a minority racial group and accepted otherwise. If a reason for declination is requested, many reasons would be available, such as, "He parks his car on the street at night", even though the company may insure other drivers who park on the street at night.

The author's third standard (page 696) is: "The classes should be exhaustive and mutually exclusive; that is, an individual should belong to at least one, but only one class with respect to each rating variable." This flies in the face of the common practice of refusing to insure many applicants, meaning that many individuals find they do not belong to any class offered by the insurer. The author summarizes this standard (page 699) by saying, "'Exclusivity' precludes two different rates for the exact same risk." This flies in the face of the common practice of most insurance managements to have at least two insurers with different rates for the same risk. It is unreasonable to insist that a part of the pricing system should "be exhaustive and mutually exclusive" when the whole pricing system is neither. It seems that as we increase the clarity of the rate classes, we simultaneously increase the unclarity of the underwriting process so that the pricing system as a whole remains unchanged.

The author's quotation of SRI's statement: "The regulator's determination of what is unfairly discriminatory should relate only to the use of variables whose predictive validity cannot be substantiated and to unequal application of a classification system." (page 704) is applicable as a condemnation of the industry's present underwriting practices - using unsubstantiated variables to decline risks or to put them in a higher rated affiliate, and to unequally apply the classification system by declining or rating up some risks that meet the classification definition of the lower rated affiliate.

The author summarizes his seven standards "into three broader categories which can describe a set of necessary and sufficient conditions for insurance classifications: i.e., homogeneous, well-defined, and practical." (page 697) The author acknowledges that insurance classifications are not homogeneous. As discussed above, the pricing system as a whole is not well defined, opening the door to unequal application of it. In view of the insurance industry's pricing system being not homogeneous and not well-defined, for someone to then advocate that the only remaining standard should be "practical" (from the insurer's point of view) conveys the appearance of an industry that is insensitive to the public's perception of the industry's pricing system. The public is more concerned about whether the system is fair to the public then whether it is practical for the insurers. And the public perceives that it is not fair to the public because of the industry's failure to meet the two standards conceded by the author as "necessary" namely that the pricing system be homogeneous and well-defined. The author's paper, by concentrating on only a part of the pricing system, rate classifications, and ignoring an equally important part, underwriting selectivity, which negates the part he describes, illustrates why the public feels that insurance prices are unfair and why the industry fails to understand why the public feels that way.

ALTERNATIVE STANDARDS FOR CLASSIFICATIONS

The author touches upon what the reviewer regards as the single, fundamental standard for classification in a competitive market. "If, for example, it costs an insurer ten dollars on each policy for all to find only a small portion of risks who could save twenty dollars, it is not worth the effort." (page 699) This could be restated as, "A classification criterion is economic and appropriate if the reduction in expected losses for those who meet the criterion exceeds the cost of measuring the criterion for all who apply. The classification criterion which is most economic is the one that achieves the highest ratio of the reduction in expected losses for those who meet the criterion versus the cost of measuring the criterion for all who apply." This definition is the logical outcome of a competitive market. Under this definition, if sex is more correlated with expected losses and is cheaper to measure than, say, psychological attitudes, it is more economic in a competitive market as a classification criterion. Since no criterion is perfectly correlated with expected losses or is costless to measure, no classification criterion is perfect. It is a question of degree and relative efficiency and cost, questions that a private competitive market is ideally suited to determine.

As the author points out, some segments of the public criticize some classification criteria on the grounds that they are not perfect. Unfortunately, we do not have a choice between a perfect system and an imperfect system. Our choice is between a system controlled by profit motivated insurers competing against each other or a system controlled by government and the political process or some combination of the two. The results of our choice will affect the economic value of insurance to the many private property owners and small businesses who depend on insurance, and will affect in a small degree the economic ability of the United States as a whole to compete in the world market against other foreign economies.

Some segments of the public also criticize some classification criteria on the grounds that they are offensive, such as race, religion, occupation and income level, regardless of how economic they may be. The suspicion exists that insurers do use such criteria but attempt to conceal such use by not openly defining and disclosing all the criteria they use in their pricing systems, which embrace both rating and underwriting procedures. This suspicion is expressed in the accusation that the pricing system is "unfair". The suspicion will persist as long as the risk classification criteria are not "well-defined" and publicly disclosed.