IMPLICATIONS OF SAMPLING THEORY FOR PACKAGE POLICY RATEMAKING

JEFFREY T. LANGE

Following the introduction of the Homeowners' policy, interest began to develop in the problem of making rates for package policies. This interest was heightened by the introduction of the commercial package policies, which departed from the indivisible premium concept of the Homeowners' policy. These policies raised a question for ratemakers: should the experience data be collected so that a single rate might replace the separate rates for each coverage? Before answering this question, it was necessary to face the more basic question of how the experience developed for several different coverages might be combined for ratemaking. While actuaries were pondering these questions, still another problem arose. The experience data for the residual fire dwelling business—those risks not insured under the package policy—were found to be extremely adverse. Thus the ratemakers were forced to expand the package policy ratemaking problem to include ratemaking for residual business.

It is the purpose of this paper to discuss package policy ratemaking from the point of view of mathematical statistics, and in particular sampling theory. Two fundamental techniques which are widely used in sampling—stratification and ratio estimation—are discussed with emphasis on why these techniques produce more precise estimates than less sophisticated methods. These techniques are then applied to package policy ratemaking. The statistics are stratified by layer of coverage and the ratio of package and non-package pure premiums within each stratum is estimated. These ratios are applied to the underlying pure premiums, developed using combined package and non-package data, to obtain underlying pure premiums for each policy form and each coverage. For package policies, the underlying pure premiums (reflecting the appropriate ratios) may be combined and loaded for expenses to obtain an indivisible premium.

The essence of the method is that package policy experience will be subdivided by coverage for ratemaking, and will be used in combination with non-package experience in determining rate levels and rate relationships. Differentials will be computed for each coverage between package and non-package data to reflect the differences between these two classes of risks.

While the method is supported by certain principles drawn from sampling theory, which are explained in some detail, it also has practical advantages. Package rates would be adjusted even when the experience was still of small volume; later when non-package volume declines, the combined experience will still be sufficient to produce adequate rate adjustment for residual classes of business. The method also provides for an accurate computation of trend, credibility and loss development factors for package policies.

The method implies that essentially the same statistical plan be used for both package and non-package data and that uniform definitions be used for all coverages. It also implies that both sets of data become available at about the same time (and for the same group of companies). Furthermore, it represents a departure from the current procedures for rating package policies. Finally, it would appear to suggest that package and non-package rates for all coverages be developed simultaneously. The net result of these implications is a radical departure from current procedure.

While the method presented in this paper is illustrated by a detailed example and is described at some length, it should be clear that it is not presented as a solution to package policy ratemaking problems. The purpose of the paper is to discuss the implication of certain principles from sampling theory for ratemaking. The ratemaking method presented is only an example of what might be developed from these principles. As is pointed out in the paper, there are certain limitations to these principles, and their applicability in general to all package policy ratemaking is not completely clear.

Sampling Theory and Ratemaking

One might well question whether sampling theory has any applicability to the general ratemaking problem. The typical sampling problem is to estimate a certain population parameter based upon a random sample of n items drawn from the total population. The theory deals with the best ways to select the sample units, the methods of computing the estimate and the relative precision of the estimate. Few companies or rating bureaus rely upon samples in establishing overall rate level changes, and hence one might argue that sampling theory has little application to ratemaking.

Whether sampling theory has any relevance for ratemaking depends upon our view of ratemaking and the insurance mechanism. If ratemaking decisions are made after an analysis of the costs of doing business—the premiums, losses and expenses—and if these statistics are considered to be historical accountings of what actually happened, then sampling theory has no application to ratemaking. On the other hand, the insurance business may be regarded "as a continuous game of chance between the company on one side, and the totality of policyholders on the other. In the course of this game, certain random events known as *claims* occur from time to time, and have to be settled by the company, while on the other hand the company receives a continuous flow of risk premiums from the policyholders."[1] Thus, the relationship of claims or losses to premiums or exposures over a period of time may be described in terms of random sequences, which is to say, "the risk business of an insurance company forms a particular case of a stochastic process."[2]

In examining this latter view, it must be borne in mind that in an insurance contract the "insured is relieved of any concern, not only as to what is going to happen, but also as to what could happen but probably will not".[3] Thus, the losses which the insurer incurs during a given time period "never actually reflect the hazard covered, but are always an isolated sample of all the possible amounts of losses which might have been incurred".[4] Thus, insurance statistics may be viewed as samples of what might have occurred. In ratemaking, these samples are used to make projections of what will occur in the future, and it is important to note that these samples will be subject to sampling variation due to pure chance fluctuation.

If ratemaking statistics are samples, then sampling theory has a great deal of significance for ratemaking. One goal of ratemaking should be to produce estimates which minimize sampling variation. In this paper, certain sampling techniques, which are utilized to reduce the variance of estimates, are examined and their implications for ratemaking are explored. In general, such techniques might be divided into two broad classifications. One class would include those techniques which present more sophisticated ways of drawing the sample—i.e. that deal with sample design. In this class fall stratification, sub-sampling, cluster sampling, etc. The other class of techniques would encompass those that present improved methods of making an estimate from the data once it has been collected. In this latter category are ratio estimates, regression estimates, etc.

Stratification

In 1926, A. L. Bowley in his paper "Measurement of the Precision Attained in Sampling"[5] pointed out that the precision of estimates can be

[1] Cramer, H., "Collective Risk Theory: A Survey from the Point of View of the Theory of Stochastic Processing" (Esselte Reklam, Stockholm 1955), p. 5.

[2] Ibid., p. 7.

[3] Bailey, A., "Sampling Theory in Casualty Insurance," *PCAS Vol. XXXIX*, p. 50.

[4] Ibid., p. 50.

[5] Bowley, A., "Measurement of the Precision Attained in Sampling," *Bulletin of the International Statistical Institute (BISI) Vol. XXII.*

improved by taking a sample which is representative of the population. In particular, a more precise estimate can often be obtained economically by stratified random sampling. When stratified random sampling is used, the population is divided into several strata (mutually exclusive subgroups of the population), the sample is apportioned in some prearranged way among the strata, and the sampling performed at random from each stratum. In apportioning the sample among the strata, Bowley suggested that the number of sample units selected in a strata should be proportional to the number of units of the population in the strata: "proportional allocation."

Most authors agree that stratification nearly always results in a smaller variance for the estimated mean than is given by a comparable simple random sample. In fact, there will be a reduction in variance if the population can be subdivided into strata which are somewhat more homogeneous than the total population. The variance is reduced by the weighted average of the squared differences of the strata means and the grand mean.[6]

Variance for stratified sampling $=$ variance $- \sum N_h (\bar{Y}_h - \bar{Y})^2/nN$
where $\bar{Y}_h$ is the mean for a strata
- $\bar{Y}$ is the grand mean
- N_h is the number of units in a strata
- N is the total number of units in population
- n is the number of sample units
- "variance" is the variance of a simple random sample

As a result, the greater the difference between the individual strata (i.e. the more homogeneous each strata), the greater the improvement due to stratification. This arises from the nature of the variance itself. In simple random sampling, the variance is computed by squaring the difference between each sample item and the grand mean, not the mean of the strata as in stratified sampling. Thus, the reduction in variance arises from the fact that the individual item within each stratum is closer to the average value for the stratum than to the average of all strata.

Neyman[7] presented an alternate method of allocation in which the sample size within the strata is proportional to both the number of units and the standard deviation within the strata: "Neyman allocation" or

---

[6] Derivation of the formula is given by Cochran, W. *Sampling Techniques* (Second Edition) (John Wiley & Sons, Inc., New York, 1963) p. 98.

[7] Neyman, J., "On the Two Different Aspects of the Representative Method: the Method of Stratified Sampling and the Method of Purposive Selection," *Journal of the Royal Statistical Society. Vol. XCVII*, p. 558.

"optimum allocation." If the variances for individual strata differ significantly from the variance of the whole sample, then by making the sample size within each strata proportional to its variance, a reduction in the overall variance proportional to the average differences in variance is possible. In other words, more information should be used in making estimates when the data exhibits greater variability. Both Hurley and Mayerson[8] arrived at a similar conclusion (for a different reason), when they examine the need for different credibility criteria for different classifications of risks.

Neyman proved that for infinite populations the variance of the sample mean for proportional allocation was always less than or equal to that for simple random sampling, and that the variance of the mean for Neymen allocation is less than or equal to that for proportional allocation. Armitage[9] extended Neyman's results to finite populations, and found that in general the results do not hold. In fact, if the means within each strata are equal, then the variance of the mean under proportional allocation is greater than that under simple random sampling. If in addition the standard deviations within each strata are equal, then variance of the mean under Neyman allocation is greater than that under simple random sampling. Thus, in the case of small samples stratification will improve precision only if the resulting strata are more homogeneous than the total population.[10]

Stratification by Coverage and Layer of Insurance

It would appear that by dividing loss statistics based upon coverage, and into layers within those coverages, the resulting strata would each be more homogeneous than the total sample. The distributions of claims and of losses by size of claim show considerable variation by line of insurance. It seems unnecessary to discuss at length differences in loss distributions between fire insurance and liability insurance, or between windstorm insurance and theft insurance. Similarly, it is generally accepted that in ratemaking estimates may be improved by giving separate consideration to various layers of insurance.[11] This is another use of stratification, and

---

[8] Hurley, R. "A Credibility Framework for Gauging Fire Classification Experience," *PCAS Vol. XLI* p. 161 and Mayerson, A. "A Bayesian View of Credibility," *PCAS Vol. LI*, p. 85.

[9] Armitage, P., "A Comparison of Stratified with Unrestricted Random Sampling from a Finite Population," *Biometrika Vol. XXXIV*, p. 273.

[10] A discussion of whether stratification will yield an improvement in precision when sampling from finite populations is given by Evans, W., "On Stratification and Optimum Allocation," *Journal of the American Statistical Association (JASA) Vol. XLVI*, p. 95.

[11] Salzmann, R., "Rating by Layer of Insurance," *PCAS Vol. L*, p. 15.

should result in improved estimates since the loss distributions for excess insurance differ substantially from those for basic coverages. Since strata by coverage and layer would exhibit different means (and variances), a considerable improvement in precision would result from making separate estimates within each strata, and then combining the estimates, as opposed to simply combining the data.

In addition to minimizing chance variation, stratification by coverage and layer of insurance would permit the application of different credibility procedures to different coverages and layers of coverage. From an actuarial standpoint, this would be more accurate than applying a single credibility factor to the overall result.[12]

Stratification by coverage would also permit the application of trend factors especially suited to each coverage, rather than an average trend factor. Bodily injury liability trends are certainly influenced by many factors (hospital costs, jury verdicts) which have little significance for fire insurance. Similarly, rising crime rates, while significant for theft insurance, have little relevance for windstorm insurance. Loss development factors, which measure the changes in the aggregate dollar losses for an accident year as reserves mature, are also probably best measured by coverage, rather than for all coverages combined.

It would seem that from an actuarial standpoint, the number of years of data to be used in ratemaking, the calculation of credibilities, the measurement of trends, and the computation of loss development factors might all best be considered independently by coverage. Furthermore, from a statistical viewpoint, the analysis of package policy statistics by coverage and layer of coverage, i.e. by strata, would serve to reduce the effect of chance variation and to increase the precision of the estimates.

Ratio Estimates

Ratio estimates, although biased, have been frequently used in applied statistical work for more than a quarter century. The Bureau of the Census, for example, has for many years produced annual estimates of items included in the decennial census by the use of sample surveys incorporating ratio estimate. In fact, the use of ratio estimation in large scale

[12] Hurley and Mayerson. It might also be noted that stratification by coverage parallels the subdivision of Workmen's Compensation data into three categories (serious, non-serious and medical) and the use of different credibility factors for each category. A discussion of the decrease in relative credibility which results from the combination of non-homogeneous data is given by L. H. Longley-Cook, "Underwriting Profit in Fire Bureau Rates," *PCAS Vol. LIII*, this issue.

sample surveys predates the theoretical examination of ratio estimates. In the application of these techniques, it was customary to note that ratio estimates are biased and to add the opinion that the bias "is usually negligible"; however, no support, mathematical or empirical, was offered for the latter statement.[13]

In order to apply any of the ratio estimation techniques each sample observation must consist of two points: an observation of the variable under study and of an auxiliary variable. The auxiliary variate is simply some item which is closely correlated with the variable to be studied. In producing the annual updating of the decennial census, the Bureau of the Census usually uses the values obtained at the time of the last complete census as the auxiliary variable. In its survey, the Bureau samples not only the current value, but also the value at the time of the last census. The ratio of the current value to the value at the time of the last census is estimated for the sample, and this ratio is applied to the total obtained in the last census to produce the estimate of this year's value.[14]

During the early 1950's, several statisticians became interested in examining the bias of the ratio estimate and its relative efficiency when compared with simple expansion. J. C. Koop[15] obtained an expression for the bias of a simple ratio estimate, and explored the possibilities of reducing the bias. When analytic expressions for the bias were developed, it became possible to evaluate the various ways of computing ratio estimates and to develop ratio estimates which were unbiased.[16]

Since the ratio estimate may be biased, one may question whether or not it is worth trying. There are two reasons for exploring its use. First, it is possible to compute unbiased ratio estimates or to compute biased ratio estimates and then estimate their bias. Thus, in practice, it is unnecessary to use a ratio estimate which is significantly biased, since if it is biased one has the option of using an unbiased ratio estimate. Second, whether or not there will be an improvement as a result of using ratio estimates can usually be estimated fairly easily.

---

[13] Hansen, M., Hurwitz, W. and Gurney, M., "Problems and Methods of a Sample Survey of Business," *JASA Vol. XLI*, p. 173 and Hurwitz, W. and Hansen, M., "On the Theory of Sampling from Finite Populations," *Annals of Mathematical Statistics Vol. XIV*, p. 333.

[14] In actual practice, the sample design is more complicated and varies according to the item sampled. For an example see Hansen, Hurwitz and Gurney.

[15] Koop, J., "A Note on the Bias of the Ratio Estimate," *BISI Vol. XXXIII Part II*, p. 141.

[16] Hartley, H. and Ross, A., "Unbiased Ratio Estimates," *Nature Vol. CLXXIV*, p. 270.

Ratio estimates produce their most significant improvement over the other forms of estimation if the regression of the variable under study and the auxiliary variable is a straight line through the origin. In this case, ratio estimates are unbiased. If the relationship of the two variables is approximately linear, then ratio estimates are probably more precise. Stated more analytically,[17] ratio estimates are better (in the sense of having lower variance) than estimates based upon simple expansion (non-regression estimates) if the correlation of the variable under study and the auxiliary random variable is greater than one-half the ratio of the coefficients of variation of the auxiliary random variable to the variable under study:

$$\text{correlation} > \tfrac{1}{2}\,\frac{\text{coefficient of variation of auxiliary variable}}{\text{coefficient of variation of variable to be studied}}$$

If for example the relative amount of variation of both variables is equal, then the ratio estimate will result in a lower variance (and an improvement in precision) if the correlation exceeds .5. If the auxiliary random variable has less variation than the variable under study (i.e. if it is the result of a larger sample), then an even lower correlation is sufficient for a reduction in variance.

The use of stratification coupled with ratio estimation has been quite widespread in sampling problems. Published comparisons[18] of the application of these techniques versus less sophisticated methods have shown that the variance may be reduced by as much as 50% to 95%. This dramatic improvement in precision is equivalent to radically increasing the sample size at no additional cost.

Ratio of Package to Non-Package Experience

When a package policy is first introduced its rates are generally constructed from the non-package rates for component coverages with appropriate discounts. These non-package rates are the result of many rate revisions and can be thought of as relatively accurate, time-tested, known values in comparison with the package rates constructed from them using judgment discounts. During the first few years of the package policy's operation, the volume of statistics developed will probably be much smaller than the non-package experience, and certainly smaller than the sum total of the experience which over the years went into the development of the non-package rates. The preliminary package policy data may be thought

---

[17] Derivation is given in Cochran, p. 165.

[18] Ibid., p. 179.

of as a sample which will be used to estimate package policy rates. The non-package rates, as modified by the latest available statistics, may be considered auxiliary variables which can be utilized in the estimation of package rates. For a given coverage, the correlation between package and non-package statistics should be fairly high, especially in the early stages, since much of the package business will represent simply a transfer from the non-package policies.

One might also expect that the package and non-package statistics would exhibit approximately the same amout of relative variation. Perhaps due to smaller volumes, the package policy data might exhibit greater variation, but this may be offset to some extent by the greater homogeneity of the population of risks written under the package policy. If the package policy data exhibits as much, or more, variation than the non-package data, and if correlation between the two sets of data is relatively high (greater than .5), then by analogy to sampling theory a gain in precision should be achieved by the use of some form of ratio estimation.

The use of ratio estimates implies that the ratio of the variable under study to an auxiliary variable for the sample is measured, and that this ratio is applied to the auxiliary variable population value to obtain our estimate of the population value for the variable under study. Applying this to package policy ratemaking, the average ratio (by class and territory) of package to non-package pure premiums (or the ratio of the averages) might be applied to the non-package underlying pure premiums to obtain package underlying pure premiums for the coverage. These package underlying pure premiums for each coverage might be added together to obtain the pure premium underlying the indivisible premium for the package policy. Presumably, the non-package rates would reflect the rate level indications of the latest experience and trend data, and also the class and territory rate relationships established from several years of data. By using the ratio estimate technique, this body of statistical information would be incorporated into the package policy rates, while simultaneously reflecting the relationship of package and non-package experience indicated by the available statistical data. The ratio technique would thus make use of all of the available statistical information.

Eventually, the volume of data developed under the package policy may exceed that developed under the non-package policies—the residual problem. Here a ratio estimate technique might be employed, using the package policy underlying pure premiums as the auxiliary variable in setting non-package rates. However, the use of ratio estimates would cease to yield much advantage when correlation between package and non-package business declines.

Example

The use of stratification and ratio estimation in ratemaking would be best illustrated by an example using actual package policy data. Unfortunately, statistics are not collected in a manner which permits an application of the method to a broad package policy. As an illustration of how the method might be applied, the Special Automobile Package Policy (SAP) was selected because statistics for that package policy and for its component coverages when purchased separately are collected under the same statistical plan and are almost comparable.

The SAP consists of a liability package with an indivisible premium and of a physical damage package. For this illustration the liability package was selected.

The basic ingredients of the liability package are bodily injury (B.I.) and property damage (P.D.) liability insurance at a $25,000 single limit[19], medical expense coverage limited to $1000 per person, uninsured motorist coverage with limits equal to the financial responsibility limits in the state, and accidental death coverage with a $1000 limit. An increased single limit of liability and increased medical expense coverage are available for an additional premium charge.

Currently, automobile liability ratemaking for non-package policies (Family Auto Policy—FAP) would treat each of these coverages independently. In order to illustrate how ratio estimates and stratification might be applied to a more sophisticated package incorporating both property and casualty coverages, the coverages will be grouped in three subdivisions representing three different approaches to ratemaking. The first will include the basic limits ($10,000/$20,000 B.I. and $5,000 P.D.) liability coverages, for which a rather sophisticated, formula ratemaking technique has been developed for FAP rates.[20] Since this approach utilizes exposures in computing premiums at present rates it is sometimes referred to as a "modified pure premium approach" and will serve as an example of casualty ratemaking procedures. The second group of coverages includes medical expense coverage and uninsured motorists coverage. The ratemaking techniques currently used for these lines may be taken as an illus-

---

[19] Limits of $15,000 for liability and $500 for medical expense are available in a few states.

[20] Stern, P., "Ratemaking Procedures for Automobile Liability Insurance," *PCAS LII*, p. 139.

tration of property insurance ratemaking in that they resemble the loss ratio approach as outlined by C. A. Kulp.[21] Finally, there are a collection of excess coverages, some mandatory (liability above 10/20/5 and accidental death coverage), some optional (liability in excess of the $25,000 single limit). Such low premium volume coverages, some with high possible single losses, will probably be found in most packages.

Since premiums at present rates are used in ratemaking for the FAP, it was necessary to subdivide the SAP indivisible premium into its components by coverage. This was accomplished by taking the original formula for computing SAP rates from the non-package rates, substituting the present FAP base rates for the original FAP rates, the present SAP base rate for the original SAP rate, and solving for the package discount.

$$\text{SAP Semi-annual Rate} = .5d[1.0750(\text{BI Rate}) + 1.0368(\text{PD Rate}) + .50(\text{Med. Pay. Rate}) + .50(\text{UM Rate})]$$

where d = complement of package discount expressed as a decimal and where (____ Rate) designates the corresponding annual FAP rate

This package discount times the present FAP 10/20 B.I. base rate related to the present SAP base rate is the percentage of the SAP premium at present rates in a given territory which should be allocated to 10/20 B.I. For example, in territory 01 where the FAP B.I. rate is $62, the SAP rate is $44 and the complement of the package discount was found to be .84, the percentage of SAP premium which should be allocated to bodily injury coverage is 59.2%:

$$\frac{.5d\,(\text{B.I. Rate})}{\text{SAP Rate}} = \frac{.5(.84)(\$62)}{\$44} = 59.2\%$$

This same procedure was applied to the other coverages.

The SAP premium at present rates for each territory could be added to the corresponding FAP premium and the sum could be incorporated in the standard ratemaking procedures for each subline. From the idenitfication of SAP losses by cause of loss, it is possible to obtain SAP losses for a given layer of coverage. The losses may be added to the corresponding FAP losses, and statewide rate changes and territory rates may then be computed using combined package and non-package data and following standard formulas. This has been illustrated with bodily injury liability data on Tables 2 and 3. For the basic limits coverages, the ratio of package to non-package data has been computed by dividing the SAP

[21] Kulp, C., "The Ratemaking Process in Property and Casualty Insurance—Goals, Techniques, and Limits," *Law and Contemporary Problems Vol. 15*, p. 493.

pure premium by the average FAP pure premium, which was computed by taking the weighted average of FAP pure premiums by class and territory utilizing the SAP exposures as weights.[22] (See Table 1.)

The Automobile Statistical Plan[23] does not require the reporting of exposures for medical payments coverage (the FAP equivalent of medical expense) or uninsured motorist coverage. SAP premiums and losses could be obtained for these two coverages as described in the preceding paragraph, and thus SAP statistics could be incorporated in the loss ratio analyses usually followed in setting rates for these sublines. In addition, the SAP premiums could be adjusted to the FAP level by dividing by the discount assumed in the allocation of SAP premiums by coverage. This would permit a comparison of SAP and FAP loss ratios so that indicated package discounts might be computed. (See Table 1.)

Several layers of coverage remain for consideration: excess B.I. and P.D. liability and accidental death coverage. Excess coverages are not normally rated on a state by state basis, so the experience for these coverages might be combined on a countrywide basis. Presumably, the SAP excess data could be reviewed simultaneously with non-package data and modifications of the existing charges made at that time. For our example, it has been assumed no modification of the existing charges for limits of coverage in excess of 10/20/5 is to be made.

The calculation of an SAP indivisible premium is shown for Territory 01. The proposed FAP rates (developed utilizing combined SAP and FAP data) are converted to underlying pure premiums and these underlyings are increased to the SAP limits of liability using the standard FAP factors for a $25,000 single limit, since no change in excess charges has been assumed. The ratio of SAP to FAP experience for each coverage is applied to the underlying for that coverage. The resulting underlying pure premiums by coverage were added together, multiplied by .5 to con-

[22] For this example, it was necessary to estimate the ratio by taking the ratio of the averages; however, a more accurate result might have been obtained by averaging the ratios of the SAP pure premium to the FAP pure premiums for each class and territory and then correcting this average ratio for the bias. See Hartley and Ross. In order to simplify the example, credibility factors have not been applied to the ratios.

[23] *Automobile Statistical Plan*, National Bureau of Casualty Underwriters, 1966.

Table 1

ESTIMATION OF RATIOS

Coverage	Average Pure Premium S.A.P.	F.A.P. *(a)*	Ratio of S.A.P. to F.A.P.
B. I. 10/20	$36.39	$39.13	.93
P.D. $5,000	14.48	15.40	.94

Coverage	Complement of Discount Assumed in Splitting Premium *(b)*	S.A.P. Loss Ratio On S.A.P. Level	On F.A.P. Level	F.A.P. Loss Ratio	Ratio of Adjusted F.A.P. to F.A.P.
(1)	(2)	(3)	(4) = (3) × (2)	(5)	(4) ÷ (5)
Medical	.41	.820	.336	.703	.48
Uninsured Motorist	.41	.939	.385	.544	.71

Coverage	S.A.P. Premium	S.A.P. Losses	Loss Ratio
Excess liability and accidental death coverage			
In standard limits package	$119,599 *(c)*	$142,140	1.188
Excess over standard limits	290,170 *(d)*	228,104	.786
Subtotal	$409,769	$370,244	.904

NOTES:

(a) F.A.P. pure premiums by class and territory were averaged using the S.A.P. exposures as weights.

(b) The statewide average complement of the package discount was found to be .82. The complement of package discount times .50 yields .41. (A 50% additional discount for medical and uninsured motorist coverages was included in the original formula.)

(c) Computed by applying the increased limits factor (minus unity) to the basic limits premium at present rates.

(d) Computed by applying the average S.A.P. additional charge to the S.A.P. premium at present rates for standard limits.

Table 2

AUTOMOBILE LIABILITY INSURANCE – PRIVATE PASSENGER NON-FLEET

Development of Statewide Rate Level Changes

Coverage	(2) Accident Year	(3) 10/20/5 Limits Earned Premium at Present Collectible Level	(4) 10/20/5 Limits Incurred Losses	(5) Number of Claims	(6) Loss & Loss Adjustment Ratio at Present Rates (4) ÷ (3)	(7) Accident Year Weights	(8) Weighted Loss & Loss Adjustment Ratio at Present Rates
B. I.	1963	$9,434,132	$6,689,194	5,872	.709	15%	.786
	1964	9,723,912	7,767,803	6,318	.799	85%	
P.D.	1963	$4,421,474	$2,892,330	17,033	.654	15%	.678
	1964	4,560,564	3,109,779	17,582	.682	85%	

Coverage (9)	(10) Factor to Adjust Losses	(11) Rate Level Loss Ratio (8) × (10)	(12) Expected Loss & Loss Adjustment Ratio	(13) Credibility	(14) Indicated Rate Level Change $\left[\frac{(11)}{(12)} - 1.0\right] \times (13)$
B. I.	1.000	.786	.662	1.000	+18.7%
P.D.	1.173	.795	.662	1.000	+20.1%

NOTES:

This table corresponds to Exhibit 7, page 178, of Stern's "Ratemaking Procedures for Automobile Liability Insurance" *(PCAS Vol. LII)*. An explanation of the terms used in the exhibit and of the derivation of the values shown in each column is set forth on pages 176-183 of Stern's paper. Modifications of Stern's example (in addition to the substitution of a different set of data) are discussed below.

Column (3) FAP and SAP earned premiums at present collectible rates were computed as described by Stern. SAP premiums were subdivided by coverage as explained previously, and then added to the FAP premiums. For 1964, the subdivision of premiums by policy form is shown below:

B.I.	1964	FAP	$8,430,213
		SAP	$1,293,699
		Total	$9,723,912

Columns (4) and (5) SAP losses for each coverage were identified by cause of loss coding. Both FAP and SAP losses were then limited and adjusted as outlined by Stern. For 1964, the subdivision of bodily injury losses by policy form is shown below:

			Column (4)	Column (5)
B.I.	1964	FAP	$6,542,253	5,325
		SAP	$1,225,550	993
		Total	$7,767,803	6,318

Column (12) was obtained by taking the weighted average of the FAP and SAP expected loss ratios. For this example, expected loss ratios of .655 and .705 respectively have been assumed.

Column (14) sets forth the combined rate change. Since the proposed differential between the two policy forms will differ from the present differential, there will be different rate changes for each policy form. For bodily injury coverage, the present package discount is .82; the comparable package discount resulting from the indications on Table 1 and the assumed difference in expense ratios is (.93) (.655 ÷ .705) or .86. By applying a rate change of +17.9% for the FAP and utilizing the .93 ratio and the .705 expected loss ratio in computing SAP rates, a 23.7% rate change (1.179) (.86 ÷ .82) is achieved for the SAP. The average of the SAP and FAP rate changes would be 18.7%.

Table 3

AUTOMOBILE LIABILITY INSURANCE – PRIVATE PASSENGER CARS

Development of Proposed Rate Level Changes by Territory

Bodily Injury – 10/20 Limits

Territory	(2) Accident Year 1964 Earned Number of Cars	(3) Present Average Rate	(4) Accident Years 1962 – 1964: 10/20 Limits Pure Premium (Incl. All Loss Adj.)	(5) Accident Years 1962 – 1964: Loss and Loss Adjustment Ratio at Present Rates (4) ÷ (3)	(6) Accident Years 1962 – 1964: Credibility	(7) Formula Loss & Loss Adjustment Ratio at Present Rates [(5) × (6)] + [1.0 – (6)] × .773	(8) Col. (7) as Ratio to Statewide Average	(9) Territorial Rate Level Change [(8) × 1.179] – 1.0	(10) Average of Present Differentials to Rate Class 1A	(11) Proposed Class 1A Rate (3) × [1.0 + (9)] ÷ (10)
					BODILY INJURY					
01	56,383	60.15	45.42	.755	1.00	.755	.974	+14.8%	.970	71
02	39,920	55.22	44.91	.813	1.00	.813	1.049	+23.7	.969	70
03	10,082	44.68	30.23	.677	.80	.696	.898	+ 5.9	.951	50
•										
•										
•										
Total	202,944	47.92	37.02	.773		.775				

NOTES: This table corresponds to Exhibit 8, page 185, of Stern's "Ratemaking Procedures for Automobile Liability Insurance" (*PCAS Vol. LII*), and an explanation of the exhibit appears on pages 183 through 187 of that paper. In addition to the use of combined SAP and FAP data, the following should be noted.

Column (3) was obtained by adding to the FAP premium at present rates, the SAP premium at present rates (apportioned by coverage as on Tables 1 and 2) and dividing by the combined FAP and SAP exposures.

Column (4) was obtained by limiting the SAP bodily injury losses (obtained from cause of loss coding) to 10/20 and adding them to the FAP losses. The result was divided by the combined exposures to obtain the pure premium.

Column (5) was obtained using the standard credibility table (Stern, page 166) and combined number of claims. Had FAP experience been used alone, credibilities would have been up to .10 lower. If SAP B.I. data were used alone, credibilities would have been .20 to .40 lower.

Column (8) did not differ significantly from the comparable values for the FAP policy alone: the maximum difference was 5%. However, when SAP B.I. data were used alone, there were substantial differences between the resulting ratios to the average and those shown in column (8) for low volume territories.

vert them to a semi-annual basis, and divided by the SAP expected loss ratio[24] to obtain the SAP indivisible premium.

(1) Coverage and Limits	(2) Proposed FAP Rate	(3) Underlying Pure Premium .655 x Rate	(4) Increased Limits Factor	(5) Ratio from Table 1	(6) SAP Underlying (3)(4)(5)
10/20 B.I.	$71.00	$46.51	1.0750	.93	$46.51
5,000 P.D.	34.00	22.27	1.0368	.94	21.71
$1,000 Med Pay	13.00	8.52	1.000	.48	4.09
UM	5.00	3.28	1.000	.71	2.33
					$74.64

SAP Semi-Annual Rate = $.5(74.64 \div .705) = \$53$

Discussion

There are a number of similarities between the preceding example and the "component method" of ratemaking outlined by Bailey, Hobbs, Hunt and Salzmann in "Commercial Package Policies—Rating and Statistics."[25] They rejected the component method in favor of the "indivisible premium method." The main feature of the latter was that statistics would be analyzed "by type of insured, according to the combination of coverages selected."[26] Since their "Model Statistical Plan" provided for the recording of exposures and for cause of loss coding,[27] it would be possible to superimpose stratification and ratio estimation on the authors' indivisible premium ratemaking procedures. The added refinement of stratification-ratio estimation will produce more meaningful and useful results in each of the four areas where the indivisible premium approach was shown by the authors to be most efficient.

The first area had to do with the philosophy of package policies, and in particular with the concept that perils insured against is a valid basis for classification. A corollary is that package loss costs for a particular insured (type of insured) might not equal the sum of the loss costs for the coverages rated individually for all insureds. Since the use of stratification and ratio estimation does not call for the combination of the ex-

[24] To illustrate how package and non-package data might be combined even if the expense provisions were different for each type of policy, a .655 expected loss ratio has been assumed for the FAP and a .705 expected loss ratio for the SAP.

[25] Bailey, R., Hobbs, E., Hunt, F. and Salzmann, R., "Commercial Package Policies—Rating and Statistics," *PCAS Vol. L*, p. 87.

[26] Ibid., p. 92.

[27] Ibid., p. 97.

perience of all insureds (as does the authors' component approach), the truth of the packaging principle could be tested. While under the indivisible premium approach it would be tested only for all perils combined, by using stratification-ratio estimation one could determine which perils produced the saving. In the SAP example in the previous section, it was found that for certain coverages, e.g. medical expense, the savings were much more significant than for other coverages (52% versus 6%). The stratification-ratio estimation approach would yield more information in testing the packaging principle, and the results of such analysis would be of greater significance in planning future packages and redesigning existing packages because they would pinpoint areas where the greatest savings were achieved. Considering the Homeowners' policy as an example of indivisible premium rate-making, it is interesting to note that although it is clear that the packaging principle was true—i.e. package loss costs were less than the sum of the individual coverages—it is not possible to determine how much burglary loss costs were reduced by making this coverage mandatory, or whether there was any reduction in windstorm loss costs, etc. Such information might have been of value in modifying the Homeowners' package or in designing new package policies. It would not emerge from an indivisible premium method although it would be routinely produced by a method employing stratification-ratio estimation.

The second area was the screening and reducing of the number of different plans available, which would be accomplished by collecting data by combination coverage. By employing stratification-ratio estimation one might determine which combinations produce no packaging savings. Coverage combination purchased by the insured could be considered another form of risk classification which is superimposed over the existing classification plan. One could more precisely pinpoint the ineffective package combinations by isolating exactly where (for what coverages) the combination produced savings and the magnitude of these savings. For example, one might find that the addition of a certain coverage to a package did not produce any reduction in pure premium for that coverage, nor did it change the results for any other coverage. From this, one might conclude that the combination including that coverage on a mandatory basis should be eliminated.

The third area was the elimination of complications caused by duplication of coverage between endorsements and the basic policy. Once again the same arguments in favor of stratification and ratio estimation may be advanced. Providing the statistical plan is set up so that coding is carried out by risk, then an analysis by coverage has all the advantages of the indivisible premium analysis. In each of these areas the advantages claimed for the indivisible premium method arise from the way the statistics are maintained—"the Indivisible Premium Statistical Plan"—not from method of analysis. Given the excellent statistical plan described in "Commercial Package Policies—Rating and Statistics," the use of stratification-ratio estimation will generally produce more meaningful results than the indivisible premium method.

The fourth area was the primary one—the coding of the data. "With the indivisible premium approach, experience would be coded *by policy;* whereas experience would be collected *by coverage* under the component rating method."[28] A method involving coding by policy will produce the advantages discussed under areas one, two and three. Without such coding, a method is deficient in all areas. The use of stratification and ratio estimation offers no obstacle to coding by policy, as long as coverages purchased are identified and cause of loss is identified. The example of stratification and ratio estimation presented in the previous section was based on statistics for an indivisible premium package collected by policy, not by coverage.

Two points seem evident from this discussion of "Commercial Package Policies—Rating and Statistics." The first involves stratification-ratio estimation while the second deals with statistics. First, stratification and ratio estimation would yield more valuable information for the design and analysis of package policies than would either the indivisible premium approach, or a feedback of statistics into the basic coverages. This advantage is in addition to the greater precision gained by the use of actuarial procedures suited to each strata (coverage-layer) and the possible advantages from utilizing the ratios of package and non-package data. Second, the key to package policy ratemaking is the statistical plan. Stratification and ratio estimation yielded more information than the indivisible premium method when the "Indivisible Premium Statistical Plan" was used. Both methods owe most of their advantages to the statistical plan assumed by their authors. Each method assumes a statistical plan which is significantly different from the current methods of coding commercial package policy data. While a statistical plan as advanced as the "Indivisible Premium Statistical Plan" is not necessary for the use of stratification and ratio estimation, it is necessary that certain features be incorporated in the statistical plan if these methods are to be used. Among the desirable features are uniform definitions and methods of compiling data by package and

[28] Ibid., p. 94.

coverage, uniform classification and territory definitions, some uniformity in exposure bases, identification of coverages purchased, and cause-of-loss coding. It should be obvious that the existing statistical plans by line of insurance cannot be stapled together and put in a package policy binder. As a corollary to this, all of the detailed coding by line of insurance cannot be preserved in the commercial lines plan.

In the discussion of ratio estimates, it has been assumed that the ratio of package and non-package data will be used if ratio estimates are appropriate at all. In sampling, the denominator of the ratio is usually some auxiliary variable which exhibits less variation than the variable under study and which is based on a broader sample. In our example in which SAP volume was much smaller than FAP volume and FAP rates were the result of many years of experience, the FAP data provided such a base. Turning to Homeowners', it is obvious that the residual fire, burglary, and comprehensive personal liability lines would not provide such a base. The problem of a proper denominator for a ratio estimate will have to be decided individually for each problem to which ratio estimates are to be applied.

In the commercial lines field, one possibility is that a statistical organization might combine the data for the various packages with the non-package data and develop pure premiums by coverage (and layer of insurance), by class, and by territory. These industry-wide pure premiums could be used by companies and rating bureaus as a standard of comparison, or as the denominator in their ratio estimates. In that way an individual package policy could be compared coverage by coverage to the total business, and the company or bureau could establish the savings achieved due to packaging together a particular combination of coverages. Presumably, manual rates for non-package business could be computed by utilizing ratios of non-package pure premiums to the average, resulting in ratios in excess of unity (a non-package surcharge). Thus, a broad statistical basis would be obtained for class and territory relativities, and for analysis of varying package savings which resulted from the coverage combinations in different packages.

Conclusion

Stratification and ratio estimation could be used in package policy ratemaking to produce more accurate results and more meaningful statistics for the evaluation of package policies. The degree of increased accuracy and the utility of the additional information produced by these techniques could probably be evaluated only by empirical studies. It would appear that these techniques have sufficient theoretical support to merit such empirical investigations. To accomplish this would require changes in the method of compiling statistics for package policies. Since similar changes[29] are being considered for other reasons, it is possible that these techniques might be experimented with in the commercial risk area.

---

[29] Simon, L., "Statistical Support for Adequate Rates," *Best's Insurance News* (Fire & Casualty Edition) Vol. 67 (No. 3), p. 10.